Biology I

Lumen Learning

Developed in conjunction with Northern Virginia Community College's Dr. David Fernandez, Dr. Leslie Orzetti, and Dr. Paula Rodgers.

Except where expressly noted otherwise, the contents of this course are based on materials originally published by OpenStax College under a Creative Commons Attribution License CC BY 4.0. Download the original material for free at http://cnx.org/content/col11407/latest. If you believe this page violates your copyright, please contact us.

Cover photo: "Gulf of Eilat (Red Sea) coral reefs" by Daviddarom. Public Domain.

Biology I by Lumen Learning is licensed under a Creative Commons Attribution 4.0 International License, except where otherwise noted.

CONTENTS

THE PROCESS OF SCIENCE

AN INTRODUCTION TO BIOLOGY

Figure 1.1. This NASA image is a composite of several satellite-based views of Earth. To make the whole-Earth image, NASA scientists combine observations of different parts of the planet. (credit: modification of work by NASA)

Viewed from space, Earth (Figure 1.1) offers few clues about the diversity of life forms that reside there. The first forms of life on Earth are thought to have been microorganisms that existed for billions of years before plants and animals appeared. The mammals, birds, and flowers so familiar to us are all relatively recent, originating 130 to 200 million years ago. Humans have inhabited this planet for only the last 2.5 million years, and only in the last 200,000 years have humans started looking like we do today.

Licensing & Attributions

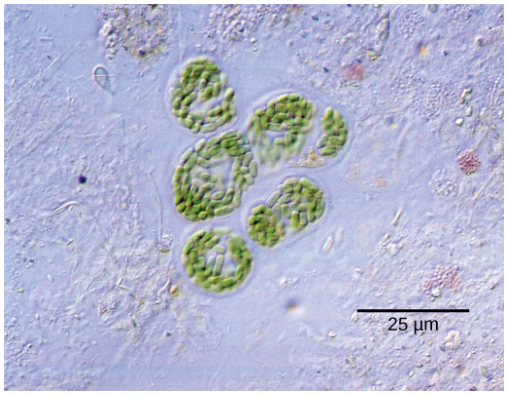
CC licensed content, Shared previously

- An Introduction to Biology. **Authored by:** OpenStax College. **Located at:** http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@8.10:1/Concepts_of_Biology. **License:** *CC BY: Attribution*

THE SCIENCE OF BIOLOGY

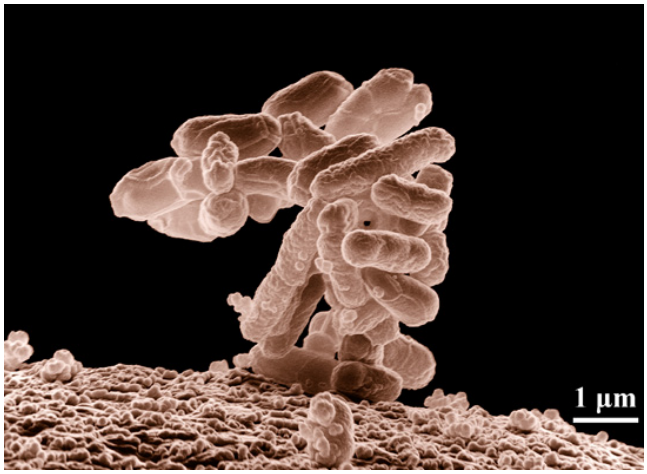
1.1 The Science of Biology

Summary



Formerly called blue-green algae, these (a) cyanobacteria, shown here at 300x magnification under a light microscope, are some of Earth's oldest life forms. These (b) stromatolites along the shores of Lake Thetis in Western Australia are ancient structures formed by the layering of cyanobacteria in shallow waters. (credit a: modification of work by NASA; credit b: modification of work by Ruth Ellison; scale-bar data from Matt Russell)

What is biology? In simple terms, biology is the study of living organisms and their interactions with one another and their environments. This is a very broad definition because the scope of biology is vast. Biologists may study anything from the microscopic or submicroscopic view of a cell to ecosystems and the whole living planet (Figure). Listening to the daily news, you will quickly realize how many aspects of biology are discussed every day. For example, recent news topics include *Escherichia coli*(Figure) outbreaks in spinach and *Salmonella* contamination in peanut butter. Other subjects include efforts toward finding a cure for AIDS, Alzheimer's disease, and cancer. On a global scale, many researchers are committed to finding ways to protect the planet, solve environmental issues, and reduce the effects of climate change. All of these diverse endeavors are related to different facets of the discipline of biology.



Escherichia coli (*E. coli*) bacteria, seen in this scanning electron micrograph, are normal residents of our digestive tracts that aid in the absorption of vitamin K and other nutrients. However, virulent strains are sometimes responsible for disease outbreaks. (credit: Eric Erbe, digital colorization by Christopher Pooley, both of USDA, ARS, EMU)

The Process of Science

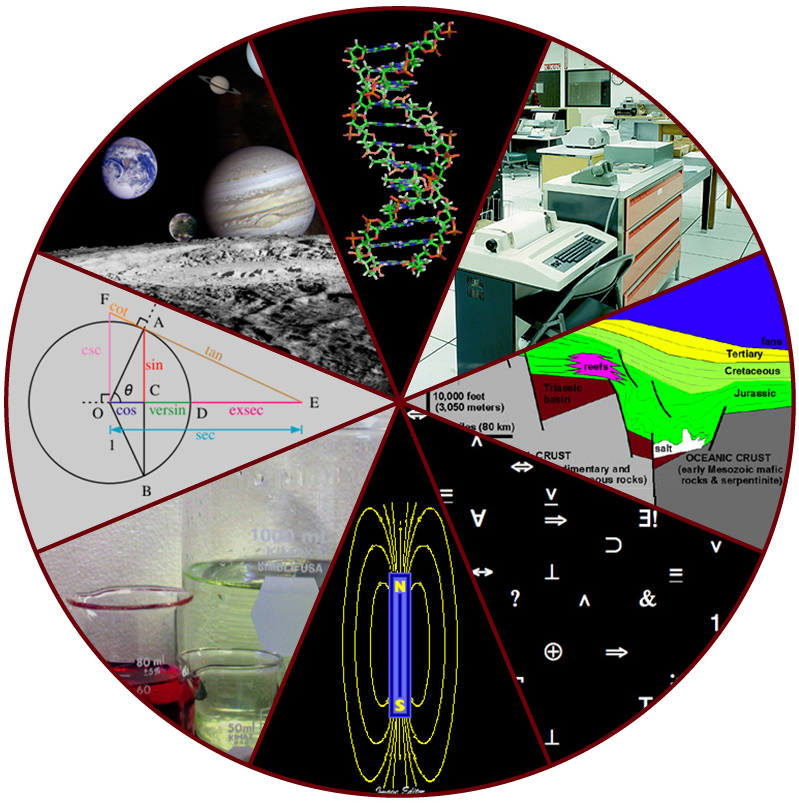
Biology is a science, but what exactly is science? What does the study of biology share with other scientific disciplines? Science(from the Latin *scientia*, meaning “knowledge”) can be defined as knowledge that covers general truths or the operation of general laws, especially when acquired and tested by the scientific method. It becomes clear from this definition that the application of the scientific method plays a major role in science. The scientific method is a method of research with defined steps that include experiments and careful observation.

The steps of the scientific method will be examined in detail later, but one of the most important aspects of this method is the testing of hypotheses by means of repeatable experiments. A hypothesis is a suggested explanation for an event, which can be tested. Although using the scientific method is inherent to science, it is inadequate in determining what science is. This is because it is relatively easy to apply the scientific method to disciplines such as physics and chemistry, but when it comes to disciplines like archaeology, psychology, and geology, the scientific method becomes less applicable as it becomes more difficult to repeat experiments.

These areas of study are still sciences, however. Consider archeology—even though one cannot perform repeatable experiments, hypotheses may still be supported. For instance, an archeologist can hypothesize that an ancient culture existed based on finding a piece of pottery. Further hypotheses could be made about various characteristics of this culture, and these hypotheses may be found to be correct or false through continued support or contradictions from other findings. A hypothesis may become a verified theory. Atheory is a tested and confirmed explanation for observations or phenomena. Science may be better defined as fields of study that attempt to comprehend the nature of the universe.

Natural Sciences

What would you expect to see in a museum of natural sciences? Frogs? Plants? Dinosaur skeletons? Exhibits about how the brain functions? A planetarium? Gems and minerals? Or, maybe all of the above? Science includes such diverse fields as astronomy, biology, computer sciences, geology, logic, physics, chemistry, and mathematics (Figure). However, those fields of science related to the physical world and its phenomena and processes are considered natural sciences. Thus, a museum of natural sciences might contain any of the items listed above.



The diversity of scientific fields includes astronomy, biology, computer science, geology, logic, physics, chemistry, mathematics, and many other fields. (credit: “Image Editor”/Flickr)

There is no complete agreement when it comes to defining what the natural sciences include, however. For some experts, the natural sciences are astronomy, biology, chemistry, earth science, and physics. Other scholars choose to divide natural sciences intolife sciences, which study living things and include biology, and physical sciences, which study nonliving matter and include astronomy, geology, physics, and chemistry. Some disciplines such as biophysics and biochemistry build on both life and physical sciences and are interdisciplinary. Natural sciences are sometimes referred to as “hard science” because they rely on the use of quantitative data; social sciences that study society and human behavior are more likely to use qualitative assessments to drive investigations and findings.

Not surprisingly, the natural science of biology has many branches or subdisciplines. Cell biologists study cell structure and function, while biologists who study anatomy investigate the structure of an entire organism. Those biologists studying physiology, however, focus on the internal functioning of an organism. Some areas of biology focus on only particular types of living things. For example, botanists explore plants, while zoologists specialize in animals.

Scientific Reasoning

One thing is common to all forms of science: an ultimate goal “to know.” Curiosity and inquiry are the driving forces for the development of science. Scientists seek to understand the world and the way it operates. To do this, they use two methods of logical thinking: inductive reasoning and deductive reasoning.

Inductive reasoning is a form of logical thinking that uses related observations to arrive at a general conclusion. This type of reasoning is common in descriptive science. A life scientist such as a biologist makes observations and records them. These data can be qualitative or quantitative, and the raw data can be supplemented with drawings, pictures, photos, or videos. From many observations, the scientist can infer conclusions (inductions) based on evidence. Inductive reasoning involves formulating generalizations inferred from careful observation and the analysis of a large amount of data. Brain studies provide an example. In this type of research, many live brains are observed while people are doing a specific activity, such as viewing images of food. The part of the brain that “lights up” during this activity is then predicted to be the part controlling the response to the selected

stimulus, in this case, images of food. The “lighting up” of the various areas of the brain is caused by excess absorption of radioactive sugar derivatives by active areas of the brain. The resultant increase in radioactivity is observed by a scanner. Then, researchers can stimulate that part of the brain to see if similar responses result.

Deductive reasoning or deduction is the type of logic used in hypothesis-based science. In deductive reason, the pattern of thinking moves in the opposite direction as compared to inductive reasoning. Deductive reasoning is a form of logical thinking that uses a general principle or law to forecast specific results. From those general principles, a scientist can extrapolate and predict the specific results that would be valid as long as the general principles are valid. Studies in climate change can illustrate this type of reasoning. For example, scientists may predict that if the climate becomes warmer in a particular region, then the distribution of plants and animals should change. These predictions have been made and tested, and many such changes have been found, such as the modification of arable areas for agriculture, with change based on temperature averages.

Both types of logical thinking are related to the two main pathways of scientific study: descriptive science and hypothesis-based science. Descriptive (or discovery) science, which is usually inductive, aims to observe, explore, and discover, while hypothesis-based science, which is usually deductive, begins with a specific question or problem and a potential answer or solution that can be tested. The boundary between these two forms of study is often blurred, and most scientific endeavors combine both approaches. The fuzzy boundary becomes apparent when thinking about how easily observation can lead to specific questions. For example, a gentleman in the 1940s observed that the burr seeds that stuck to his clothes and his dog’s fur had a tiny hook structure. On closer inspection, he discovered that the burrs’ gripping device was more reliable than a zipper. He eventually developed a company and produced the hook-and-loop fastener popularly known today as Velcro. Descriptive science and hypothesis-based science are in continuous dialogue.

The Scientific Method

Biologists study the living world by posing questions about it and seeking science-based responses. This approach is common to other sciences as well and is often referred to as the scientific method. The scientific method was used even in ancient times, but it was first documented by England’s Sir Francis Bacon (1561–1626) (Figure), who set up inductive methods for scientific inquiry. The scientific method is not exclusively used by biologists but can be applied to almost all fields of study as a logical, rational problem-solving method.

Sir Francis Bacon (1561–1626) is credited with being the first to define the scientific method. (credit: Paul van Somer)

The scientific process typically starts with an observation (often a problem to be solved) that leads to a question. Let’s think about a simple problem that starts with an observation and apply the scientific method to solve the problem. One Monday morning, a student arrives at class and quickly discovers that the classroom is too warm.

That is an observation that also describes a problem: the classroom is too warm. The student then asks a question: “Why is the classroom so warm?”

Proposing a Hypothesis

Recall that a hypothesis is a suggested explanation that can be tested. To solve a problem, several hypotheses may be proposed. For example, one hypothesis might be, “The classroom is warm because no one turned on the air conditioning.” But there could be other responses to the question, and therefore other hypotheses may be proposed. A second hypothesis might be, “The classroom is warm because there is a power failure, and so the air conditioning doesn’t work.”

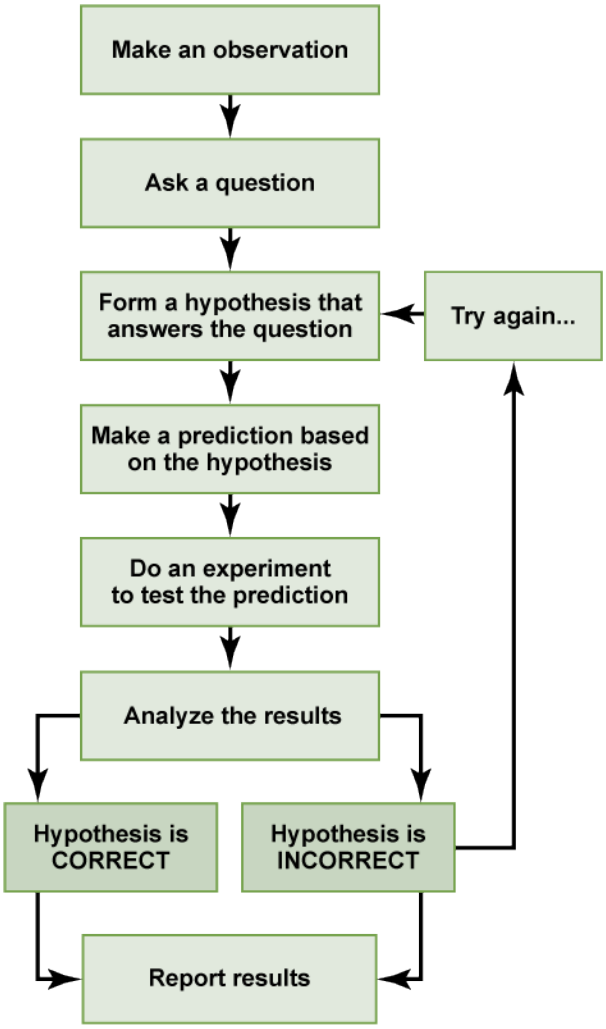
Once a hypothesis has been selected, the student can make a prediction. A prediction is similar to a hypothesis but it typically has the format “If . . . then . . . .” For example, the prediction for the first hypothesis might be, “*If* the student turns on the air conditioning,*then* the classroom will no longer be too warm.”

Testing a Hypothesis

A valid hypothesis must be testable. It should also be falsifiable, meaning that it can be disproven by experimental results. Importantly, science does not claim to “prove” anything because scientific understandings are always subject to modification with further information. This step—openness to disproving ideas—is what distinguishes sciences from non-sciences. The presence of the supernatural, for instance, is neither testable nor falsifiable. To test a hypothesis, a researcher will conduct one or more experiments designed to eliminate one or more of the hypotheses. Each experiment will have one or more variables and one or more controls. A variable is any part of the experiment that can vary or change during the experiment. The control group contains every feature of the experimental group except it is not given the manipulation that is hypothesized about. Therefore, if the results of the experimental group differ from the control group, the difference must be due to the hypothesized manipulation, rather than some outside factor. Look for the variables and controls in the examples that follow. To test the first hypothesis, the student would find out if the air conditioning is on. If the air conditioning is turned on but does not work, there should be another reason, and this hypothesis should be rejected. To test the second hypothesis, the student could check if the lights in the classroom are functional. If so, there is no power failure and this hypothesis should be rejected. Each hypothesis should be tested by carrying out appropriate experiments. Be aware that rejecting one hypothesis does not determine whether or not the other hypotheses can be accepted; it simply eliminates one hypothesis that is not valid (Figure). Using the scientific method, the hypotheses that are inconsistent with experimental data are rejected.

While this “warm classroom” example is based on observational results, other hypotheses and experiments might have clearer controls. For instance, a student might attend class on Monday and realize she had difficulty concentrating on the lecture. One observation to explain this occurrence might be, “When I eat breakfast before class, I am better able to pay attention.” The student could then design an experiment with a control to test this hypothesis.

In hypothesis-based science, specific results are predicted from a general premise. This type of reasoning is called deductive reasoning: deduction proceeds from the general to the particular. But the reverse of the process is also possible: sometimes, scientists reach a general conclusion from a number of specific observations. This type of reasoning is called inductive reasoning, and it proceeds from the particular to the general. Inductive and deductive reasoning are often used in tandem to advance scientific knowledge (Figure).
ART CONNECTION



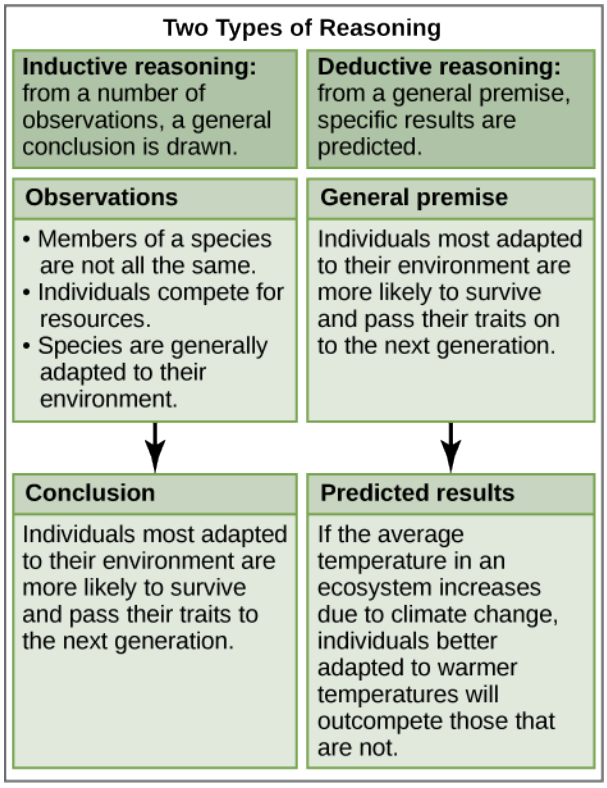
The scientific method consists of a series of well-defined steps. If a hypothesis is not supported by experimental data, a new hypothesis can be proposed.

In the example below, the scientific method is used to solve an everyday problem. Order the scientific method steps (numbered items) with the process of solving the everyday problem (lettered items). Based on the results of the experiment, is the hypothesis correct? If it is incorrect, propose some alternative hypotheses.

1. Observation
2. Question
3. Hypothesis (answer)
4. Prediction
5. Experiment
6. Result

1. There is something wrong with the electrical outlet.
2. If something is wrong with the outlet, my coffeemaker also won't work when plugged into it.
3. My toaster doesn't toast my bread.
4. I plug my coffee maker into the outlet.
5. My coffeemaker works.
6. Why doesn't my toaster work?

ART CONNECTION



Scientists use two types of reasoning, inductive and deductive reasoning, to advance scientific knowledge. As is the case in this example, the conclusion from inductive reasoning can often become the premise for inductive reasoning.

Decide if each of the following is an example of inductive or deductive reasoning.

1. All flying birds and insects have wings. Birds and insects flap their wings as they move through the air. Therefore, wings enable flight.
2. Insects generally survive mild winters better than harsh ones. Therefore, insect pests will become more problematic if global temperatures increase.
3. Chromosomes, the carriers of DNA, separate into daughter cells during cell division. Therefore, DNA is the genetic material.
4. Animals as diverse as humans, insects, and wolves all exhibit social behavior. Therefore, social behavior must have an evolutionary advantage.

The scientific method may seem too rigid and structured. It is important to keep in mind that, although scientists often follow this sequence, there is flexibility. Sometimes an experiment leads to conclusions that favor a change in approach; often, an experiment brings entirely new scientific questions to the puzzle. Many times, science does not operate in a linear fashion; instead, scientists continually draw inferences and make generalizations, finding patterns as their research proceeds. Scientific reasoning is more complex than the scientific method alone suggests. Notice, too, that the scientific method can be applied to solving problems that aren't necessarily scientific in nature.

Two Types of Science: Basic Science and Applied Science

The scientific community has been debating for the last few decades about the value of different types of science. Is it valuable to pursue science for the sake of simply gaining knowledge, or does scientific knowledge only have worth if we can apply it to solving a specific problem or to bettering our lives? This question focuses on the differences between two types of science: basic science and applied science.

Basic science or "pure" science seeks to expand knowledge regardless of the short-term application of that knowledge. It is not focused on developing a product or a service of immediate public or commercial value. The

immediate goal of basic science is knowledge for knowledge's sake, though this does not mean that, in the end, it may not result in a practical application.

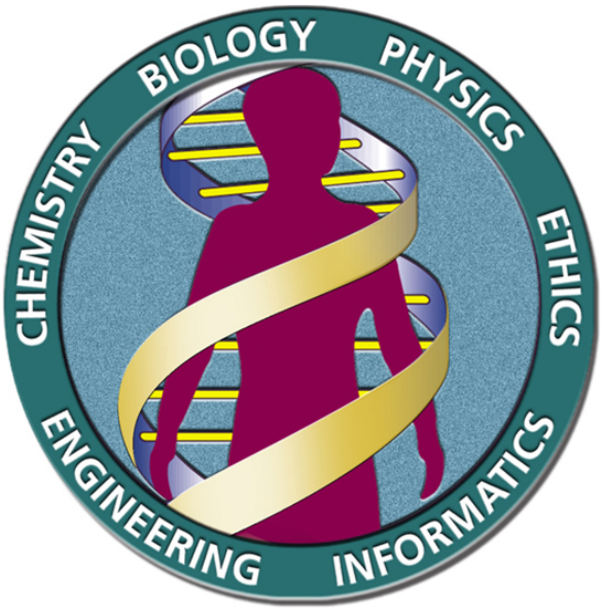
In contrast, applied science or "technology," aims to use science to solve real-world problems, making it possible, for example, to improve a crop yield, find a cure for a particular disease, or save animals threatened by a natural disaster (Figure). In applied science, the problem is usually defined for the researcher.

After Hurricane Ike struck the Gulf Coast in 2008, the U.S. Fish and Wildlife Service rescued this brown pelican. Thanks to applied science, scientists knew how to rehabilitate the bird. (credit: FEMA)

Some individuals may perceive applied science as "useful" and basic science as "useless." A question these people might pose to a scientist advocating knowledge acquisition would be, "What for?" A careful look at the history of science, however, reveals that basic knowledge has resulted in many remarkable applications of great value. Many scientists think that a basic understanding of science is necessary before an application is developed; therefore, applied science relies on the results generated through basic science. Other scientists think that it is time to move on from basic science and instead to find solutions to actual problems. Both approaches are valid. It is true that there are problems that demand immediate attention; however, few solutions would be found without the help of the wide knowledge foundation generated through basic science.

One example of how basic and applied science can work together to solve practical problems occurred after the discovery of DNA structure led to an understanding of the molecular mechanisms governing DNA replication. Strands of DNA, unique in every human, are found in our cells, where they provide the instructions necessary for life. During DNA replication, DNA makes new copies of itself, shortly before a cell divides. Understanding the mechanisms of DNA replication enabled scientists to develop laboratory techniques that are now used to identify genetic diseases, pinpoint individuals who were at a crime scene, and determine paternity. Without basic science, it is unlikely that applied science would exist.

Another example of the link between basic and applied research is the Human Genome Project, a study in which each human chromosome was analyzed and mapped to determine the precise sequence of DNA subunits and the exact location of each gene. (The gene is the basic unit of heredity; an individual's complete collection of genes is his or her genome.) Other less complex organisms have also been studied as part of this project in order to gain a better understanding of human chromosomes. The Human Genome Project (Figure) relied on basic research carried out with simple organisms and, later, with the human genome. An important end goal eventually became using the data for applied research, seeking cures and early diagnoses for genetically related diseases.



The Human Genome Project was a 13-year collaborative effort among researchers working in several different fields of science. The project, which sequenced the entire human genome, was completed in 2003. (credit: the U.S. Department of Energy Genome Programs (http://genomics.energy.gov))

While research efforts in both basic science and applied science are usually carefully planned, it is important to note that some discoveries are made by serendipity, that is, by means of a fortunate accident or a lucky surprise. Penicillin was discovered when biologist Alexander Fleming accidentally left a petri dish of *Staphylococcus* bacteria open. An unwanted mold grew on the dish, killing the bacteria. The mold turned out to be *Penicillium*, and a new antibiotic was discovered. Even in the highly organized world of science, luck—when combined with an observant, curious mind—can lead to unexpected breakthroughs.

Reporting Scientific Work

Whether scientific research is basic science or applied science, scientists must share their findings in order for other researchers to expand and build upon their discoveries. Collaboration with other scientists—when planning, conducting, and analyzing results—are all important for scientific research. For this reason, important aspects of a scientist's work are communicating with peers and disseminating results to peers. Scientists can share results by presenting them at a scientific meeting or conference, but this approach can reach only the select few who are present. Instead, most scientists present their results in peer-reviewed manuscripts that are published in scientific journals. Peer-reviewed manuscripts are scientific papers that are reviewed by a scientist's colleagues, or peers. These colleagues are qualified individuals, often experts in the same research area, who judge whether or not the scientist's work is suitable for publication. The process of peer review helps to ensure that the research described in a scientific paper or grant proposal is original, significant, logical, and thorough. Grant proposals, which are requests for research funding, are also subject to peer review. Scientists publish their work so other scientists can reproduce their experiments under similar or different conditions to expand on the findings. The experimental results must be consistent with the findings of other scientists.

A scientific paper is very different from creative writing. Although creativity is required to design experiments, there are fixed guidelines when it comes to presenting scientific results. First, scientific writing must be brief, concise, and accurate. A scientific paper needs to be succinct but detailed enough to allow peers to reproduce the experiments.

The scientific paper consists of several specific sections—introduction, materials and methods, results, and discussion. This structure is sometimes called the "IMRaD" format. There are usually acknowledgment and reference sections as well as an abstract (a concise summary) at the beginning of the paper. There might be additional sections depending on the type of paper and the journal where it will be published; for example, some review papers require an outline.

The introduction starts with brief, but broad, background information about what is known in the field. A good introduction also gives the rationale of the work; it justifies the work carried out and also briefly mentions the end of the paper, where the hypothesis or research question driving the research will be presented. The introduction refers to the published scientific work of others and therefore requires citations following the style of the journal. Using the work or ideas of others without proper citation is consideredplagiarism.

The materials and methods section includes a complete and accurate description of the substances used, and the method and techniques used by the researchers to gather data. The description should be thorough enough to allow another researcher to repeat the experiment and obtain similar results, but it does not have to be verbose. This section will also include information on how measurements were made and what types of calculations and statistical analyses were used to examine raw data. Although the materials and methods section gives an accurate description of the experiments, it does not discuss them.

Some journals require a results section followed by a discussion section, but it is more common to combine both. If the journal does not allow the combination of both sections, the results section simply narrates the findings without any further interpretation. The results are presented by means of tables or graphs, but no duplicate information should be presented. In the discussion section, the researcher will interpret the results, describe how variables may be related, and attempt to explain the observations. It is indispensable to conduct an extensive literature search to put the results in the context of previously published scientific research. Therefore, proper citations are included in this section as well.

Finally, the conclusion section summarizes the importance of the experimental findings. While the scientific paper almost certainly answered one or more scientific questions that were stated, any good research should lead to more questions. Therefore, a well-done scientific paper leaves doors open for the researcher and others to continue and expand on the findings.

Review articles do not follow the IMRAD format because they do not present original scientific findings, or primary literature; instead, they summarize and comment on findings that were published as primary literature and typically include extensive reference sections.

Section Summary

Biology is the science that studies living organisms and their interactions with one another and their environments. Science attempts to describe and understand the nature of the universe in whole or in part by rational means. Science has many fields; those fields related to the physical world and its phenomena are considered natural sciences.

Science can be basic or applied. The main goal of basic science is to expand knowledge without any expectation of short-term practical application of that knowledge. The primary goal of applied research, however, is to solve practical problems.

Two types of logical reasoning are used in science. Inductive reasoning uses particular results to produce general scientific principles. Deductive reasoning is a form of logical thinking that predicts results by applying general principles. The common thread throughout scientific research is the use of the scientific method, a step-based process that consists of making observations, defining a problem, posing hypotheses, testing these hypotheses, and drawing one or more conclusions. The testing uses proper controls. Scientists present their results in peer-reviewed scientific papers published in scientific journals. A scientific research paper consists of several well-defined sections: introduction, materials and methods, results, and, finally, a concluding discussion. Review papers summarize the research done in a particular field over a period of time.

Art Connections

Figure In the example below, the scientific method is used to solve an everyday problem. Order the scientific method steps (numbered items) with the process of solving the everyday problem (lettered items). Based on the results of the experiment, is the hypothesis correct? If it is incorrect, propose some alternative hypotheses.

1. Observation
2. Question

3. Hypothesis (answer)
4. Prediction
5. Experiment
6. Result

1. There is something wrong with the electrical outlet.
2. If something is wrong with the outlet, my coffeemaker also won't work when plugged into it.
3. My toaster doesn't toast my bread.
4. I plug my coffee maker into the outlet.
5. My coffeemaker works.
6. Why doesn't my toaster work?

Figure Decide if each of the following is an example of inductive or deductive reasoning.

1. All flying birds and insects have wings. Birds and insects flap their wings as they move through the air. Therefore, wings enable flight.
2. Insects generally survive mild winters better than harsh ones. Therefore, insect pests will become more problematic if global temperatures increase.
3. Chromosomes, the carriers of DNA, separate into daughter cells during cell division. Therefore, DNA is the genetic material.
4. Animals as diverse as humans, insects, and wolves all exhibit social behavior. Therefore, social behavior must have an evolutionary advantage.

Review Questions

The first forms of life on Earth were _____.

1. plants
2. microorganisms
3. birds
4. dinosaurs

A suggested and testable explanation for an event is called a _____.

1. hypothesis
2. variable
3. theory
4. control

Which of the following sciences is not considered a natural science?

1. biology
2. astronomy
3. physics
4. computer science

The type of logical thinking that uses related observations to arrive at a general conclusion is called _____.

1. deductive reasoning
2. the scientific method
3. hypothesis-based science
4. inductive reasoning

The process of _____ helps to ensure that a scientist's research is original, significant, logical, and thorough.

1. publication
2. public speaking
3. peer review
4. the scientific method

A person notices that her houseplants that are regularly exposed to music seem to grow more quickly than those in rooms with no music. As a result, she determines that plants grow better when exposed to music. This example most closely resembles which type of reasoning?

1. inductive reasoning
2. deductive reasoning
3. neither, because no hypothesis was made
4. both inductive and deductive reasoning

Free Response

Although the scientific method is used by most of the sciences, it can also be applied to everyday situations. Think about a problem that you may have at home, at school, or with your car, and apply the scientific method to solve it.

Give an example of how applied science has had a direct effect on your daily life.

Name two topics that are likely to be studied by biologists, and two areas of scientific study that would fall outside the realm of biology.

Thinking about the topic of cancer, write a basic science question and an applied science question that a researcher interested in this topic might ask

Glossary

abstract opening section of a scientific paper that summarizes the research and conclusions

applied science form of science that aims to solve real-world problems

basic science science that seeks to expand knowledge and understanding regardless of the short-term application of that knowledge

biology the study of living organisms and their interactions with one another and their environments

conclusion section of a scientific paper that summarizes the importance of the experimental findings

control part of an experiment that does not change during the experiment

deductive reasoning form of logical thinking that uses a general inclusive statement to forecast specific results

descriptive science (also, discovery science) form of science that aims to observe, explore, and investigate

discussion section of a scientific paper in which the author interprets experimental results, describes how variables may be related, and attempts to explain the phenomenon in question

falsifiable able to be disproven by experimental results

hypothesis suggested explanation for an observation, which can be tested

hypothesis-based science form of science that begins with a specific question and potential testable answers

inductive reasoning form of logical thinking that uses related observations to arrive at a general conclusion

introduction opening section of a scientific paper, which provides background information about what was known in the field prior to the research reported in the paper

life science field of science, such as biology, that studies living things

materials and methods section of a scientific paper that includes a complete description of the substances, methods, and techniques used by the researchers to gather data

natural science field of science that is related to the physical world and its phenomena and processes

peer-reviewed manuscript scientific paper that is reviewed by a scientist's colleagues who are experts in the field of study

physical science field of science, such as geology, astronomy, physics, and chemistry, that studies nonliving matter

plagiarism using other people's work or ideas without proper citation, creating the false impression that those are the author's original ideas

results section of a scientific paper in which the author narrates the experimental findings and presents relevant figures, pictures, diagrams, graphs, and tables, without any further interpretation

review article paper that summarizes and comments on findings that were published as primary literature

science knowledge that covers general truths or the operation of general laws, especially when acquired and tested by the scientific method

scientific method method of research with defined steps that include observation, formulation of a hypothesis, testing, and confirming or falsifying the hypothesis

serendipity fortunate accident or a lucky surprise

theory tested and confirmed explanation for observations or phenomena

variable part of an experiment that the experimenter can vary or change

- Downloads
- History
- Attribution
- More Information

Licensing & Attributions

CC licensed content, Shared previously

- the science of biology. **Authored by**: open stax. **Provided by**: open stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:agVo2CPX@12/The-Science-of-Biology. **License**: *Public Domain: No Known Copyright*
- **License**: *CC BY: Attribution*

THEMES AND CONCEPTS OF BIOLOGY

Learning Objectives

By the end of this section, you will be able to:

- Identify and describe the properties of life
- Describe the levels of organization among living things
- List examples of different sub disciplines in biology

Biology is the science that studies life. What exactly is life? This may sound like a silly question with an obvious answer, but it is not easy to define life. For example, a branch of biology called virology studies viruses, which exhibit some of the characteristics of living entities but lack others. It turns out that although viruses can attack living organisms, cause diseases, and even reproduce, they do not meet the criteria that biologists use to define life.

From its earliest beginnings, biology has wrestled with four questions: What are the shared properties that make something "alive"? How do those various living things function? When faced with the remarkable diversity of life,

how do we organize the different kinds of organisms so that we can better understand them? And, finally—what biologists ultimately seek to understand—how did this diversity arise and how is it continuing? As new organisms are discovered every day, biologists continue to seek answers to these and other questions.

Properties of Life

All groups of living organisms share several key characteristics or functions: order, sensitivity or response to stimuli, reproduction, adaptation, growth and development, regulation, homeostasis, and energy processing. When viewed together, these eight characteristics serve to define life.

Order

Organisms are highly organized structures that consist of one or more cells. Even very simple, single-celled organisms are remarkably complex. Inside each cell, atoms make up molecules. These in turn make up cell components or organelles. Multicellular organisms, which may consist of millions of individual cells, have an advantage over single-celled organisms in that their cells can be specialized to perform specific functions, and even sacrificed in certain situations for the good of the organism as a whole. How these specialized cells come together to form organs such as the heart, lung, or skin in organisms like the toad shown in Figure 1 will be discussed later.

Figure 1. A toad represents a highly organized structure consisting of cells, tissues, organs, and organ systems. (credit: "Ivengo(RUS)"/Wikimedia Commons)

Sensitivity or Response to Stimuli

Organisms respond to diverse stimuli. For example, plants can grow toward a source of light or respond to touch (Figure 2). Even tiny bacteria can move toward or away from chemicals (a process called chemotaxis) or light (phototaxis). Movement toward a stimulus is considered a positive response, while movement away from a stimulus is considered a negative response.

Figure 2. The leaves of this sensitive plant (Mimosa pudica) will instantly droop and fold when touched. After a few minutes, the plant returns to its normal state. (credit: Alex Lomas)

Concept in Action

Watch this short 13 second video to see how the sensitive plant responds to a touch stimulus.
Watch this video online: https://youtu.be/j_23Auf7Nvc

Reproduction

Single-celled organisms reproduce by first duplicating their DNA, which is the genetic material, and then dividing it equally as the cell prepares to divide to form two new cells. Many multicellular organisms (those made up of more than one cell) produce specialized reproductive cells that will form new individuals. When reproduction occurs, DNA containing genes is passed along to an organism's offspring. These genes are the reason that the offspring will belong to the same species and will have characteristics similar to the parent, such as fur color and blood type.

Adaptation

All living organisms exhibit a "fit" to their environment. Biologists refer to this fit as adaptation and it is a consequence of evolution by natural selection, which operates in every lineage of reproducing organisms. Examples of adaptations are as diverse as unique heat-resistant Archaea that live in boiling hot springs to the tongue length of a nectar-feeding moth that matches the size of the flower from which it feeds. All adaptations enhance the reproductive potential of the individual exhibiting them, including their ability to survive to reproduce. Adaptations are not constant. As an environment changes, natural selection causes the characteristics of the individuals in a population to track those changes.

Growth and Development

All organisms grow and develop according to specific instructions coded for by their genes. These genes provide instructions that will direct cellular growth and development, ensuring that a species' young (Figure 3) will grow up to exhibit many of the same characteristics as its parents.

Figure 3. Although no two look alike, these kittens have inherited genes from both parents and share many of the same characteristics. (credit: Pieter & Renée Lanser)

Regulation

Even the smallest organisms are complex and require multiple regulatory mechanisms to coordinate internal functions, such as the transport of nutrients, response to stimuli, and coping with environmental stresses. For example, organ systems such as the digestive or circulatory systems perform specific functions like carrying oxygen throughout the body, removing wastes, delivering nutrients to every cell, and cooling the body.

Homeostasis

To function properly, cells require appropriate conditions such as proper temperature, pH, and concentrations of diverse chemicals. These conditions may, however, change from one moment to the next. Organisms are able to maintain internal conditions within a narrow range almost constantly, despite environmental changes, through a process called **homeostasis** or "steady state"—the ability of an organism to maintain constant internal conditions. For example, many organisms regulate their body temperature in a process known as thermoregulation. Organisms that live in cold climates, such as the polar bear (Figure 4), have body structures that help them withstand low temperatures and conserve body heat. In hot climates, organisms have methods (such as perspiration in humans or panting in dogs) that help them to shed excess body heat.

Figure 4. Polar bears and other mammals living in ice-covered regions maintain their body temperature by generating heat and reducing heat loss through thick fur and a dense layer of fat under their skin. (credit: "longhorndave"/Flickr)

Energy Processing

All organisms (such as the California condor shown in Figure 5) use a source of energy for their metabolic activities. Some organisms capture energy from the Sun and convert it into chemical energy in food; others use chemical energy from molecules they take in.

Figure 5. A lot of energy is required for a California condor to fly. Chemical energy derived from food is used to power flight. California condors are an endangered species; scientists have strived to place a wing tag on each bird to help them identify and locate each individual bird. (credit: Pacific Southwest Region U.S. Fish and Wildlife)

Levels of Organization of Living Things

Living things are highly organized and structured, following a hierarchy on a scale from small to large. The **atom** is the smallest and most fundamental unit of matter. It consists of a nucleus surrounded by electrons. Atoms form molecules. A **molecule** is a chemical structure consisting of at least two atoms held together by a chemical bond. Many molecules that are biologically important are **macromolecules**, large molecules that are typically formed by combining smaller units called monomers. An example of a macromolecule is deoxyribonucleic acid (DNA) (Figure 6), which contains the instructions for the functioning of the organism that contains it.Concept in Action

Concept in Action

Take a look at this animation of a rotating DNA molecule.

Some cells contain aggregates of macromolecules surrounded by membranes; these are called **organelles**. Organelles are small structures that exist within cells and perform specialized functions. All living things are made of cells; the **cell** itself is the smallest fundamental unit of structure and function in living organisms. (This requirement is why viruses are not considered living: they are not made of cells. To make new viruses, they have to invade and hijack a living cell; only then can they obtain the materials they need to reproduce.) Some organisms consist of a single cell and others are multicellular. Cells are classified as prokaryotic or eukaryotic.**Prokaryotes** are single-celled organisms that lack organelles surrounded by a membrane and do not have nuclei surrounded by nuclear membranes; in contrast, the cells of **eukaryotes** do have membrane-bound organelles and nuclei.

In most multicellular organisms, cells combine to make **tissues**, which are groups of similar cells carrying out the same function. **Organs** are collections of tissues grouped together based on a common function. Organs are present not only in animals but also in plants. An **organ system** is a higher level of organization that consists of functionally related organs. For example vertebrate animals have many organ systems, such as the circulatory system that transports blood throughout the body and to and from the lungs; it includes organs such as the heart and blood vessels. **Organisms** are individual living entities. For example, each tree in a forest is an organism. Single-celled prokaryotes and single-celled eukaryotes are also considered organisms and are typically referred to as microorganisms.

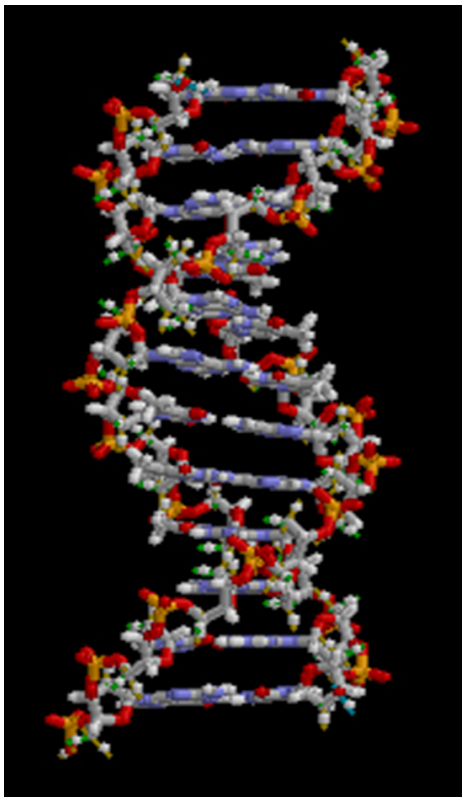


Figure 6. A molecule, like this large DNA molecule, is composed of atoms. (credit: "Brian0918"/Wikimedia Commons)

Art Connection

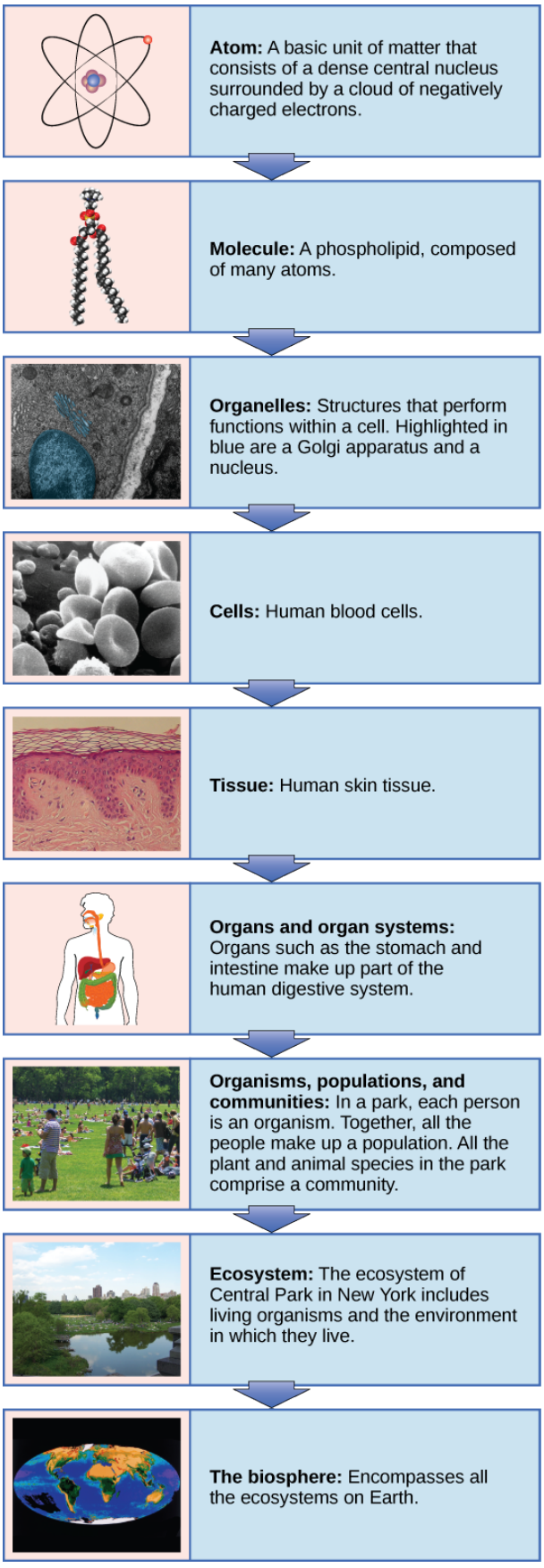

Atom: A basic unit of matter that consists of a dense central nucleus surrounded by a cloud of negatively charged electrons.
Molecule: A phospholipid, composed of many atoms.
Organelles: Structures that perform functions within a cell. Highlighted in blue are a Golgi apparatus and a nucleus.
Cells: Human blood cells.
Tissue: Human skin tissue.
Organs and organ systems: Organs such as the stomach and intestine make up part of the human digestive system.
Organisms, populations, and communities: In a park, each person is an organism. Together, all the people make up a population. All the plant and animal species in the park comprise a community.
Ecosystem: The ecosystem of Central Park in New York includes living organisms and the environment in which they live.
The biosphere: Encompasses all the ecosystems on Earth.

Figure 7. From an atom to the entire Earth, biology examines all aspects of life. (credit "molecule": modification of work by Jane Whitney; credit "organelles": modification of work by Louisa Howard; credit "cells": modification of work by Bruce Wetzel, Harry Schaefer, National Cancer Institute; credit "tissue": modification of work by "Kilbad"/Wikimedia Commons; credit "organs": modification of work by Mariana Ruiz Villareal, Joaquim Alves Gaspar; credit "organisms": modification of work by Peter Dutton; credit "ecosystem": modification of work by "gigi4791"/Flickr; credit "biosphere": modification of work by NASA)

Which of the following statements is false?

1. Tissues exist within organs which exist within organ systems.
2. Communities exist within populations which exist within ecosystems.
3. Organelles exist within cells which exist within tissues.
4. Communities exist within ecosystems which exist in the biosphere.

All the individuals of a species living within a specific area are collectively called a **population**. For example, a forest may include many white pine trees. All of these pine trees represent the population of white pine trees in this forest. Different populations may live in the same specific area. For example, the forest with the pine trees includes populations of flowering plants and also insects and microbial populations. A **community** is the set of populations inhabiting a particular area. For instance, all of the trees, flowers, insects, and other populations in a forest form the forest's community. The forest itself is an ecosystem. An **ecosystem** consists of all the living things in a particular area together with the abiotic, or non-living, parts of that environment such as nitrogen in the soil or rainwater. At the highest level of organization (Figure 7), the **biosphere** is the collection of all ecosystems, and it represents the zones of life on Earth. It includes land, water, and portions of the atmosphere.

The Diversity of Life

The science of biology is very broad in scope because there is a tremendous diversity of life on Earth. The source of this diversity is **evolution**, the process of gradual change during which new species arise from older species. Evolutionary biologists study the evolution of living things in everything from the microscopic world to ecosystems.

In the eighteenth century, a scientist named Carl Linnaeus first proposed organizing the known species of organisms into a hierarchical taxonomy. In this system, species that are most similar to each other are put together within a grouping known as a genus. Furthermore, similar genera (the plural of genus) are put together within a family. This grouping continues until all organisms are collected together into groups at the highest level. The current taxonomic system now has eight levels in its hierarchy, from lowest to highest, they are: species, genus, family, order, class, phylum, kingdom, domain. Thus species are grouped within genera, genera are grouped within families, families are grouped within orders, and so on (Figure 8).

DOMAIN **Eukarya**	Dog	Wolf	Coyote	Fox	Lion	Seal	Mouse	Human	Whale	Bat	Fish	Snake	Earthworm	Moth	Paramecium	Tree
KINGDOM **Animalia**	Dog	Wolf	Coyote	Fox	Lion	Seal	Mouse	Human	Whale	Bat	Fish	Snake	Earthworm	Moth		
PHYLUM **Chordata**	Dog	Wolf	Coyote	Fox	Lion	Seal	Mouse	Human	Whale	Bat	Fish	Snake				
CLASS **Mammalia**	Dog	Wolf	Coyote	Fox	Lion	Seal	Mouse	Human	Whale	Bat						
ORDER **Carnivora**	Dog	Wolf	Coyote	Fox	Lion	Seal										
FAMILY **Canidae**	Dog	Wolf	Coyote	Fox												
GENUS **Canis**	Dog	Wolf	Coyote													
SPECIES ***Canis lupus***	Dog	Wolf														

Figure 8. This diagram shows the levels of taxonomic hierarchy for a dog, from the broadest category—domain—to the most specific—species.

The highest level, domain, is a relatively new addition to the system since the 1990s. Scientists now recognize three domains of life, the Eukarya, the Archaea, and the Bacteria. The domain Eukarya contains organisms that have cells with nuclei. It includes the kingdoms of fungi, plants, animals, and several kingdoms of protists. The Archaea, are single-celled organisms without nuclei and include many extremophiles that live in harsh environments like hot springs. The Bacteria are another quite different group of single-celled organisms without nuclei (Figure 9). Both the Archaea and the Bacteria are prokaryotes, an informal name for cells without nuclei. The recognition in the 1990s that certain "bacteria," now known as the Archaea, were as different genetically and biochemically from other bacterial cells as they were from eukaryotes, motivated the recommendation to divide life into three domains. This dramatic change in our knowledge of the tree of life demonstrates that classifications are not permanent and will change when new information becomes available.

In addition to the hierarchical taxonomic system, Linnaeus was the first to name organisms using two unique names, now called the binomial naming system. Before Linnaeus, the use of common names to refer to organisms caused confusion because there were regional differences in these common names. Binomial names consist of the genus name (which is capitalized) and the species name (all lower-case). Both names are set in italics when they are printed. Every species is given a unique binomial which is recognized the world over, so that a scientist in any location can know which organism is being referred to. For example, the North American blue jay is known uniquely as *Cyanocitta cristata*. Our own species is *Homo sapiens*.

Figure 9. These images represent different domains. The scanning electron micrograph shows (a) bacterial cells belong to the domain Bacteria, while the (b) extremophiles, seen all together as colored mats in this hot spring, belong to domain Archaea. Both the (c) sunflower and (d) lion are part of domain Eukarya. (credit a: modification of work by Rocky Mountain Laboratories, NIAID, NIH; credit b: modification of work by Steve Jurvetson; credit c: modification of work by Michael Arrighi; credit d: modification of work by Frank Vassen)

Evolution in Action

Carl Woese and the Phylogenetic Tree

The evolutionary relationships of various life forms on Earth can be summarized in a phylogenetic tree. A **phylogenetic tree** is a diagram showing the evolutionary relationships among biological species based on similarities and differences in genetic or physical traits or both. A phylogenetic tree is composed of branch points, or nodes, and branches. The internal nodes represent ancestors and are points in evolution when, based on scientific evidence, an ancestor is thought to have diverged to form two new species. The length of each branch can be considered as estimates of relative time.

In the past, biologists grouped living organisms into five kingdoms: animals, plants, fungi, protists, and bacteria. The pioneering work of American microbiologist Carl Woese in the early 1970s has shown, however, that life on Earth has evolved along three lineages, now called domains—Bacteria, Archaea, and Eukarya. Woese proposed the domain as a new taxonomic level and Archaea as a new domain, to reflect the new phylogenetic tree (Figure 10). Many organisms belonging to the Archaea domain live under extreme conditions and are called extremophiles. To construct his tree, Woese used genetic relationships rather than similarities based on morphology (shape). Various genes were used in phylogenetic studies. Woese's tree was constructed from comparative sequencing of the genes that are universally distributed, found in some

slightly altered form in every organism, conserved (meaning that these genes have remained only slightly changed throughout evolution), and of an appropriate length.

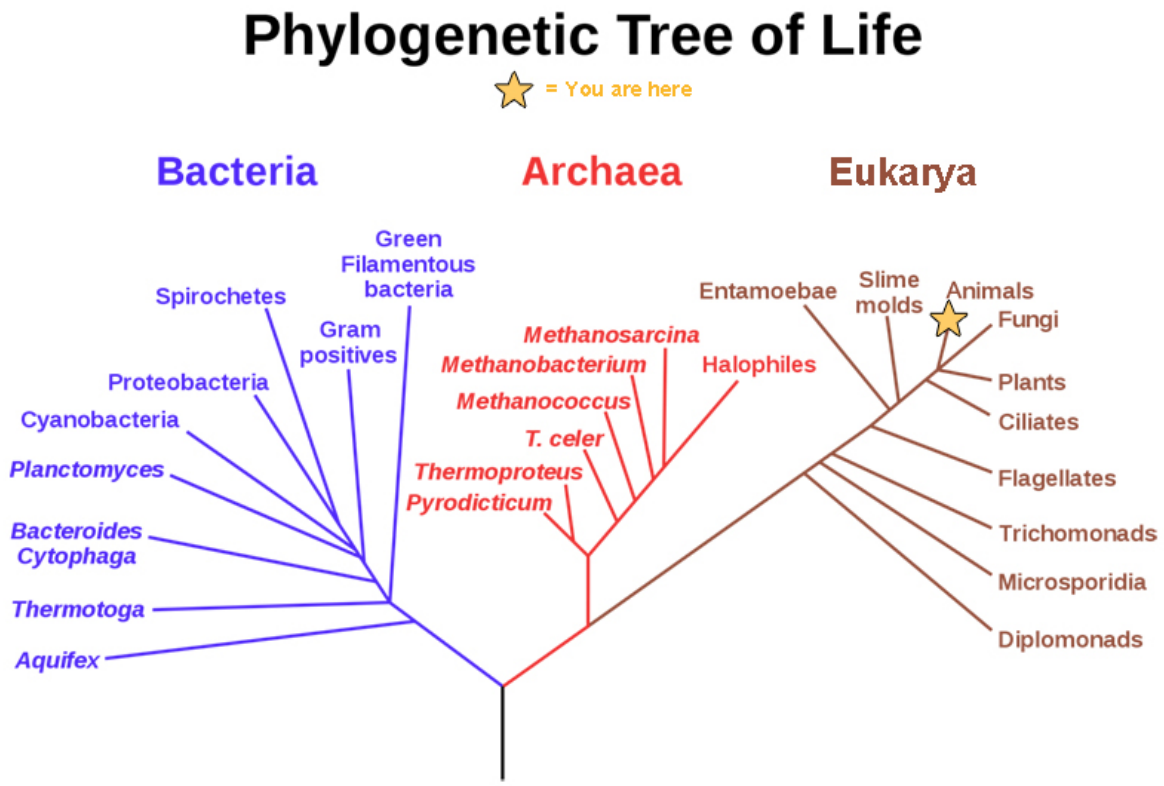


Figure 10. This phylogenetic tree was constructed by microbiologist Carl Woese using genetic relationships. The tree shows the separation of living organisms into three domains: Bacteria, Archaea, and Eukarya. Bacteria and Archaea are organisms without a nucleus or other organelles surrounded by a membrane and, therefore, are prokaryotes. (credit: modification of work by Eric Gaba)

Branches of Biological Study

The scope of biology is broad and therefore contains many branches and sub disciplines. Biologists may pursue one of those sub disciplines and work in a more focused field. For instance, molecular biology studies biological processes at the molecular level, including interactions among molecules such as DNA, RNA, and proteins, as well as the way they are regulated. Microbiology is the study of the structure and function of microorganisms. It is quite a broad branch itself, and depending on the subject of study, there are also microbial physiologists, ecologists, and geneticists, among others.

Another field of biological study, neurobiology, studies the biology of the nervous system, and although it is considered a branch of biology, it is also recognized as an interdisciplinary field of study known as neuroscience. Because of its interdisciplinary nature, this sub discipline studies different functions of the nervous system using molecular, cellular, developmental, medical, and computational approaches.

Paleontology, another branch of biology, uses fossils to study life's history (Figure 11). Zoology and botany are the study of animals and plants, respectively. Biologists can also specialize as biotechnologists, ecologists, or physiologists, to name just a few areas. Biotechnologists apply the knowledge of biology to create useful products. Ecologists study the interactions of organisms in their environments. Physiologists study the workings of cells, tissues and organs. This is just a small sample of the many fields that biologists can pursue. From our own bodies to the world we live in, discoveries in biology can affect us in very direct and important ways. We depend on these discoveries for our health, our food sources, and the benefits provided by our ecosystem. Because of this, knowledge of biology can benefit us in making decisions in our day-to-day lives.

Figure 11. Researchers work on excavating dinosaur fossils at a site in Castellón, Spain. (credit: Mario Modesto)

The development of technology in the twentieth century that continues today, particularly the technology to describe and manipulate the genetic material, DNA, has transformed biology. This transformation will allow biologists to continue to understand the history of life in greater detail, how the human body works, our human origins, and how humans can survive as a species on this planet despite the stresses caused by our increasing numbers. Biologists continue to decipher huge mysteries about life suggesting that we have only begun to understand life on the planet, its history, and our relationship to it. For this and other reasons, the knowledge of biology gained through this textbook and other printed and electronic media should be a benefit in whichever field you enter.

Visit this page in your course online to check your understanding.

Careers in Action

Forensic Scientist

Forensic science is the application of science to answer questions related to the law. Biologists as well as chemists and biochemists can be forensic scientists. Forensic scientists provide scientific evidence for use in courts, and their job involves examining trace material associated with crimes. Interest in forensic science has increased in the last few years, possibly because of popular television shows that feature forensic scientists on the job. Also, the development of molecular techniques and the establishment of DNA databases have updated the types of work that forensic scientists can do. Their job activities are primarily related to crimes against people such as murder, rape, and assault. Their work involves analyzing samples such as hair, blood, and other body fluids and also processing DNA (Figure 12) found in many different environments and materials. Forensic scientists also analyze other biological evidence left at crime scenes, such as insect parts or pollen grains. Students who want to pursue

careers in forensic science will most likely be required to take chemistry and biology courses as well as some intensive math courses.

Figure 12. This forensic scientist works in a DNA extraction room at the U.S. Army Criminal Investigation Laboratory. (credit: U.S. Army CID Command Public Affairs)

Section Summary

Biology is the science of life. All living organisms share several key properties such as order, sensitivity or response to stimuli, reproduction, adaptation, growth and development, regulation, homeostasis, and energy processing. Living things are highly organized following a hierarchy that includes atoms, molecules, organelles, cells, tissues, organs, and organ systems. Organisms, in turn, are grouped as populations, communities, ecosystems, and the biosphere. Evolution is the source of the tremendous biological diversity on Earth today. A diagram called a phylogenetic tree can be used to show evolutionary relationships among organisms. Biology is very broad and includes many branches and sub disciplines. Examples include molecular biology, microbiology, neurobiology, zoology, and botany, among others.

Licensing & Attributions

CC licensed content, Shared previously

- Themes and Concepts of Biology. **Authored by**: OpenStax College. **Located at**: http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@8.10:1/Concepts_of_Biology. **License**: *CC BY: Attribution*
- themes and concepts of biology. **Authored by**: open stax. **Provided by**: open stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:gNLp76vu@13/Themes-and-Concepts-of-Biology. **License**: *CC BY: Attribution*

All rights reserved content

- Touch-me-not (Mimosa pudica) leaves folding inward after being touched. **Authored by**: User:Brandizzi. **Provided by**: The Nature Box. **Located at**: https://www.youtube.com/watch?v=j_23Auf7Nvc. **License**: *Other.* **License Terms**: Standard YouTube License

STUDY GUIDE: INTRO AND MAJOR THEMES

Study Questions

Use this page to check your understanding of the content.
Objective: Apply the process of scientific inquiry to design experiments, analyze data, and evaluate sources of information.
Vocabulary:

1. Science
2. The scientific method
3. Hypothesis
4. Theory
5. Independent variable
6. Dependent variable
7. Standardized variable
8. Experimental group
9. Control group
10. Positive control
11. Negative control
12. Observation
13. Inference

Study Guide

1. Be able to answer (and explain) any of the “What is Science” survey questions.
2. What is the scientific definition of the word “theory”? How does this differ from the casually used definition of the word “theory”?
3. What is the difference between a hypothesis and a theory?

4. Be able to clearly describe the types of questions that SCIENCE can be used to answer…and the types of questions that SCIENCE cannot be used to answer.
5. What do you think about The Scientific Method? Should it be taught in this class? Why or why not?
6. What are the critical components of experimental design?
7. Distinguish between positive and negative controls.
8. Distinguish between observations and inferences.
9. Distinguish between independent, dependent and standardized variables.
10. Be able to analyze the design of a given experiment and make suggestions for improvement.
11. Design an experiment to answer any question.
12. Identify the properties of life, what makes something living vs not living
13. Understand the hierarchy of life from Biosphere to atom
14. Understand classification system from Domain to species

Licensing & Attributions

CC licensed content, Shared previously

- Study Guide: Process of Science. **Authored by**: Wendy Riggs. **Provided by**: College of the Redwoods. **Project**: Kaleidoscope. **License**: *CC BY: Attribution*

STUDY GUIDE: LIFE

Study Guide: Life

Study Questions

Use this page to check your understanding of the content.

Objective: Use the characteristics of life to determine whether or not something is alive.

Vocabulary

1. Emergent property
2. Atom
3. Molecule
4. Biomolecule
5. Organ
6. Tissue
7. Population
8. Ecosystem
9. Community
10. Biosphere
11. Organ system
12. Homeostasis

Study Guide Questions

1. Describe the 5 basic characteristics that identify living systems.
2. Why is it difficult to define “life”?
3. What is an emergent property? Why is life considered an emergent property?
4. Describe all the levels of organization in living systems. Be able to put them in order from simplest to most complex, and vice versa.
5. Compare and contrast the unity and diversity of life. (Your answer to this question will improve as we continue through the course.)

Licensing & Attributions

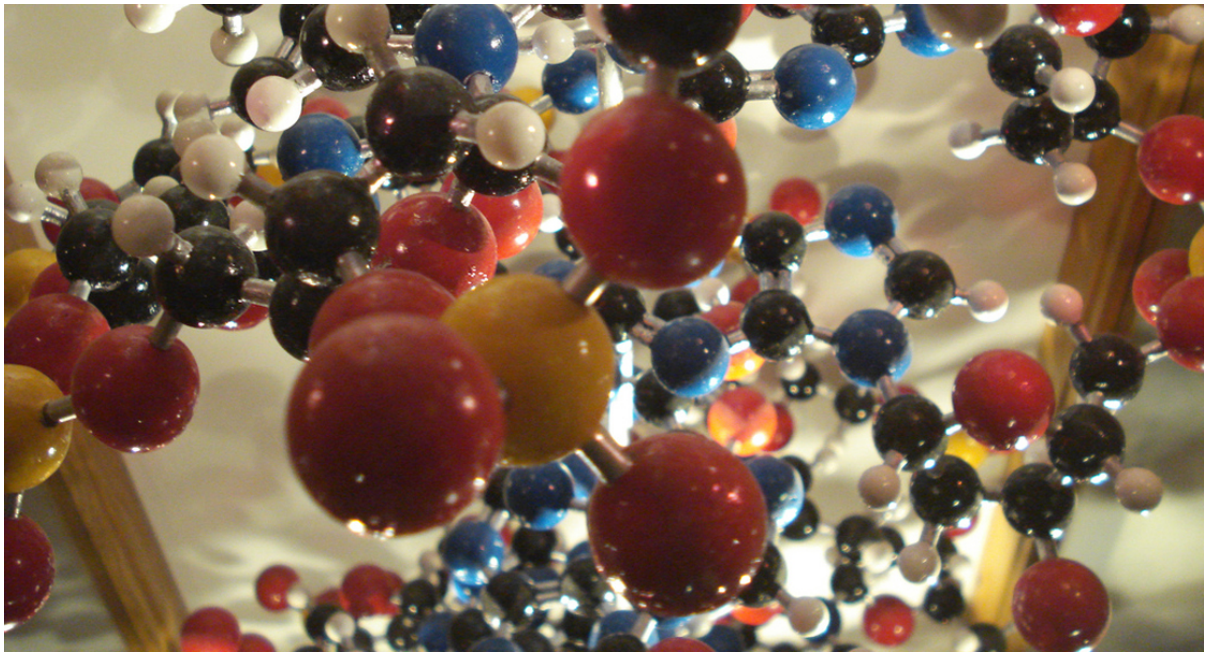
CC licensed content, Shared previously

- Study Guide: Life. **Authored by**: Wendy Riggs. **Provided by**: College of the Redwoods. **Project**: Kaleidoscope. **License**: *CC BY: Attribution*

THE CHEMICAL FOUNDATION OF LIFE

INTRODUCTION

Chapter 2. Introduction



Atoms are the building blocks of molecules found in the universe—air, soil, water, rocks . . . and also the cells of all living organisms. In this model of an organic molecule, the atoms of carbon (black), hydrogen (white), nitrogen (blue), oxygen (red), and sulfur (yellow) are shown in proportional atomic size. The silver rods indicate chemical bonds. (credit: modification of work by Christian Guthier)

Elements in various combinations comprise all matter, including living things. Some of the most abundant elements in living organisms include carbon, hydrogen, nitrogen, oxygen, sulfur, and phosphorus. These form the nucleic acids, proteins, carbohydrates, and lipids that are the fundamental components of living matter. Biologists must understand these important building blocks and the unique structures of the atoms that make up molecules, allowing for the formation of cells, tissues, organ systems, and entire organisms.

All biological processes follow the laws of physics and chemistry, so in order to understand how biological systems work, it is important to understand the underlying physics and chemistry. For example, the flow of blood within the circulatory system follows the laws of physics that regulate the modes of fluid flow. The breakdown of the large, complex molecules of food into smaller molecules—and the conversion of these to release energy to be stored in adenosine triphosphate (ATP)—is a series of chemical reactions that follow chemical laws. The properties of water and the formation of hydrogen bonds are key to understanding living processes. Recognizing the properties of acids and bases is important, for example, to our understanding of the digestive process. Therefore, the fundamentals of physics and chemistry are important for gaining insight into biological processes.

Licensing & Attributions

CC licensed content, Shared previously

- introduction to chemistry. **Authored by**: open stax. **Provided by**: open stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:djajv-uA@2/Introduction. **License**: *CC BY: Attribution*

ATOMS, ISOTOPES, IONS, AND MOLECULES: THE BUILDING BLOCKS

At its most fundamental level, life is made up of matter. Matter is any substance that occupies space and has mass. Elements are unique forms of matter with specific chemical and physical properties that cannot be broken down into smaller substances by ordinary chemical reactions. There are 118 elements, but only 92 occur naturally. The remaining elements are synthesized in laboratories and are unstable.

Each element is designated by its chemical symbol, which is a single capital letter or, when the first letter is already "taken" by another element, a combination of two letters. Some elements follow the English term for the element, such as C for carbon and Ca for calcium. Other elements' chemical symbols derive from their Latin names; for example, the symbol for sodium is Na, referring to*natrium*, the Latin word for sodium.

The four elements common to all living organisms are oxygen (O), carbon (C), hydrogen (H), and nitrogen (N). In the non-living world, elements are found in different proportions, and some elements common to living organisms are relatively rare on the earth as a whole, as shown in Table. For example, the atmosphere is rich in nitrogen and oxygen but contains little carbon and hydrogen, while the earth's crust, although it contains oxygen and a small amount of hydrogen, has little nitrogen and carbon. In spite of their differences in abundance, all elements and the chemical reactions between them obey the same chemical and physical laws regardless of whether they are a part of the living or non-living world.

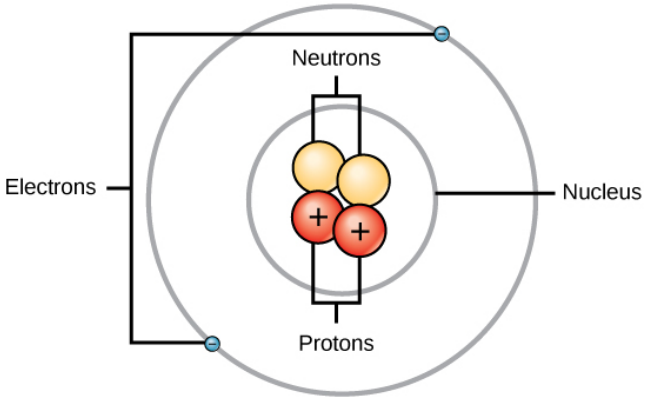
Approximate Percentage of Elements in Living Organisms (Humans) Compared to the Non-living World

Element	Life (Humans)	Atmosphere	Earth's Crust
Oxygen (O)	65%	21%	46%
Carbon (C)	18%	trace	trace
Hydrogen (H)	10%	trace	0.1%
Nitrogen (N)	3%	78%	trace

The Structure of the Atom

To understand how elements come together, we must first discuss the smallest component or building block of an element, the atom. An atom is the smallest unit of matter that retains all of the chemical properties of an element. For example, one gold atom has all of the properties of gold in that it is a solid metal at room temperature. A gold coin is simply a very large number of gold atoms molded into the shape of a coin and containing small amounts of other elements known as impurities. Gold atoms cannot be broken down into anything smaller while still retaining the properties of gold.

An atom is composed of two regions: the nucleus, which is in the center of the atom and contains protons and neutrons, and the outermost region of the atom which holds its electrons in orbit around the nucleus, as illustrated in Figure. Atoms contain protons, electrons, and neutrons, among other subatomic particles. The only exception is hydrogen (H), which is made of one proton and one electron with no neutrons.



Elements, such as helium, depicted here, are made up of atoms. Atoms are made up of protons and neutrons located within the nucleus, with electrons in orbitals surrounding the nucleus.

Protons and neutrons have approximately the same mass, about 1.67×10^{-24} grams. Scientists arbitrarily define this amount of mass as one atomic mass unit (amu) or one Dalton, as shown in Table. Although similar in mass, protons and neutrons differ in their electric charge. A proton is positively charged whereas a neutron is uncharged. Therefore, the number of neutrons in an atom contributes significantly to its mass, but not to its charge. Electrons are much smaller in mass than protons, weighing only 9.11×10^{-28} grams, or about 1/1800 of an atomic mass unit. Hence, they do not contribute much to an element's overall atomic mass. Therefore, when considering atomic mass, it is customary to ignore the mass of any electrons and calculate the atom's mass based on the number of protons and neutrons alone. Although not significant contributors to mass, electrons do contribute greatly to the atom's charge, as each electron has a negative charge equal to the positive charge of a proton. In uncharged, neutral atoms, the number of electrons orbiting the nucleus is equal to the number of protons inside the nucleus. In these atoms, the positive and negative charges cancel each other out, leading to an atom with no net charge.

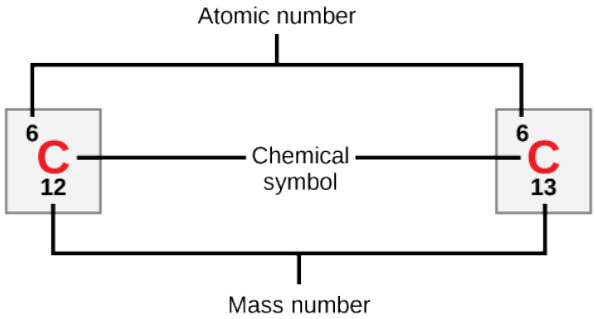
Accounting for the sizes of protons, neutrons, and electrons, most of the volume of an atom—greater than 99 percent—is, in fact, empty space. With all this empty space, one might ask why so-called solid objects do not just pass through one another. The reason they do not is that the electrons that surround all atoms are negatively charged and negative charges repel each other.

Protons, Neutrons, and Electrons			
	Charge	Mass (amu)	Location
Proton	+1	1	nucleus
Neutron	0	1	nucleus
Electron	–1	0	orbitals

Atomic Number and Mass

Atoms of each element contain a characteristic number of protons and electrons. The number of protons determines an element'satomic number and is used to distinguish one element from another. The number of neutrons is variable, resulting in isotopes, which are different forms of the same atom that vary only in the number of neutrons they possess. Together, the number of protons and the number of neutrons determine an element's mass number, as illustrated in Figure. Note that the small contribution of mass from electrons is disregarded in calculating the mass number. This approximation of mass can be used to easily calculate how many neutrons an element has by simply subtracting the number of protons from the mass number. Since an element's isotopes will have slightly different mass numbers, scientists also determine the atomic mass, which is the calculated mean of the mass number for its naturally occurring isotopes. Often, the resulting number contains a fraction. For example,

the atomic mass of chlorine (Cl) is 35.45 because chlorine is composed of several isotopes, some (the majority) with atomic mass 35 (17 protons and 18 neutrons) and some with atomic mass 37 (17 protons and 20 neutrons).
ART CONNECTION



Carbon has an atomic number of six, and two stable isotopes with mass numbers of twelve and thirteen, respectively. Its relative atomic mass is 12.011

How many neutrons do carbon-12 and carbon-13 have, respectively?

Isotopes

Isotopes are different forms of an element that have the same number of protons but a different number of neutrons. Some elements—such as carbon, potassium, and uranium—have naturally occurring isotopes. Carbon-12 contains six protons, six neutrons, and six electrons; therefore, it has a mass number of 12 (six protons and six neutrons). Carbon-14 contains six protons, eight neutrons, and six electrons; its atomic mass is 14 (six protons and eight neutrons). These two alternate forms of carbon are isotopes. Some isotopes may emit neutrons, protons, and electrons, and attain a more stable atomic configuration (lower level of potential energy); these are radioactive isotopes, or radioisotopes. Radioactive decay (carbon-14 losing neutrons to eventually become carbon-12) describes the energy loss that occurs when an unstable atom's nucleus releases radiation.
EVOLUTION CONNECTION

Carbon DatingCarbon is normally present in the atmosphere in the form of gaseous compounds like carbon dioxide and methane. Carbon-14 (^{14}C) is a naturally occurring radioisotope that is created in the atmosphere from atmospheric ^{14}N (nitrogen) by the addition of a neutron and the loss of a proton because of cosmic rays. This is a continuous process, so more ^{14}C is always being created. As a living organism incorporates ^{14}C initially as carbon dioxide fixed in the process of photosynthesis, the relative amount of^{14}C in its body is equal to the concentration of ^{14}C in the atmosphere. When an organism dies, it is no longer ingesting ^{14}C, so the ratio between ^{14}C and ^{12}C will decline as ^{14}C decays gradually to ^{14}N by a process called beta decay—the emission of electrons or positrons. This decay gives off energy in a slow process.

After approximately 5,730 years, half of the starting concentration of ^{14}C will have been converted back to ^{14}N. The time it takes for half of the original concentration of an isotope to decay back to its more stable form is called its half-life. Because the half-life of ^{14}C is long, it is used to date formerly living objects such as old bones or wood. Comparing the ratio of the ^{14}C concentration found in an object to the amount of ^{14}C detected in the atmosphere, the amount of the isotope that has not yet decayed can be determined. On the basis of this amount, the age of the material, such as the pygmy mammoth shown in Figure, can be calculated with accuracy if it is not much older than about 50,000 years. Other elements have isotopes with different half lives. For example, ^{40}K (potassium-40) has a half-life of 1.25 billion years, and ^{235}U (Uranium 235) has a half-life of about 700 million years. Through the use of radiometric dating, scientists can study the age of fossils or other remains of extinct organisms to understand how organisms have evolved from earlier species.

The age of carbon-containing remains less than about 50,000 years old, such as this pygmy mammoth, can be determined using carbon dating. (credit: Bill Faulkner, NPS)

Link to Learning

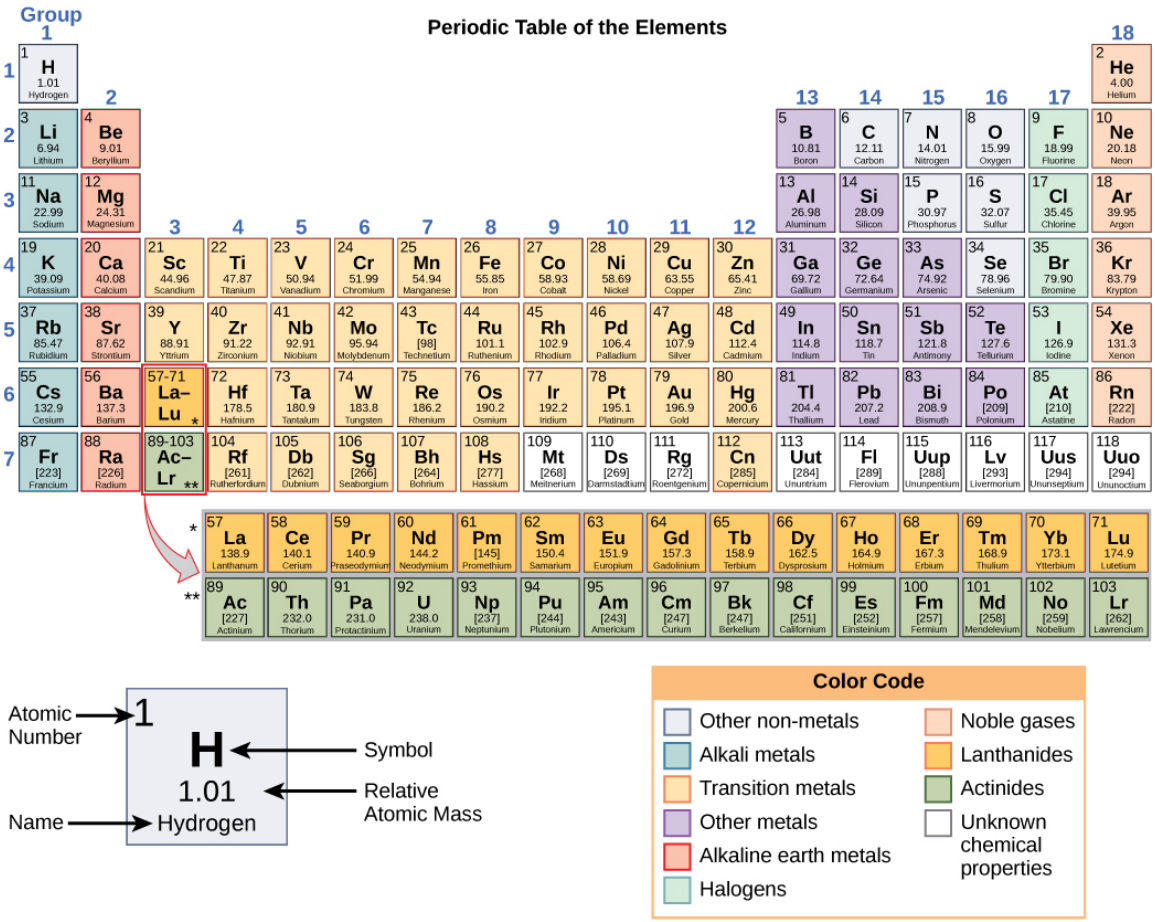
To learn more about atoms, isotopes, and how to tell one isotope from another, run the simulation.

Visit this page in your course online to use this activity.

The Periodic Table

The different elements are organized and displayed in the periodic table. Devised by Russian chemist Dmitri Mendeleev (1834–1907) in 1869, the table groups elements that, although unique, share certain chemical properties with other elements. The properties of elements are responsible for their physical state at room temperature: they may be gases, solids, or liquids. Elements also have specific chemical reactivity, the ability to combine and to chemically bond with each other.

In the periodic table, shown in Figure, the elements are organized and displayed according to their atomic number and are arranged in a series of rows and columns based on shared chemical and physical properties. In addition to providing the atomic number for each element, the periodic table also displays the element's atomic mass. Looking at carbon, for example, its symbol (C) and name appear, as well as its atomic number of six (in the upper left-hand corner) and its atomic mass of 12.11.



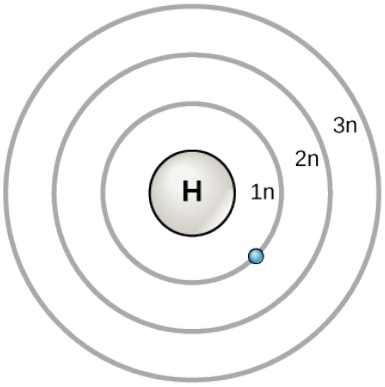
The periodic table shows the atomic mass and atomic number of each element. The atomic number appears above the symbol for the element and the approximate atomic mass appears below it.

The periodic table groups elements according to chemical properties. The differences in chemical reactivity between the elements are based on the number and spatial distribution of an atom's electrons. Atoms that chemically react and bond to each other form molecules. Molecules are simply two or more atoms chemically bonded together. Logically, when two atoms chemically bond to form a molecule, their electrons, which form the outermost region of each atom, come together first as the atoms form a chemical bond.

Electron Shells and the Bohr Model

It should be stressed that there is a connection between the number of protons in an element, the atomic number that distinguishes one element from another, and the number of electrons it has. In all electrically neutral atoms, the number of electrons is the same as the number of protons. Thus, each element, at least when electrically neutral, has a characteristic number of electrons equal to its atomic number.

An early model of the atom was developed in 1913 by Danish scientist Niels Bohr (1885–1962). The Bohr model shows the atom as a central nucleus containing protons and neutrons, with the electrons in circular orbitals at specific distances from the nucleus, as illustrated in Figure. These orbits form electron shells or energy levels, which are a way of visualizing the number of electrons in the outermost shells. These energy levels are designated by a number and the symbol "n." For example, 1n represents the first energy level located closest to the nucleus.

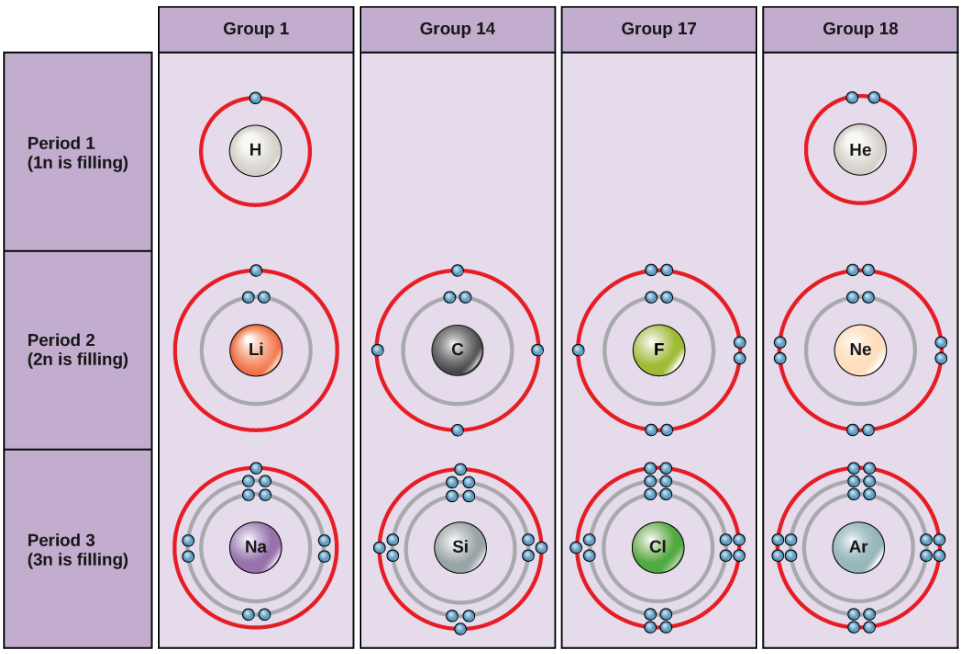


The Bohr model was developed by Niels Bohrs in 1913. In this model, electrons exist within principal shells. An electron normally exists in the lowest energy shell available, which is the one closest to the nucleus. Energy from a photon of light can bump it up to a higher energy shell, but this situation is unstable, and the electron quickly decays back to the ground state. In the process, a photon of light is released.

Electrons fill orbitals in a consistent order: they first fill the orbitals closest to the nucleus, then they continue to fill orbitals of increasing energy further from the nucleus. If there are multiple orbitals of equal energy, they will be filled with one electron in each energy level before a second electron is added. The electrons of the outermost energy level determine the energetic stability of the atom and its tendency to form chemical bonds with other atoms to form molecules.

Under standard conditions, atoms fill the inner shells first, often resulting in a variable number of electrons in the outermost shell. The innermost shell has a maximum of two electrons but the next two electron shells can each have a maximum of eight electrons. This is known as the octet rule, which states, with the exception of the innermost shell, that atoms are more stable energetically when they have eight electrons in their valence shell, the outermost electron shell. Examples of some neutral atoms and their electron configurations are shown in Figure. Notice that in this Figure, helium has a complete outer electron shell, with two electrons filling its first and only shell. Similarly, neon has a complete outer 2n shell containing eight electrons. In contrast, chlorine and sodium have seven and one in their outer shells, respectively, but theoretically they would be more energetically stable if they followed the octet rule and had eight.

ART CONNECTION



Bohr diagrams indicate how many electrons fill each principal shell. Group 18 elements (helium, neon, and argon are shown) have a full outer, or valence, shell. A full valence shell is the most stable electron configuration. Elements in other groups have partially filled valence shells and gain or lose electrons to achieve a stable electron configuration.

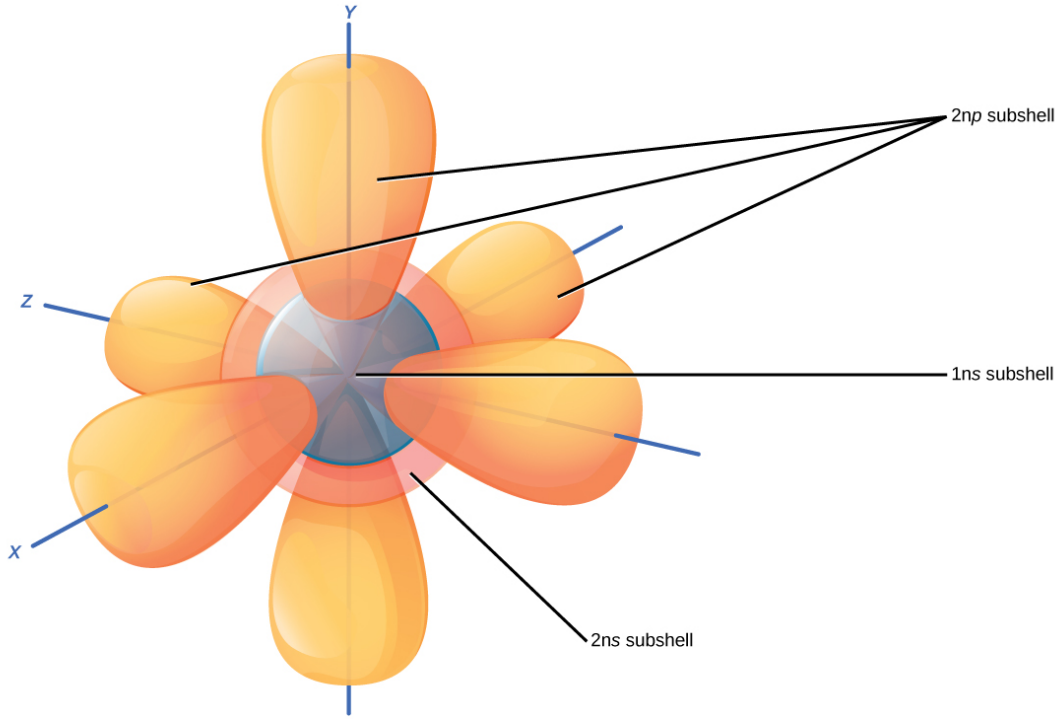
An atom may give, take, or share electrons with another atom to achieve a full valence shell, the most stable electron configuration. Looking at this figure, how many electrons do elements in group 1 need to lose in order to achieve a stable electron configuration? How many electrons do elements in groups 14 and 17 need to gain to achieve a stable configuration?

Understanding that the organization of the periodic table is based on the total number of protons (and electrons) helps us know how electrons are distributed among the outer shell. The periodic table is arranged in columns and rows based on the number of electrons and where these electrons are located. Take a closer look at the some of the elements in the table's far right column inFigure. The group 18 atoms helium (He), neon (Ne), and argon (Ar) all have filled outer electron shells, making it unnecessary for them to share electrons with other atoms to attain stability; they are highly stable as single atoms. Their non-reactivity has resulted in their being named the inert gases (or noble gases). Compare this to the group 1 elements in the left-hand column. These elements, including hydrogen (H), lithium (Li), and sodium (Na), all have one electron in their outermost shells. That means that they can achieve a stable configuration and a filled outer shell by donating or sharing one electron with another atom or a molecule such as water. Hydrogen will donate or share its electron to achieve this configuration, while lithium and sodium will donate their electron to become stable. As a result of losing a negatively charged electron, they become positively charged ions. Group 17 elements, including fluorine and chlorine, have seven electrons in their outmost shells, so they tend to fill this shell with an electron from other atoms or molecules, making them negatively charged ions. Group 14 elements, of which carbon is the most important to living systems, have four electrons in their outer shell allowing them to make several covalent bonds (discussed below) with other atoms. Thus, the columns of the periodic table represent the potential shared state of these elements' outer electron shells that is responsible for their similar chemical characteristics.

Electron Orbitals

Although useful to explain the reactivity and chemical bonding of certain elements, the Bohr model of the atom does not accurately reflect how electrons are spatially distributed surrounding the nucleus. They do not circle the nucleus like the earth orbits the sun, but are found in electron orbitals. These relatively complex shapes result from the fact that electrons behave not just like particles, but also like waves. Mathematical equations from quantum mechanics known as wave functions can predict within a certain level of probability where an electron might be at any given time. The area where an electron is most likely to be found is called its orbital.

Recall that the Bohr model depicts an atom's electron shell configuration. Within each electron shell are subshells, and each subshell has a specified number of orbitals containing electrons. While it is impossible to calculate exactly where an electron is located, scientists know that it is most probably located within its orbital path. Subshells are designated by the letter s, *p*, *d*, and *f*. The *s*subshell is spherical in shape and has one orbital. Principal shell 1n has only a single *s* orbital, which can hold two electrons. Principal shell 2n has one *s* and one *p* subshell, and can hold a total of eight electrons. The *p* subshell has three dumbbell-shaped orbitals, as illustrated in Figure. Subshells *d* and *f* have more complex shapes and contain five and seven orbitals, respectively. These are not shown in the illustration. Principal shell 3n has *s*, *p*, and *d* subshells and can hold 18 electrons. Principal shell 4n has*s*, *p*, *d* and *f* orbitals and can hold 32 electrons. Moving away from the nucleus, the number of electrons and orbitals found in the energy levels increases. Progressing from one atom to the next in the periodic table, the electron structure can be worked out by fitting an extra electron into the next available orbital.



The *s* subshells are shaped like spheres. Both the 1n and 2n principal shells have an *s* orbital, but the size of the sphere is larger in the 2n orbital. Each sphere is a single orbital. *p* subshells are made up of three dumbbell-shaped orbitals. Principal shell 2n has a *p*subshell, but shell 1 does not.

The closest orbital to the nucleus, called the 1s orbital, can hold up to two electrons. This orbital is equivalent to the innermost electron shell of the Bohr model of the atom. It is called the 1*s* orbital because it is spherical around the nucleus. The 1*s* orbital is the closest orbital to the nucleus, and it is always filled first, before any other orbital can be filled. Hydrogen has one electron; therefore, it has only one spot within the 1*s* orbital occupied. This is designated as $1s^1$, where the superscripted 1 refers to the one electron within the 1*s* orbital. Helium has two electrons; therefore, it can completely fill the 1*s* orbital with its two electrons. This is designated as $1s^2$, referring to the two electrons of helium in the 1*s* orbital. On the periodic table Figure, hydrogen and helium are the only two elements in the first row (period); this is because they only have electrons in their first shell, the 1*s* orbital. Hydrogen and helium are the only two elements that have the 1*s* and no other electron orbitals in the electrically neutral state.

The second electron shell may contain eight electrons. This shell contains another spherical *s* orbital and three "dumbbell" shaped *p*orbitals, each of which can hold two electrons, as shown in Figure. After the 1*s* orbital is filled, the second electron shell is filled, first filling its 2*s* orbital and then its three *p* orbitals. When filling the *p* orbitals, each takes a single electron; once each *p* orbital has an electron, a second may be added. Lithium (Li) contains three electrons that occupy the first and second shells. Two electrons fill the 1*s* orbital, and the third electron then fills the 2*s* orbital. Its electron configuration is $1s^2 2s^1$. Neon (Ne), on the other hand, has a total of ten electrons: two are in its innermost 1*s* orbital and eight fill its second shell (two each in the 2*s* and three *p* orbitals); thus, it is an inert gas and energetically stable as a single atom that will rarely form a chemical bond with other atoms. Larger elements have additional orbitals, making up the third electron shell. While the concepts of electron shells and orbitals are closely related, orbitals provide a more accurate depiction of the electron configuration of an atom because the orbital model specifies the different shapes and special orientations of all the places that electrons may occupy.

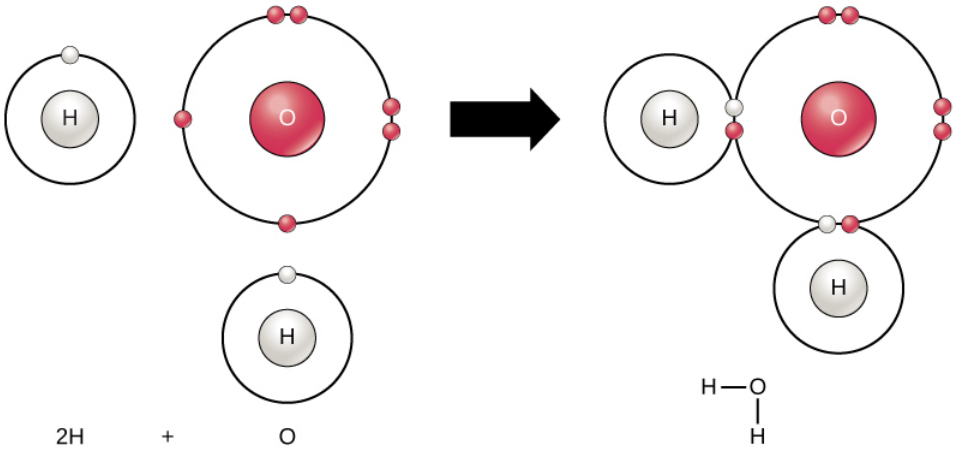
Link to Learning

Watch this visual animation to see the spatial arrangement of the p and s orbitals.

Visit this page in your course online to use this activity.

Chemical Reactions and Molecules

All elements are most stable when their outermost shell is filled with electrons according to the octet rule. This is because it is energetically favorable for atoms to be in that configuration and it makes them stable. However, since not all elements have enough electrons to fill their outermost shells, atoms form chemical bonds with other atoms thereby obtaining the electrons they need to attain a stable electron configuration. When two or more atoms chemically bond with each other, the resultant chemical structure is a molecule. The familiar water molecule, H_2O, consists of two hydrogen atoms and one oxygen atom; these bond together to form water, as illustrated in Figure. Atoms can form molecules by donating, accepting, or sharing electrons to fill their outer shells.

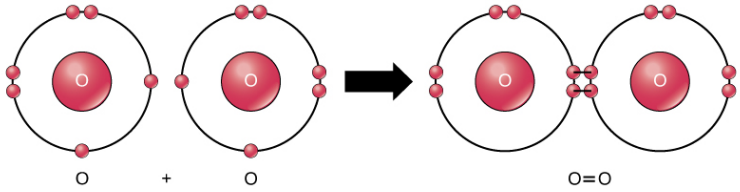


Two or more atoms may bond with each other to form a molecule. When two hydrogens and an oxygen share electrons via covalent bonds, a water molecule is formed.

Chemical reactions occur when two or more atoms bond together to form molecules or when bonded atoms are broken apart. The substances used in the beginning of a chemical reaction are called the reactants (usually found on the left side of a chemical equation), and the substances found at the end of the reaction are known as the products (usually found on the right side of a chemical equation). An arrow is typically drawn between the reactants and products to indicate the direction of the chemical reaction; this direction is not always a “one-way street.” For the creation of the water molecule shown above, the chemical equation would be:
2H + O → H2O2H + O → H2O

An example of a simple chemical reaction is the breaking down of hydrogen peroxide molecules, each of which consists of two hydrogen atoms bonded to two oxygen atoms (H_2O_2). The reactant hydrogen peroxide is broken down into water, containing one oxygen atom bound to two hydrogen atoms (H_2O), and oxygen, which consists of two bonded oxygen atoms (O_2). In the equation below, the reaction includes two hydrogen peroxide molecules and two water molecules. This is an example of a balanced chemical equation, wherein the number of atoms of each element is the same on each side of the equation. According to the law of conservation of matter, the number of atoms before and after a chemical reaction should be equal, such that no atoms are, under normal circumstances, created or destroyed.
2H2O2 (hydrogen peroxide) → 2H2O (water) + O2 (oxygen)2H2O2 (hydrogen peroxide) → 2H2O (water) + O2 (oxygen)

Even though all of the reactants and products of this reaction are molecules (each atom remains bonded to at least one other atom), in this reaction only hydrogen peroxide and water are representatives of compounds: they contain atoms of more than one type of element. Molecular oxygen, on the other hand, as shown in Figure,consists of two doubly bonded oxygen atoms and is not classified as a compound but as a mononuclear molecule.



The oxygen atoms in an O_2 molecule are joined by a double bond.

Some chemical reactions, such as the one shown above, can proceed in one direction until the reactants are all used up. The equations that describe these reactions contain a unidirectional arrow and are irreversible. Reversible reactions are those that can go in either direction. In reversible reactions, reactants are turned into products, but when the concentration of product goes beyond a certain threshold (characteristic of the particular reaction), some of these products will be converted back into reactants; at this point, the designations of products and reactants are reversed. This back and forth continues until a certain relative balance between reactants and products occurs—a state called equilibrium. These situations of reversible reactions are often denoted by a chemical equation with a double headed arrow pointing towards both the reactants and products.

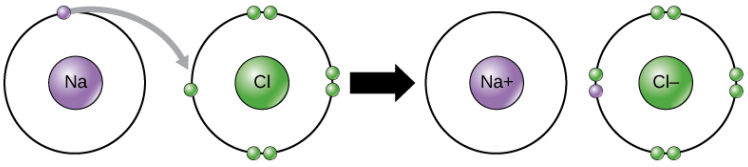
For example, in human blood, excess hydrogen ions (H^+) bind to bicarbonate ions (HCO_3^-) forming an equilibrium state with carbonic acid (H_2CO_3). If carbonic acid were added to this system, some of it would be converted to bicarbonate and hydrogen ions.
HCO3−+ H+↔H2CO3HCO3− + H+ ↔ H2CO3

In biological reactions, however, equilibrium is rarely obtained because the concentrations of the reactants or products or both are constantly changing, often with a product of one reaction being a reactant for another. To return to the example of excess hydrogen ions in the blood, the formation of carbonic acid will be the major direction of the reaction. However, the carbonic acid can also leave the body as carbon dioxide gas (via exhalation) instead of being converted back to bicarbonate ion, thus driving the reaction to the right by the chemical law known as law of mass action. These reactions are important for maintaining the homeostasis of our blood.
HCO3− + H+ ↔ H2CO3 ↔ CO2 + H2OHCO3− + H+ ↔ H2CO3 ↔ CO2 + H2O

Ions and Ionic Bonds

Some atoms are more stable when they gain or lose an electron (or possibly two) and form ions. This fills their outermost electron shell and makes them energetically more stable. Because the number of electrons does not equal the number of protons, each ion has a net charge. Cations are positive ions that are formed by losing electrons. Negative ions are formed by gaining electrons and are called anions. Anions are designated by their elemental name being altered to end in “-ide”: the anion of chlorine is called chloride, and the anion of sulfur is called sulfide, for example.

This movement of electrons from one element to another is referred to as electron transfer. As Figure illustrates, sodium (Na) only has one electron in its outer electron shell. It takes less energy for sodium to donate that one electron than it does to accept seven more electrons to fill the outer shell. If sodium loses an electron, it now has 11 protons, 11 neutrons, and only 10 electrons, leaving it with an overall charge of +1. It is now referred to as a sodium ion. Chlorine (Cl) in its lowest energy state (called the ground state) has seven electrons in its outer shell. Again, it is more energy-efficient for chlorine to gain one electron than to lose seven. Therefore, it tends to gain an electron to create an ion with 17 protons, 17 neutrons, and 18 electrons, giving it a net negative (–1) charge. It is now referred to as a chloride ion. In this example, sodium will donate its one electron to empty its shell, and chlorine will accept that electron to fill its shell. Both ions now satisfy the octet rule and have complete outermost shells. Because the number of electrons is no longer equal to the number of protons, each is now an ion and has a +1 (sodium cation) or –1 (chloride anion) charge. Note that these transactions can normally only take place simultaneously: in order for a sodium atom to lose an electron, it must be in the presence of a suitable recipient like a chlorine atom.



In the formation of an ionic compound, metals lose electrons and nonmetals gain electrons to achieve an octet.

Ionic bonds are formed between ions with opposite charges. For instance, positively charged sodium ions and negatively charged chloride ions bond together to make crystals of sodium chloride, or table salt, creating a crystalline molecule with zero net charge.

Certain salts are referred to in physiology as electrolytes (including sodium, potassium, and calcium), ions necessary for nerve impulse conduction, muscle contractions and water balance. Many sports drinks and dietary supplements provide these ions to replace those lost from the body via sweating during exercise.

Covalent Bonds and Other Bonds and Interactions

Another way the octet rule can be satisfied is by the sharing of electrons between atoms to form covalent bonds. These bonds are stronger and much more common than ionic bonds in the molecules of living organisms. Covalent bonds are commonly found in carbon-based organic molecules, such as our DNA and proteins. Covalent bonds are also found in inorganic molecules like H_2O, CO_2, and O_2. One, two, or three pairs of electrons may be shared, making single, double, and triple bonds, respectively. The more covalent bonds between two atoms, the stronger their connection. Thus, triple bonds are the strongest.

The strength of different levels of covalent bonding is one of the main reasons living organisms have a difficult time in acquiring nitrogen for use in constructing their molecules, even though molecular nitrogen, N_2, is the most abundant gas in the atmosphere. Molecular nitrogen consists of two nitrogen atoms triple bonded to each other and, as with all molecules, the sharing of these three pairs of electrons between the two nitrogen atoms allows for the filling of their outer electron shells, making the molecule more stable than the individual nitrogen atoms. This strong triple bond makes it difficult for living systems to break apart this nitrogen in order to use it as constituents of proteins and DNA.

The formation of water molecules provides an example of covalent bonding. The hydrogen and oxygen atoms that combine to form water molecules are bound together by covalent bonds, as shown in Figure. The electron from the hydrogen splits its time between the incomplete outer shell of the hydrogen atoms and the incomplete outer shell of the oxygen atoms. To completely fill the outer shell of oxygen, which has six electrons in its outer shell but which would be more stable with eight, two electrons (one from each hydrogen atom) are needed: hence the well-known formula H_2O. The electrons are shared between the two elements to fill the outer shell of each, making both elements more stable.

Link to Learning

View this short video to see an animation of ionic and covalent bonding.

Visit this page in your course online to use this activity.

Polar Covalent Bonds

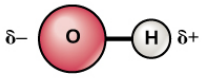
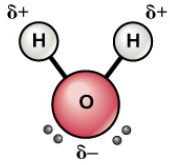
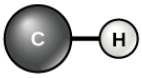
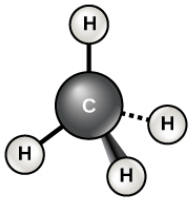
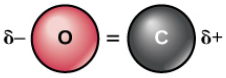
There are two types of covalent bonds: polar and nonpolar. In a polar covalent bond, shown in Figure, the electrons are unequally shared by the atoms and are attracted more to one nucleus than the other. Because of the unequal distribution of electrons between the atoms of different elements, a slightly positive ($\delta+$) or slightly negative ($\delta-$) charge develops. This partial charge is an important property of water and accounts for many of its characteristics.

Water is a polar molecule, with the hydrogen atoms acquiring a partial positive charge and the oxygen a partial negative charge. This occurs because the nucleus of the oxygen atom is more attractive to the electrons of the hydrogen atoms than the hydrogen nucleus is to the oxygen's electrons. Thus oxygen has a higher electronegativity than hydrogen and the shared electrons spend more time in the vicinity of the oxygen nucleus than they do near the nucleus of the hydrogen atoms, giving the atoms of oxygen and hydrogen slightly negative and positive charges, respectively. Another way of stating this is that the probability of finding a shared electron near an oxygen nucleus is more likely than finding it near a hydrogen nucleus. Either way, the atom's relative electronegativity contributes to the development of partial charges whenever one element is significantly more electronegative than the other, and the charges generated by these polar bonds may then be used for the formation of hydrogen bonds based on the attraction of opposite partial charges. (Hydrogen bonds, which are discussed in detail below, are weak bonds between slightly positively charged hydrogen atoms to slightly negatively charged atoms in other molecules.) Since macromolecules often have atoms within them that differ in electronegativity, polar bonds are often present in organic molecules.

Nonpolar Covalent Bonds

Nonpolar covalent bonds form between two atoms of the same element or between different elements that share electrons equally. For example, molecular oxygen (O_2) is nonpolar because the electrons will be equally distributed between the two oxygen atoms.

Another example of a nonpolar covalent bond is methane (CH_4), also shown in Figure. Carbon has four electrons in its outermost shell and needs four more to fill it. It gets these four from four hydrogen atoms, each atom providing one, making a stable outer shell of eight electrons. Carbon and hydrogen do not have the same electronegativity but are similar; thus, nonpolar bonds form. The hydrogen atoms each need one electron for their outermost shell, which is filled when it contains two electrons. These elements share the electrons equally among the carbons and the hydrogen atoms, creating a nonpolar covalent molecule.

	Bond type	Molecular shape	Molecular type
Water	 Polar covalent	 Bent	Polar
Methane	 Nonpolar covalent	 Tetrahedral	Nonpolar
Carbon dioxide	 Polar covalent	 Linear	Nonpolar

Whether a molecule is polar or nonpolar depends both on bond type and molecular shape. Both water and carbon dioxide have polar covalent bonds, but carbon dioxide is linear, so the partial charges on the molecule cancel each other out.

Hydrogen Bonds and Van Der Waals Interactions

Ionic and covalent bonds between elements require energy to break. Ionic bonds are not as strong as covalent, which determines their behavior in biological systems. However, not all bonds are ionic or covalent bonds. Weaker bonds can also form between molecules. Two weak bonds that occur frequently are hydrogen bonds and van der Waals interactions. Without these two types of bonds, life as we know it would not exist. Hydrogen bonds provide many of the critical, life-sustaining properties of water and also stabilize the structures of proteins and DNA, the building block of cells.

When polar covalent bonds containing hydrogen form, the hydrogen in that bond has a slightly positive charge because hydrogen's electron is pulled more strongly toward the other element and away from the hydrogen. Because the hydrogen is slightly positive, it will be attracted to neighboring negative charges. When this happens, a weak interaction occurs between the $\delta^+$of the hydrogen from one molecule and the δ– charge on the more electronegative atoms of another molecule, usually oxygen or nitrogen, or within the same molecule. This interaction is called a hydrogen bond. This type of bond is common and occurs regularly between water molecules. Individual hydrogen bonds are weak and easily broken; however, they occur in very large numbers in water and in organic polymers, creating a major force in combination. Hydrogen bonds are also responsible for zipping together the DNA double helix.

Like hydrogen bonds, van der Waals interactions are weak attractions or interactions between molecules. Van der Waals attractions can occur between any two or more molecules and are dependent on slight fluctuations of the electron densities, which are not always symmetrical around an atom. For these attractions to happen, the molecules need to be very close to one another. These bonds—along with ionic, covalent, and hydrogen bonds—contribute to the three-dimensional structure of the proteins in our cells that is necessary for their proper function.

CAREER CONNECTION

Pharmaceutical ChemistPharmaceutical chemists are responsible for the development of new drugs and trying to determine the mode of action of both old and new drugs. They are involved in every step of the drug development process. Drugs can be found in the natural environment or can be synthesized in the laboratory. In many cases, potential drugs found in nature are changed chemically in the laboratory to make them safer and more effective, and sometimes synthetic versions of drugs substitute for the version found in nature.

After the initial discovery or synthesis of a drug, the chemist then develops the drug, perhaps chemically altering it, testing it to see if the drug is toxic, and then designing methods for efficient large-scale production. Then, the process of getting the drug approved for human use begins. In the United States, drug approval is handled by the Food and Drug Administration (FDA) and involves a series of large-scale experiments using human subjects to make sure the drug is not harmful and effectively treats the condition it aims to treat. This process often takes several years and requires the participation of physicians and scientists, in addition to chemists, to complete testing and gain approval.

An example of a drug that was originally discovered in a living organism is Paclitaxel (Taxol), an anti-cancer drug used to treat breast cancer. This drug was discovered in the bark of the pacific yew tree. Another example is aspirin, originally isolated from willow tree bark. Finding drugs often means testing hundreds of samples of plants, fungi, and other forms of life to see if any biologically active compounds are found within them. Sometimes, traditional medicine can give modern medicine clues to where an active compound can be found. For example, the use of willow bark to make medicine has been known for thousands of years, dating back to ancient Egypt. It was not until the late 1800s, however, that the aspirin molecule, known as acetylsalicylic acid, was purified and marketed for human use.

Occasionally, drugs developed for one use are found to have unforeseen effects that allow these drugs to be used in other, unrelated ways. For example, the drug minoxidil (Rogaine) was originally developed to treat high blood pressure. When tested on humans, it was noticed that individuals taking the drug would grow new hair. Eventually the drug was marketed to men and women with baldness to restore lost hair.

The career of the pharmaceutical chemist may involve detective work, experimentation, and drug development, all with the goal of making human beings healthier.

Section Summary

Matter is anything that occupies space and has mass. It is made up of elements. All of the 92 elements that occur naturally have unique qualities that allow them to combine in various ways to create molecules, which in turn combine to form cells, tissues, organ systems, and organisms. Atoms, which consist of protons, neutrons, and electrons, are the smallest units of an element that retain all of the properties of that element. Electrons can be transferred, shared, or cause charge disparities between atoms to create bonds, including ionic, covalent, and hydrogen bonds, as well as van der Waals interactions.

Art Connections

Figure How many neutrons do carbon-12 and carbon-13 have, respectively?

Figure An atom may give, take, or share electrons with another atom to achieve a full valence shell, the most stable electron configuration. Looking at this figure, how many electrons do elements in group 1 need to lose in order to achieve a stable electron configuration? How many electrons do elements in groups 14 and 17 need to gain to achieve a stable configuration?

Review Questions

If xenon has an atomic number of 54 and a mass number of 108, how many neutrons does it have?

1. 54
2. 27
3. 100
4. 108

Atoms that vary in the number of neutrons found in their nuclei are called _____.

1. ions
2. neutrons
3. neutral atoms
4. isotopes

Potassium has an atomic number of 19. What is its electron configuration?

1. shells 1 and 2 are full, and shell 3 has nine electrons
2. shells 1, 2 and 3 are full and shell 4 has three electrons
3. shells 1, 2 and 3 are full and shell 4 has one electron
4. shells 1, 2 and 3 are full and no other electrons are present

Which type of bond represents a weak chemical bond?

1. hydrogen bond
2. atomic bond
3. covalent bond
4. nonpolar covalent bond

Free Response

What makes ionic bonds different from covalent bonds?

Why are hydrogen bonds and van der Waals interactions necessary for cells?

Glossary

anion negative ion that is formed by an atom gaining one or more electrons

atom the smallest unit of matter that retains all of the chemical properties of an element

atomic mass calculated mean of the mass number for an element's isotopes

atomic number total number of protons in an atom

balanced chemical equation statement of a chemical reaction with the number of each type of atom equalized for both the products and reactants

cation positive ion that is formed by an atom losing one or more electrons

chemical bond interaction between two or more of the same or different atoms that results in the formation of molecules

chemical reaction process leading to the rearrangement of atoms in molecules

chemical reactivity the ability to combine and to chemically bond with each other

compound substance composed of molecules consisting of atoms of at least two different elements

covalent bond type of strong bond formed between two of the same or different elements; forms when electrons are shared between atoms

electrolyte ion necessary for nerve impulse conduction, muscle contractions and water balance

electron negatively charged subatomic particle that resides outside of the nucleus in the electron orbital; lacks functional mass and has a negative charge of –1 unit

electron configuration arrangement of electrons in an atom's electron shell (for example, $1s^2 2s^2 2p^6$)

electron orbital how electrons are spatially distributed surrounding the nucleus; the area where an electron is most likely to be found

electron transfer movement of electrons from one element to another; important in creation of ionic bonds

electronegativity ability of some elements to attract electrons (often of hydrogen atoms), acquiring partial negative charges in molecules and creating partial positive charges on the hydrogen atoms

element one of 118 unique substances that cannot be broken down into smaller substances; each element has unique properties and a specified number of protons

equilibrium steady state of relative reactant and product concentration in reversible chemical reactions in a closed system

hydrogen bond weak bond between slightly positively charged hydrogen atoms to slightly negatively charged atoms in other molecules

inert gas (also, noble gas) element with filled outer electron shell that is unreactive with other atoms

ion atom or chemical group that does not contain equal numbers of protons and electrons

ionic bond chemical bond that forms between ions with opposite charges (cations and anions)

irreversible chemical reaction chemical reaction where reactants proceed uni-directionally to form products

isotope one or more forms of an element that have different numbers of neutrons

law of mass action chemical law stating that the rate of a reaction is proportional to the concentration of the reacting substances

mass number total number of protons and neutrons in an atom

matter anything that has mass and occupies space

molecule two or more atoms chemically bonded together

neutron uncharged particle that resides in the nucleus of an atom; has a mass of one amu

noble gas see inert gas

nonpolar covalent bond type of covalent bond that forms between atoms when electrons are shared equally between them

nucleus core of an atom; contains protons and neutrons

octet rule rule that atoms are most stable when they hold eight electrons in their outermost shells

orbital region surrounding the nucleus; contains electrons

periodic table organizational chart of elements indicating the atomic number and atomic mass of each element; provides key information about the properties of the elements

polar covalent bond type of covalent bond that forms as a result of unequal sharing of electrons, resulting in the creation of slightly positive and slightly negative charged regions of the molecule

product molecule found on the right side of a chemical equation

proton positively charged particle that resides in the nucleus of an atom; has a mass of one amu and a charge of +1

radioisotope isotope that emits radiation composed of subatomic particles to form more stable elements

reactant molecule found on the left side of a chemical equation

reversible chemical reaction chemical reaction that functions bi-directionally, where products may turn into reactants if their concentration is great enough

valence shell outermost shell of an atom

van der Waals interaction very weak interaction between molecules due to temporary charges attracting atoms that are very close together

Licensing & Attributions

CC licensed content, Shared previously

- Atoms, Isotopes, Ions, and Molecules: The Building Blocks. **Authored by**: open stax. **Provided by**: open stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:vogY0C26@17/Atoms-Isotopes-Ions-and-Molecu. **License**: *Public Domain: No Known Copyright*

WATER AND PH

2.2 Water

Summary

Why do scientists spend time looking for water on other planets? Why is water so important? It is because water is essential to life as we know it. Water is one of the more abundant molecules and the one most critical to life on Earth. Approximately 60–70 percent of the human body is made up of water. Without it, life as we know it simply would not exist.

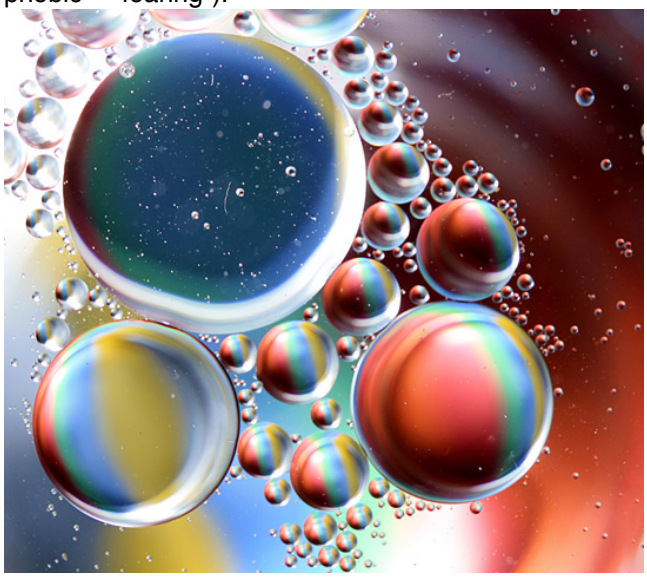
The polarity of the water molecule and its resulting hydrogen bonding make water a unique substance with special properties that are intimately tied to the processes of life. Life originally evolved in a watery environment, and most of an organism's cellular chemistry and metabolism occur inside the watery contents of the cell's cytoplasm. Special properties of water are its high heat capacity and heat of vaporization, its ability to dissolve polar molecules, its cohesive and adhesive properties, and its dissociation into ions that leads to the generation of pH. Understanding these characteristics of water helps to elucidate its importance in maintaining life.

Water's Polarity

One of water's important properties is that it is composed of polar molecules: the hydrogen and oxygen within water molecules (H_2O) form polar covalent bonds. While there is no net charge to a water molecule, the polarity of water creates a slightly positive charge on hydrogen and a slightly negative charge on oxygen, contributing to water's properties of attraction. Water's charges are generated because oxygen is more electronegative than hydrogen, making it more likely that a shared electron would be found near the oxygen nucleus than the hydrogen nucleus, thus generating the partial negative charge near the oxygen.

As a result of water's polarity, each water molecule attracts other water molecules because of the opposite charges between water molecules, forming hydrogen bonds. Water also attracts or is attracted to other polar molecules and ions. A polar substance that interacts readily with or dissolves in water is referred to as hydrophilic (hydro- = "water"; -philic = "loving"). In contrast, non-polar molecules such as oils and fats do not interact well with

water, as shown in Figure and separate from it rather than dissolve in it, as we see in salad dressings containing oil and vinegar (an acidic water solution). These nonpolar compounds are called hydrophobic(hydro- = “water”; -phobic = “fearing”).



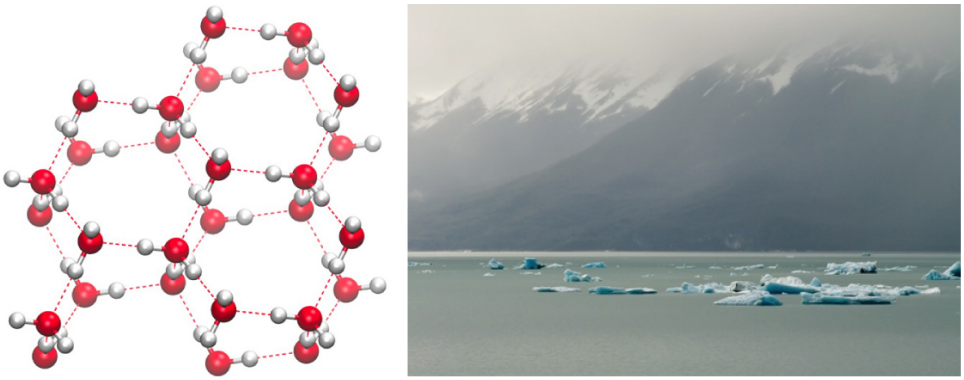
Oil and water do not mix. As this macro image of oil and water shows, oil does not dissolve in water but forms droplets instead. This is due to it being a nonpolar compound. (credit: Gautam Dogra).

Water’s States: Gas, Liquid, and Solid

The formation of hydrogen bonds is an important quality of the liquid water that is crucial to life as we know it. As water molecules make hydrogen bonds with each other, water takes on some unique chemical characteristics compared to other liquids and, since living things have a high water content, understanding these chemical features is key to understanding life. In liquid water, hydrogen bonds are constantly formed and broken as the water molecules slide past each other. The breaking of these bonds is caused by the motion (kinetic energy) of the water molecules due to the heat contained in the system. When the heat is raised as water is boiled, the higher kinetic energy of the water molecules causes the hydrogen bonds to break completely and allows water molecules to escape into the air as gas (steam or water vapor). On the other hand, when the temperature of water is reduced and water freezes, the water molecules form a crystalline structure maintained by hydrogen bonding (there is not enough energy to break the hydrogen bonds) that makes ice less dense than liquid water, a phenomenon not seen in the solidification of other liquids.

Water’s lower density in its solid form is due to the way hydrogen bonds are oriented as it freezes: the water molecules are pushed farther apart compared to liquid water. With most other liquids, solidification when the temperature drops includes the lowering of kinetic energy between molecules, allowing them to pack even more tightly than in liquid form and giving the solid a greater density than the liquid.

The lower density of ice, illustrated and pictured in Figure, an anomaly, causes it to float at the surface of liquid water, such as in an iceberg or in the ice cubes in a glass of ice water. In lakes and ponds, ice will form on the surface of the water creating an insulating barrier that protects the animals and plant life in the pond from freezing. Without this layer of insulating ice, plants and animals living in the pond would freeze in the solid block of ice and could not survive. The detrimental effect of freezing on living organisms is caused by the expansion of ice relative to liquid water. The ice crystals that form upon freezing rupture the delicate membranes essential for the function of living cells, irreversibly damaging them. Cells can only survive freezing if the water in them is temporarily replaced by another liquid like glycerol.



(a) (b)

Hydrogen bonding makes ice less dense than liquid water. The (a) lattice structure of ice makes it less dense than the freely flowing molecules of liquid water, enabling it to (b) float on water. (credit a: modification of work by Jane Whitney, image created using Visual Molecular Dynamics (VMD) software[1]; credit b: modification of work by Carlos Ponte)

Link to Learning

Click here to see a 3-D animation of the structure of an ice lattice. (Image credit: Jane Whitney. Image created using Visual Molecular Dynamics VMD software.[2])

Water's High Heat Capacity

Water's high heat capacity is a property caused by hydrogen bonding among water molecules. Water has the highest specific heat capacity of any liquids. Specific heat is defined as the amount of heat one gram of a substance must absorb or lose to change its temperature by one degree Celsius. For water, this amount is one calorie. It therefore takes water a long time to heat and long time to cool. In fact, the specific heat capacity of water is about five times more than that of sand. This explains why the land cools faster than the sea. Due to its high heat capacity, water is used by warm blooded animals to more evenly disperse heat in their bodies: it acts in a similar manner to a car's cooling system, transporting heat from warm places to cool places, causing the body to maintain a more even temperature.

Water's Heat of Vaporization

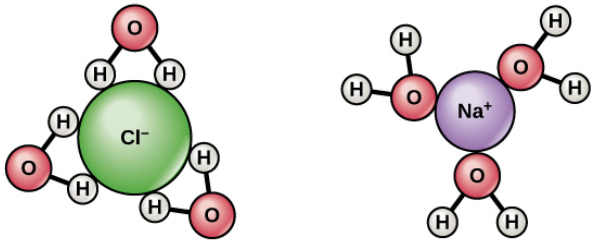
Water also has a high heat of vaporization, the amount of energy required to change one gram of a liquid substance to a gas. A considerable amount of heat energy (586 cal) is required to accomplish this change in water. This process occurs on the surface of water. As liquid water heats up, hydrogen bonding makes it difficult to separate the liquid water molecules from each other, which is required for it to enter its gaseous phase (steam). As a result, water acts as a heat sink or heat reservoir and requires much more heat to boil than does a liquid such as ethanol (grain alcohol), whose hydrogen bonding with other ethanol molecules is weaker than water's hydrogen bonding. Eventually, as water reaches its boiling point of 100° Celsius (212° Fahrenheit), the heat is able to break the hydrogen bonds between the water molecules, and the kinetic energy (motion) between the water molecules allows them to escape from the liquid as a gas. Even when below its boiling point, water's individual molecules acquire enough energy from other water molecules such that some surface water molecules can escape and vaporize: this process is known as evaporation.

The fact that hydrogen bonds need to be broken for water to evaporate means that a substantial amount of energy is used in the process. As the water evaporates, energy is taken up by the process, cooling the environment where the evaporation is taking place. In many living organisms, including in humans, the evaporation of sweat, which is 90 percent water, allows the organism to cool so that homeostasis of body temperature can be maintained.

Water's Solvent Properties

Since water is a polar molecule with slightly positive and slightly negative charges, ions and polar molecules can readily dissolve in it. Therefore, water is referred to as a solvent, a substance capable of dissolving other polar molecules and ionic compounds. The charges associated with these molecules will form hydrogen bonds with water, surrounding the particle with water molecules. This is referred to as a sphere of hydration, or a hydration shell, as illustrated in Figure and serves to keep the particles separated or dispersed in the water.

When ionic compounds are added to water, the individual ions react with the polar regions of the water molecules and their ionic bonds are disrupted in the process of dissociation. Dissociation occurs when atoms or groups of atoms break off from molecules and form ions. Consider table salt (NaCl, or sodium chloride): when NaCl crystals are added to water, the molecules of NaCl dissociate into Na^+ and Cl^- ions, and spheres of hydration form around the ions, illustrated in Figure. The positively charged sodium ion is surrounded by the partially negative charge of the water molecule's oxygen. The negatively charged chloride ion is surrounded by the partially positive charge of the hydrogen on the water molecule.

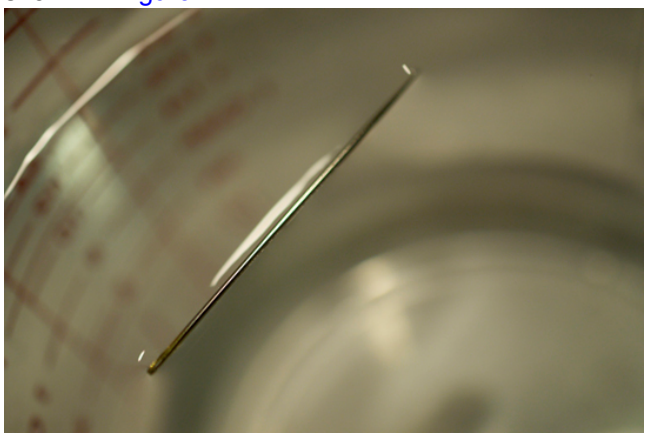


When table salt (NaCl) is mixed in water, spheres of hydration are formed around the ions.

Water's Cohesive and Adhesive Properties

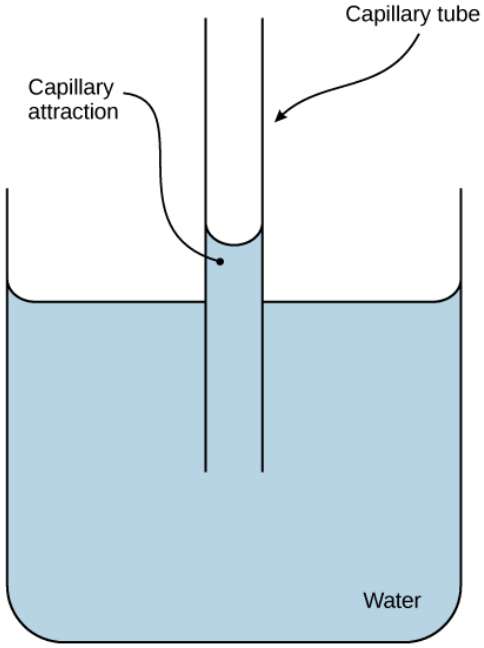
Have you ever filled a glass of water to the very top and then slowly added a few more drops? Before it overflows, the water forms a dome-like shape above the rim of the glass. This water can stay above the glass because of the property of cohesion. In cohesion, water molecules are attracted to each other (because of hydrogen bonding), keeping the molecules together at the liquid-gas (water-air) interface, although there is no more room in the glass.

Cohesion allows for the development of surface tension, the capacity of a substance to withstand being ruptured when placed under tension or stress. This is also why water forms droplets when placed on a dry surface rather than being flattened out by gravity. When a small scrap of paper is placed onto the droplet of water, the paper floats on top of the water droplet even though paper is denser (heavier) than the water. Cohesion and surface tension keep the hydrogen bonds of water molecules intact and support the item floating on the top. It's even possible to "float" a needle on top of a glass of water if it is placed gently without breaking the surface tension, as shown in Figure.



The weight of the needle is pulling the surface downward; at the same time, the surface tension is pulling it up, suspending it on the surface of the water and keeping it from sinking. Notice the indentation in the water around the needle. (credit: Cory Zanker)

These cohesive forces are related to water's property of adhesion, or the attraction between water molecules and other molecules. This attraction is sometimes stronger than water's cohesive forces, especially when the water is exposed to charged surfaces such as those found on the inside of thin glass tubes known as capillary tubes. Adhesion is observed when water "climbs" up the tube placed in a glass of water: notice that the water appears to be higher on the sides of the tube than in the middle. This is because the water molecules are attracted to the charged glass walls of the capillary more than they are to each other and therefore adhere to it. This type of adhesion is called capillary action, and is illustrated in Figure.



Capillary action in a glass tube is caused by the adhesive forces exerted by the internal surface of the glass exceeding the cohesive forces between the water molecules themselves. (credit: modification of work by Pearson-Scott Foresman, donated to the Wikimedia Foundation)

Why are cohesive and adhesive forces important for life? Cohesive and adhesive forces are important for the transport of water from the roots to the leaves in plants. These forces create a "pull" on the water column. This pull results from the tendency of water molecules being evaporated on the surface of the plant to stay connected to water molecules below them, and so they are pulled along. Plants use this natural phenomenon to help transport water from their roots to their leaves. Without these properties of water, plants would be unable to receive the water and the dissolved minerals they require. In another example, insects such as the water strider, shown in Figure, use the surface tension of water to stay afloat on the surface layer of water and even mate there.

Water's cohesive and adhesive properties allow this water strider (*Gerris* sp.) to stay afloat. (credit: Tim Vickers)

pH, Buffers, Acids, and Bases

The pH of a solution indicates its acidity or alkalinity.
H2O(l)↔H+(aq)+OH−(aq)H2O(l) ↔H+(aq)+OH-(aq)

litmus or pH paper, filter paper that has been treated with a natural water-soluble dye so it can be used as a pH indicator, to test how much acid (acidity) or base (alkalinity) exists in a solution. You might have even used some to test whether the water in a swimming pool is properly treated. In both cases, the pH test measures the concentration of hydrogen ions in a given solution.

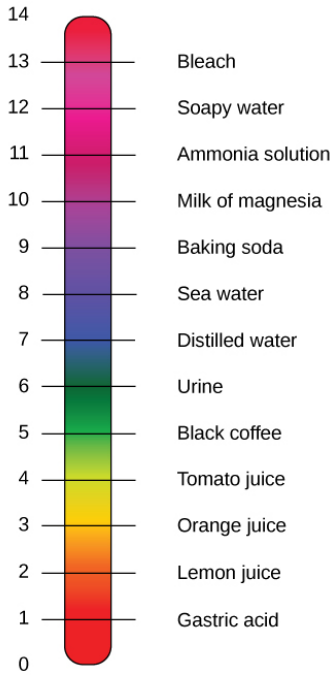
Hydrogen ions are spontaneously generated in pure water by the dissociation (ionization) of a small percentage of water molecules into equal numbers of hydrogen (H^+) ions and hydroxide (OH^-) ions. While the hydroxide ions are kept in solution by their hydrogen bonding with other water molecules, the hydrogen ions, consisting of naked protons, are immediately attracted to un-ionized water molecules, forming hydronium ions (H_30^+). Still, by convention, scientists refer to hydrogen ions and their concentration as if they were free in this state in liquid water.

The concentration of hydrogen ions dissociating from pure water is 1×10^{-7} moles H^+ ions per liter of water. Moles (mol) are a way to express the amount of a substance (which can be atoms, molecules, ions, etc), with one mole being equal to 6.02×10^{23} particles of the substance. Therefore, 1 mole of water is equal to 6.02×10^{23} water molecules. The pH is calculated as the negative of the base 10 logarithm of this concentration. The log10 of 1×10^{-7} is -7.0, and the negative of this number (indicated by the “p” of “pH”) yields a pH of 7.0, which is also known as neutral pH. The pH inside of human cells and blood are examples of two areas of the body where near-neutral pH is maintained.

Non-neutral pH readings result from dissolving acids or bases in water. Using the negative logarithm to generate positive integers, high concentrations of hydrogen ions yield a low pH number, whereas low levels of hydrogen ions result in a high pH. An acid is a substance that increases the concentration of hydrogen ions (H^+) in a solution, usually by having one of its hydrogen atoms dissociate. A base provides either hydroxide ions (OH^-) or other negatively charged ions that combine with hydrogen ions, reducing their concentration in the solution and thereby raising the pH. In cases where the base releases hydroxide ions, these ions bind to free hydrogen ions, generating new water molecules.

The stronger the acid, the more readily it donates H^+. For example, hydrochloric acid (HCl) completely dissociates into hydrogen and chloride ions and is highly acidic, whereas the acids in tomato juice or vinegar do not completely dissociate and are considered weak acids. Conversely, strong bases are those substances that readily donate OH^- or take up hydrogen ions. Sodium hydroxide (NaOH) and many household cleaners are highly alkaline and give up OH^- rapidly when placed in water, thereby raising the pH. An example of a weak basic solution is seawater, which has a pH near 8.0, close enough to neutral pH that marine organisms adapted to this saline environment are able to thrive in it.

The pH scale is, as previously mentioned, an inverse logarithm and ranges from 0 to 14 (Figure). Anything below 7.0 (ranging from 0.0 to 6.9) is acidic, and anything above 7.0 (from 7.1 to 14.0) is alkaline. Extremes in pH in either direction from 7.0 are usually considered inhospitable to life. The pH inside cells (6.8) and the pH in the blood (7.4) are both very close to neutral. However, the environment in the stomach is highly acidic, with a pH of 1 to 2. So how do the cells of the stomach survive in such an acidic environment? How do they homeostatically maintain the near neutral pH inside them? The answer is that they cannot do it and are constantly dying. New stomach cells are constantly produced to replace dead ones, which are digested by the stomach acids. It is estimated that the lining of the human stomach is completely replaced every seven to ten days.



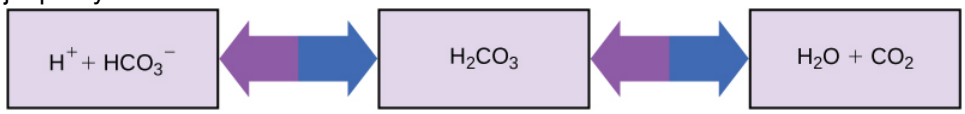
The pH scale measures the concentration of hydrogen ions (H^+) in a solution. (credit: modification of work by Edward Stevens)

Link to Learning

Watch this video for a straightforward explanation of pH and its logarithmic scale.

Visit this page in your course online to use this activity.

So how can organisms whose bodies require a near-neutral pH ingest acidic and basic substances (a human drinking orange juice, for example) and survive? Buffers are the key. Buffers readily absorb excess H^+ or OH^-, keeping the pH of the body carefully maintained in the narrow range required for survival. Maintaining a constant blood pH is critical to a person's well-being. The buffer maintaining the pH of human blood involves carbonic acid (H_2CO_3), bicarbonate ion (HCO_3^-), and carbon dioxide (CO_2). When bicarbonate ions combine with free hydrogen ions and become carbonic acid, hydrogen ions are removed, moderating pH changes. Similarly, as shown in Figure, excess carbonic acid can be converted to carbon dioxide gas and exhaled through the lungs. This prevents too many free hydrogen ions from building up in the blood and dangerously reducing the blood's pH. Likewise, if too much OH^- is introduced into the system, carbonic acid will combine with it to create bicarbonate, lowering the pH. Without this buffer system, the body's pH would fluctuate enough to put survival in jeopardy.



This diagram shows the body's buffering of blood pH levels. The blue arrows show the process of raising pH as more CO_2 is made. The purple arrows indicate the reverse process: the lowering of pH as more bicarbonate is created.

Other examples of buffers are antacids used to combat excess stomach acid. Many of these over-the-counter medications work in the same way as blood buffers, usually with at least one ion capable of absorbing hydrogen and moderating pH, bringing relief to those that suffer "heartburn" after eating. The unique properties of water that contribute to this capacity to balance pH—as well as water's other characteristics—are essential to sustaining life on Earth.

Link to Learning

To learn more about water. Visit the U.S. Geological Survey Water Science for Schools All About Water! website.

Section Summary

Water has many properties that are critical to maintaining life. It is a polar molecule, allowing for the formation of hydrogen bonds. Hydrogen bonds allow ions and other polar molecules to dissolve in water. Therefore, water is an excellent solvent. The hydrogen bonds between water molecules cause the water to have a high heat capacity, meaning it takes a lot of added heat to raise its temperature. As the temperature rises, the hydrogen bonds between water continually break and form anew. This allows for the overall temperature to remain stable, although energy is added to the system. Water also exhibits a high heat of vaporization, which is key to how organisms cool themselves by the evaporation of sweat. Water's cohesive forces allow for the property of surface tension, whereas its adhesive properties are seen as water rises inside capillary tubes. The pH value is a measure of hydrogen ion concentration in a solution and is one of many chemical characteristics that is highly regulated in living organisms through homeostasis. Acids and bases can change pH values, but buffers tend to moderate the changes they cause. These properties of water are intimately connected to the biochemical and physical processes performed by living organisms, and life would be very different if these properties were altered, if it could exist at all.

Review Questions

Which of the following statements is not true?

1. Water is polar.
2. Water stabilizes temperature.
3. Water is essential for life.
4. Water is the most abundant molecule in the Earth's atmosphere.

When acids are added to a solution, the pH should _____.

1. decrease
2. increase
3. stay the same
4. cannot tell without testing

A molecule that binds up excess hydrogen ions in a solution is called a(n) _____.

1. acid
2. isotope
3. base
4. donator

Which of the following statements is true?

1. Acids and bases cannot mix together.
2. Acids and bases will neutralize each other.
3. Acids, but not bases, can change the pH of a solution.
4. Acids donate hydroxide ions (OH^-); bases donate hydrogen ions (H^+).

Free Response

Discuss how buffers help prevent drastic swings in pH.

Why can some insects walk on water?

Footnotes

1. 1 W. Humphrey W., A. Dalke, and K. Schulten, "VMD—Visual Molecular Dynamics," *Journal of Molecular Graphics* 14 (1996): 33-38.
2. 2 W. Humphrey W., A. Dalke, and K. Schulten, "VMD—Visual Molecular Dynamics," *Journal of Molecular Graphics* 14 (1996): 33-38.

Glossary

acid molecule that donates hydrogen ions and increases the concentration of hydrogen ions in a solution

adhesion attraction between water molecules and other molecules

base molecule that donates hydroxide ions or otherwise binds excess hydrogen ions and decreases the concentration of hydrogen ions in a solution

buffer substance that prevents a change in pH by absorbing or releasing hydrogen or hydroxide ions

calorie amount of heat required to change the temperature of one gram of water by one degree Celsius

capillary action occurs because water molecules are attracted to charges on the inner surfaces of narrow tubular structures such as glass tubes, drawing the water molecules to the sides of the tubes

cohesion intermolecular forces between water molecules caused by the polar nature of water; responsible for surface tension

dissociation release of an ion from a molecule such that the original molecule now consists of an ion and the charged remains of the original, such as when water dissociates into H^+ and OH^-

evaporation separation of individual molecules from the surface of a body of water, leaves of a plant, or the skin of an organism

heat of vaporization of water high amount of energy required for liquid water to turn into water vapor

hydrophilic describes ions or polar molecules that interact well with other polar molecules such as water

hydrophobic describes uncharged non-polar molecules that do not interact well with polar molecules such as water

litmus paper (also, pH paper) filter paper that has been treated with a natural water-soluble dye that changes its color as the pH of the environment changes so it can be used as a pH indicator

pH paper see litmus paper

pH scale scale ranging from zero to 14 that is inversely proportional to the concentration of hydrogen ions in a solution

solvent substance capable of dissolving another substance

specific heat capacity the amount of heat one gram of a substance must absorb or lose to change its temperature by one degree Celsius

sphere of hydration when a polar water molecule surrounds charged or polar molecules thus keeping them dissolved and in solution

surface tension tension at the surface of a body of liquid that prevents the molecules from separating; created by the attractive cohesive forces between the molecules of the liquid

Licensing & Attributions

CC licensed content, Shared previously

- water and pH. **Authored by**: open stax. **Provided by**: open stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:pPjfgsd4@10/Water. **License**: *Public Domain: No Known Copyright*

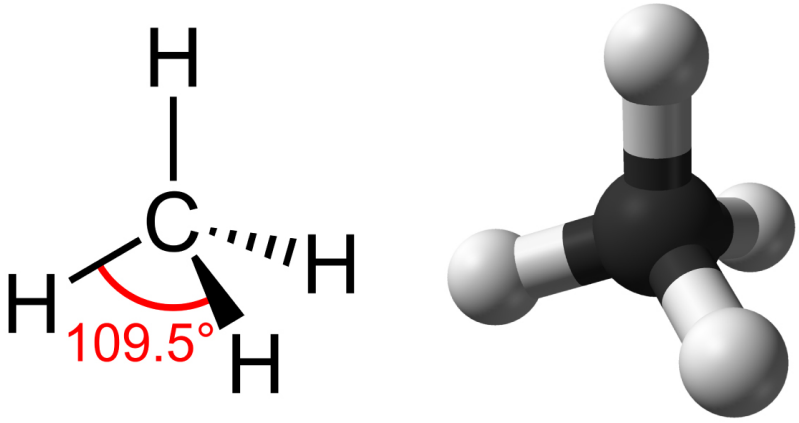
CARBON

Cells are made of many complex molecules called macromolecules, such as proteins, nucleic acids (RNA and DNA), carbohydrates, and lipids. The macromolecules are a subset of organic molecules (any carbon-containing liquid, solid, or gas) that are especially important for life. The fundamental component for all of these macromolecules is carbon. The carbon atom has unique properties that allow it to form covalent bonds to as many as four different atoms, making this versatile element ideal to serve as the basic structural component, or "backbone," of the macromolecules.

Individual carbon atoms have an incomplete outermost electron shell. With an atomic number of 6 (six electrons and six protons), the first two electrons fill the inner shell, leaving four in the second shell. Therefore, carbon atoms can form up to four covalent bonds with other atoms to satisfy the octet rule. The methane molecule provides an example: it has the chemical formula CH_4. Each of its four hydrogen atoms forms a single covalent bond with the carbon atom by sharing a pair of electrons. This results in a filled outermost shell.

Hydrocarbons

Hydrocarbons are organic molecules consisting entirely of carbon and hydrogen, such as methane (CH_4) described above. We often use hydrocarbons in our daily lives as fuels—like the propane in a gas grill or the butane in a lighter. The many covalent bonds between the atoms in hydrocarbons store a great amount of energy, which is released when these molecules are burned (oxidized). Methane, an excellent fuel, is the simplest hydrocarbon molecule, with a central carbon atom bonded to four different hydrogen atoms, as illustrated in Figure. The geometry of the methane molecule, where the atoms reside in three dimensions, is determined by the shape of its electron orbitals. The carbons and the four hydrogen atoms form a shape known as a tetrahedron, with four triangular faces; for this reason, methane is described as having tetrahedral geometry.



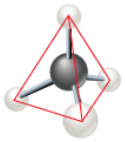
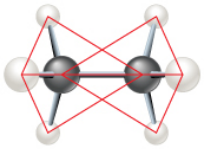
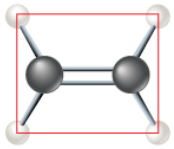
Methane has a tetrahedral geometry, with each of the four hydrogen atoms spaced 109.5° apart.

As the backbone of the large molecules of living things, hydrocarbons may exist as linear carbon chains, carbon rings, or combinations of both. Furthermore, individual carbon-to-carbon bonds may be single, double, or triple covalent bonds, and each type of bond affects the geometry of the molecule in a specific way. This three-dimensional shape or conformation of the large molecules of life (macromolecules) is critical to how they function.

Hydrocarbon Chains

Hydrocarbon chains are formed by successive bonds between carbon atoms and may be branched or unbranched. Furthermore, the overall geometry of the molecule is altered by the different geometries of single, double, and triple covalent bonds, illustrated inFigure. The hydrocarbons ethane, ethene, and ethyne serve as examples of how different carbon-to-carbon bonds affect the geometry of the molecule. The names of all three

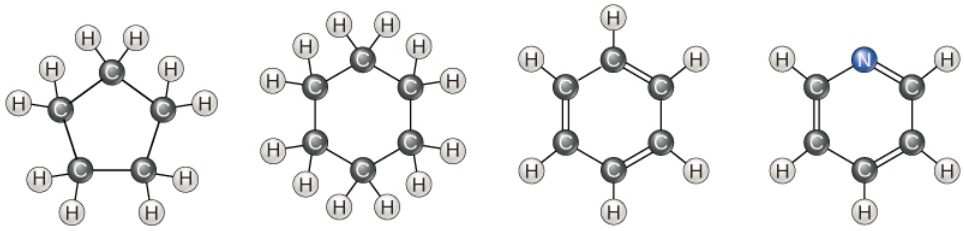
molecules start with the prefix “eth-,” which is the prefix for two carbon hydrocarbons. The suffixes “-ane,” “-ene,” and “-yne” refer to the presence of single, double, or triple carbon-carbon bonds, respectively. Thus, propane, propene, and propyne follow the same pattern with three carbon molecules, butane, butane, and butyne for four carbon molecules, and so on. Double and triple bonds change the geometry of the molecule: single bonds allow rotation along the axis of the bond, whereas double bonds lead to a planar configuration and triple bonds to a linear one. These geometries have a significant impact on the shape a particular molecule can assume.

Methane (CH_4)	Ethane (C_2H_6)	Ethene (C_2H_4)
		
Tetrahedral (single bond)	Tetrahedral (single bond)	Planar (double bond)

When carbon forms single bonds with other atoms, the shape is tetrahedral. When two carbon atoms form a double bond, the shape is planar, or flat. Single bonds, like those found in ethane, are able to rotate. Double bonds, like those found in ethene cannot rotate, so the atoms on either side are locked in place.

Hydrocarbon Rings

So far, the hydrocarbons we have discussed have been aliphatic hydrocarbons, which consist of linear chains of carbon atoms. Another type of hydrocarbon, aromatic hydrocarbons, consists of closed rings of carbon atoms. Ring structures are found in hydrocarbons, sometimes with the presence of double bonds, which can be seen by comparing the structure of cyclohexane to benzene in Figure. Examples of biological molecules that incorporate the benzene ring include some amino acids and cholesterol and its derivatives, including the hormones estrogen and testosterone. The benzene ring is also found in the herbicide 2,4-D. Benzene is a natural component of crude oil and has been classified as a carcinogen. Some hydrocarbons have both aliphatic and aromatic portions; beta-carotene is an example of such a hydrocarbon.



Cyclopentane Cyclohexane Benzene Pyridine

Carbon can form five-and six membered rings. Single or double bonds may connect the carbons in the ring, and nitrogen may be substituted for carbon.

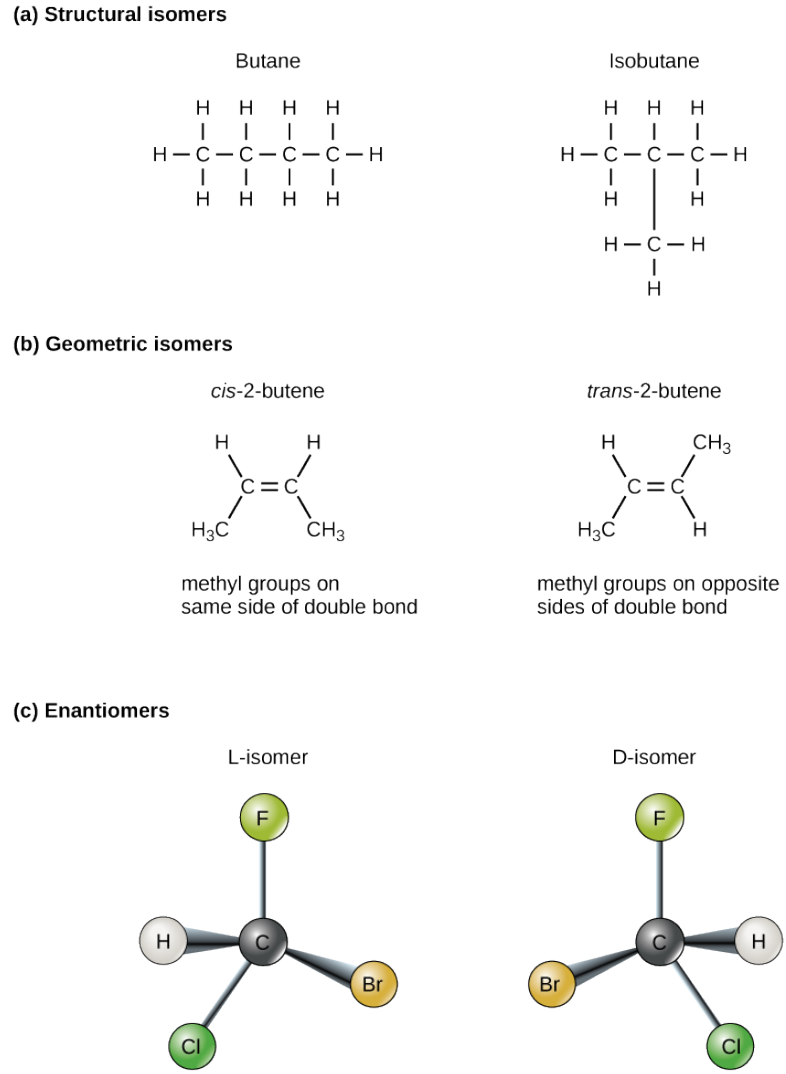
Isomers

The three-dimensional placement of atoms and chemical bonds within organic molecules is central to understanding their chemistry. Molecules that share the same chemical formula but differ in the placement (structure) of their atoms and/or chemical bonds are known as isomers. Structural isomers (like butane and isobutene shown in Figure **a**) differ in the placement of their covalent bonds: both molecules have four carbons and ten hydrogens (C_4H_{10}), but the different arrangement of the atoms within the molecules leads to differences in their chemical properties. For example, due to their different chemical properties, butane is suited for use as a fuel for cigarette lighters and torches, whereas isobutene is suited for use as a refrigerant and a propellant in spray cans.

Geometric isomers, on the other hand, have similar placements of their covalent bonds but differ in how these bonds are made to the surrounding atoms, especially in carbon-to-carbon double bonds. In the simple molecule butene (C_4H_8), the two methyl groups (CH_3) can be on either side of the double covalent bond central to the molecule, as illustrated in Figure **b**. When the carbons are bound on the same side of the double bond, this is the *cis* configuration; if they are on opposite sides of the double bond, it is a *trans*configuration. In the *trans*

configuration, the carbons form a more or less linear structure, whereas the carbons in the *cis*configuration make a bend (change in direction) of the carbon backbone.

ART CONNECTION



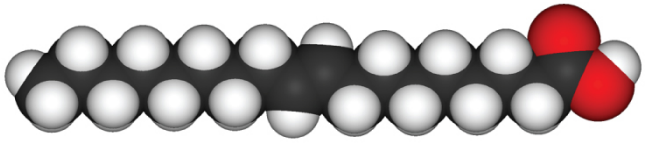
Molecules that have the same number and type of atoms arranged differently are called isomers. (a) Structural isomers have a different covalent arrangement of atoms. (b) Geometric isomers have a different arrangement of atoms around a double bond. (c) Enantiomers are mirror images of each other.

Which of the following statements is false?

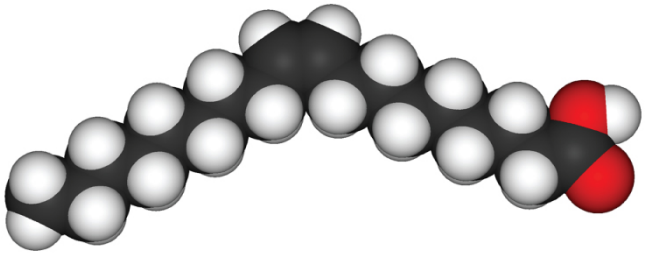
1. Molecules with the formulas CH_3CH_2COOH and $C_3H_6O_2$ could be structural isomers.
2. Molecules must have a double bond to be *cis–trans* isomers.
3. To be enantiomers, a molecule must have at least three different atoms or groups connected to a central carbon.
4. To be enantiomers, a molecule must have at least four different atoms or groups connected to a central carbon.

In triglycerides (fats and oils), long carbon chains known as fatty acids may contain double bonds, which can be in either the *cis* or*trans* configuration, illustrated in Figure. Fats with at least one double bond between carbon atoms are unsaturated fats. When some of these bonds are in the *cis* configuration, the resulting bend in the carbon backbone of the chain means that triglyceride molecules cannot pack tightly, so they remain liquid (oil) at room temperature. On the other hand, triglycerides with *trans* double bonds (popularly called trans fats), have relatively linear fatty acids that are able to pack tightly together at room temperature and form solid fats. In the human diet, trans fats are linked to an increased risk of cardiovascular disease, so many food manufacturers have reduced or eliminated their use in recent years. In contrast to unsaturated fats, triglycerides without double bonds between

carbon atoms are called saturated fats, meaning that they contain all the hydrogen atoms available. Saturated fats are a solid at room temperature and usually of animal origin.



Eliadic acid



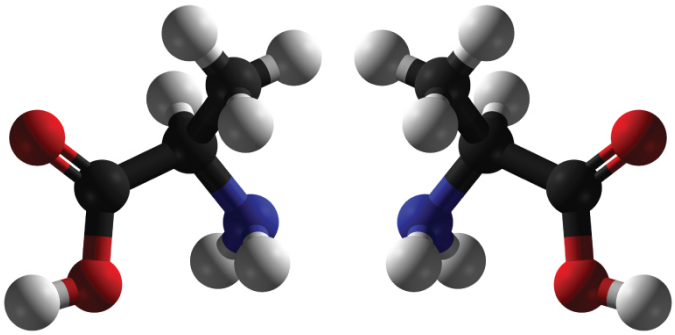
Oleic acid

These space-filling models show a *cis* (oleic acid) and a *trans* (eliadic acid) fatty acid. Notice the bend in the molecule cause by the *cis* configuration.

Enantiomers

Enantiomers are molecules that share the same chemical structure and chemical bonds but differ in the three-dimensional placement of atoms so that they are mirror images. As shown in Figure, an amino acid alanine example, the two structures are non-superimposable. In nature, only the L-forms of amino acids are used to make proteins. Some D forms of amino acids are seen in the cell walls of bacteria, but never in their proteins. Similarly, the D-form of glucose is the main product of photosynthesis and the L-form of the molecule is rarely seen in nature.

D-alanine L-alanine



D-alanine and L-alanine are examples of enantiomers or mirror images. Only the L-forms of amino acids are used to make proteins.

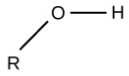
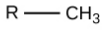
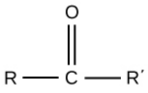
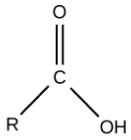
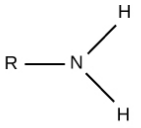
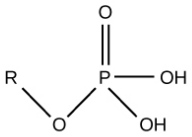
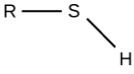
Functional Groups

Functional groups are groups of atoms that occur within molecules and confer specific chemical properties to those molecules. They are found along the “carbon backbone” of macromolecules. This carbon backbone is formed by chains and/or rings of carbon atoms with the occasional substitution of an element such as nitrogen or oxygen. Molecules with other elements in their carbon backbone are substituted hydrocarbons.

The functional groups in a macromolecule are usually attached to the carbon backbone at one or several different places along its chain and/or ring structure. Each of the four types of macromolecules—proteins, lipids,

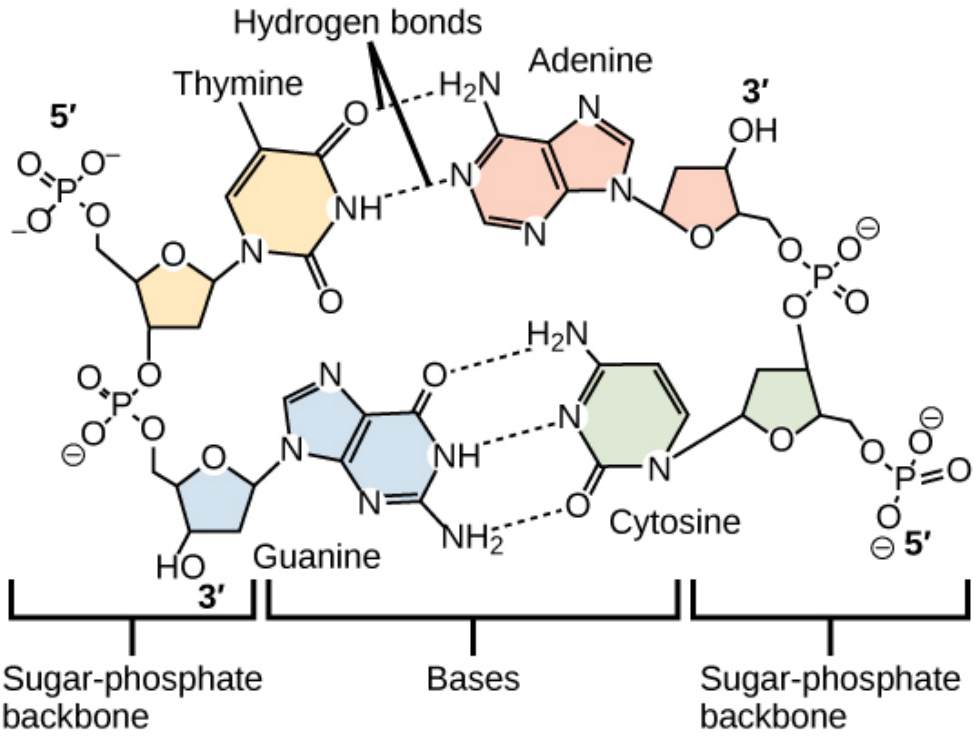
carbohydrates, and nucleic acids—has its own characteristic set of functional groups that contributes greatly to its differing chemical properties and its function in living organisms.

A functional group can participate in specific chemical reactions. Some of the important functional groups in biological molecules are shown in Figure; they include: hydroxyl, methyl, carbonyl, carboxyl, amino, phosphate, and sulfhydryl. These groups play an important role in the formation of molecules like DNA, proteins, carbohydrates, and lipids. Functional groups are usually classified as hydrophobic or hydrophilic depending on their charge or polarity characteristics. An example of a hydrophobic group is the non-polar methane molecule. Among the hydrophilic functional groups is the carboxyl group found in amino acids, some amino acid side chains, and the fatty acids that form triglycerides and phospholipids. This carboxyl group ionizes to release hydrogen ions (H^+) from the COOH group resulting in the negatively charged COO^- group; this contributes to the hydrophilic nature of whatever molecule it is found on. Other functional groups, such as the carbonyl group, have a partially negatively charged oxygen atom that may form hydrogen bonds with water molecules, again making the molecule more hydrophilic.

Functional Group	Structure	Properties
Hydroxyl		Polar
Methyl		Nonpolar
Carbonyl		Polar
Carboxyl		Charged, ionizes to release H^+. Since carboxyl groups can release H^+ ions into solution, they are considered acidic.
Amino		Charged, accepts H^+ to form NH_3^+. Since amino groups can remove H^+ from solution, they are considered basic.
Phosphate		Charged, ionizes to release H^+. Since phosphate groups can release H^+ ions into solution, they are considered acidic.
Sulfhydryl		Polar

The functional groups shown here are found in many different biological molecules.

Hydrogen bonds between functional groups (within the same molecule or between different molecules) are important to the function of many macromolecules and help them to fold properly into and maintain the appropriate shape for functioning. Hydrogen bonds are also involved in various recognition processes, such as DNA complementary base pairing and the binding of an enzyme to its substrate, as illustrated in Figure.



Hydrogen bonds connect two strands of DNA together to create the double-helix structure.

Section Summary

The unique properties of carbon make it a central part of biological molecules. Carbon binds to oxygen, hydrogen, and nitrogen covalently to form the many molecules important for cellular function. Carbon has four electrons in its outermost shell and can form four bonds. Carbon and hydrogen can form hydrocarbon chains or rings. Functional groups are groups of atoms that confer specific properties to hydrocarbon (or substituted hydrocarbon) chains or rings that define their overall chemical characteristics and function.

Art Connections

Figure Which of the following statements is false?

1. Molecules with the formulas CH_3CH_2COOH and $C_3H_6O_2$ could be structural isomers.
2. Molecules must have a double bond to be *cis–trans* isomers.
3. To be enantiomers, a molecule must have at least three different atoms or groups connected to a central carbon.
4. To be enantiomers, a molecule must have at least four different atoms or groups connected to a central carbon.

Review Questions

Each carbon molecule can bond with as many as_____ other atom(s) or molecule(s).

1. one
2. two
3. six
4. four

Which of the following is not a functional group that can bond with carbon?

1. sodium
2. hydroxyl
3. phosphate
4. carbonyl

Free Response

What property of carbon makes it essential for organic life?

Compare and contrast saturated and unsaturated triglycerides.

Glossary

aliphatic hydrocarbon hydrocarbon consisting of a linear chain of carbon atoms

aromatic hydrocarbon hydrocarbon consisting of closed rings of carbon atoms

enantiomers molecules that share overall structure and bonding patterns, but differ in how the atoms are three dimensionally placed such that they are mirror images of each other

functional group group of atoms that provides or imparts a specific function to a carbon skeleton

geometric isomer isomer with similar bonding patterns differing in the placement of atoms alongside a double covalent bond

hydrocarbon molecule that consists only of carbon and hydrogen

isomers molecules that differ from one another even though they share the same chemical formula

organic molecule any molecule containing carbon (except carbon dioxide)

structural isomers molecules that share a chemical formula but differ in the placement of their chemical bonds

substituted hydrocarbon hydrocarbon chain or ring containing an atom of another element in place of one of the backbone carbons

Licensing & Attributions

CC licensed content, Shared previously

- Carbon. **Authored by**: open stax. **Provided by**: open stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:Ur_l0X-K@6/Carbon. **License**: *CC BY: Attribution*

ACTIVITY: BUILD AN ATOM

Click on the link below to build an atom out of protons, neutrons, and electrons, and see how the element, charge, and mass change. Then play a game to test your ideas!

Build an Atom

Licensing & Attributions

CC licensed content, Shared previously

- Activity: Build an Atom. **Authored by**: Dr. David Fernandez, Dr. Leslie Orzett. **Provided by**: NOVA. **Located at**: http://www.nvcc.edu. **Project**: Kaleidoscope. **License**: *CC BY: Attribution*

STUDY GUIDE: CHEMISTRY

Study Questions

Objective: Understand Chemistry for biology students.

Use this page to check your understanding of the content.

Vocabulary

1. element
2. atom
3. atomic number
4. atomic mass
5. proton
6. neutron
7. electron
8. orbitals/electron shell
9. valence electrons
10. chemical compound
11. chemical formula
12. molecular formula
13. structural formula
14. mole
15. covalent bonds
16. polar covalent bonds
17. ionic bond
18. anion
19. cation
20. hydrogen bond
21. oxidation and reduction
22. polar molecule

Study guide

1. Be able to draw a Bohr model of a single atom
2. Be able to draw a Bohr model of an ionic bond and a covalent bond
3. Know what type of bond different atoms will participate in due to their electron configuration
4. Be able to calculate the number of grams/mole of a substance
5. Be able to calculate the charge or atomic mass of an atom

Licensing & Attributions

CC licensed content, Original

- **Authored by**: Danielle McGrath. **Provided by**: San Jacinto College. **License**: *CC BY: Attribution*

STUDY GUIDE: WATER

Study Questions

Objective: Relate hydrogen ion concentration to pH. Relate the characteristics of water to processes critical for life.

Use this page to check your understanding of the content.

Vocabulary

1. Kinetic energy
2. Polar molecule
3. Hydrophobic
4. Hydrophilic
5. Solvent
6. Solute
7. Surface tension
8. pH
9. Concentration
10. 1 mole

Study Guide

1. Explain the relationship between the kinetic energy of molecules or atoms in a system, and temperature.
2. Explain what it means when someone says, “Water is a polar molecule.” (Bet you hear that all the time!)
3. What are hydrogen bonds?
4. Why is it difficult to heat up (or cool down) water?
5. Why is water such a good solvent? Are there any substances water cannot dissolve? Why or why not?
6. What is surface tension? Give some biological examples of why surface tension in water is important.
7. What is a surfactant?
8. Be able to clearly relate pH and hydrogen ion concentration.
9. What exactly does a pH of 4 mean?
10. What impact does changing pH have on proteins?
11. What are buffers? What do they do? Why are they important?

Licensing & Attributions

CC licensed content, Shared previously

- Study Guide: Intro to Chemistry. **Authored by**: Wendy Riggs. **Provided by**: College of the Redwoods. **Project**: Kaleidoscope. **License**: *CC BY: Attribution*

BIOLOGICAL MACROMOLECULES

INTRODUCTION

Chapter 3. Introduction

Foods such as bread, fruit, and cheese are rich sources of biological macromolecules. (credit: modification of work by Bengt Nyman)

Food provides the body with the nutrients it needs to survive. Many of these critical nutrients are biological macromolecules, or large molecules, necessary for life. These macromolecules (polymers) are built from different combinations of smaller organic molecules (monomers). What specific types of biological macromolecules do living things require? How are these molecules formed? What functions do they serve? In this chapter, these questions will be explored.

SYNTHESIS OF BIOLOGICAL MACROMOLECULES

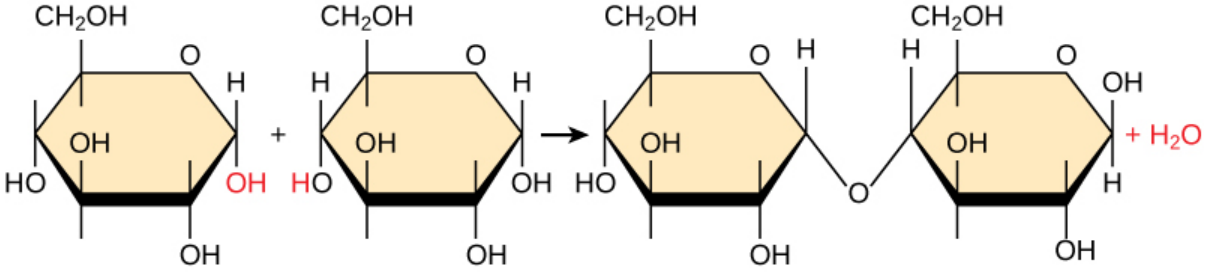
3.1 Synthesis of Biological Macromolecules

Summary

As you've learned, biological macromolecules are large molecules, necessary for life, that are built from smaller organic molecules. There are four major classes of biological macromolecules (carbohydrates, lipids, proteins, and nucleic acids); each is an important cell component and performs a wide array of functions. Combined, these molecules make up the majority of a cell's dry mass (recall that water makes up the majority of its complete mass). Biological macromolecules are organic, meaning they contain carbon. In addition, they may contain hydrogen, oxygen, nitrogen, and additional minor elements.

Dehydration Synthesis

Most macromolecules are made from single subunits, or building blocks, called monomers. The monomers combine with each other using covalent bonds to form larger molecules known as polymers. In doing so, monomers release water molecules as byproducts. This type of reaction is known as dehydration synthesis, which means "to put together while losing water."

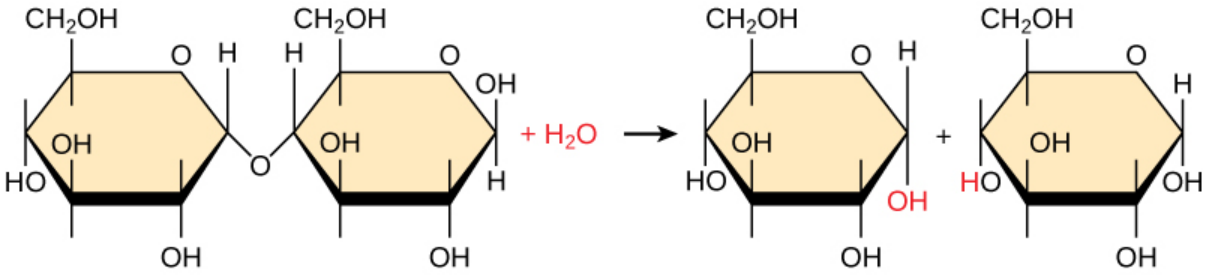


In the dehydration synthesis reaction depicted above, two molecules of glucose are linked together to form the disaccharide maltose. In the process, a water molecule is formed.

In a dehydration synthesis reaction (Figure), the hydrogen of one monomer combines with the hydroxyl group of another monomer, releasing a molecule of water. At the same time, the monomers share electrons and form covalent bonds. As additional monomers join, this chain of repeating monomers forms a polymer. Different types of monomers can combine in many configurations, giving rise to a diverse group of macromolecules. Even one kind of monomer can combine in a variety of ways to form several different polymers: for example, glucose monomers are the constituents of starch, glycogen, and cellulose.

Hydrolysis

Polymers are broken down into monomers in a process known as hydrolysis, which means "to split water," a reaction in which a water molecule is used during the breakdown (Figure). During these reactions, the polymer is broken into two components: one part gains a hydrogen atom (H+) and the other gains a hydroxyl molecule (OH–) from a split water molecule.



In the hydrolysis reaction shown here, the disaccharide maltose is broken down to form two glucose monomers with the addition of a water molecule. Note that this reaction is the reverse of the synthesis reaction shown in Figure.

Dehydration and hydrolysis reactions are catalyzed, or "sped up," by specific enzymes; dehydration reactions involve the formation of new bonds, requiring energy, while hydrolysis reactions break bonds and release energy. These reactions are similar for most macromolecules, but each monomer and polymer reaction is specific for its class. For example, in our bodies, food is hydrolyzed, or broken down, into smaller molecules by catalytic enzymes in the digestive system. This allows for easy absorption of nutrients by cells in the intestine. Each macromolecule is broken down by a specific enzyme. For instance, carbohydrates are broken down by amylase, sucrase, lactase, or maltase. Proteins are broken down by the enzymes pepsin and peptidase, and by hydrochloric acid. Lipids are broken down by lipases. Breakdown of these macromolecules provides energy for cellular activities.

Link to Learning

Visit this site to see visual representations of dehydration synthesis and hydrolysis.

Section Summary

Proteins, carbohydrates, nucleic acids, and lipids are the four major classes of biological macromolecules—large molecules necessary for life that are built from smaller organic molecules. Macromolecules are made up of single units known as monomers that are joined by covalent bonds to form larger polymers. The polymer is more than the sum of its parts: it acquires new characteristics, and leads to an osmotic pressure that is much lower than that formed by its ingredients; this is an important advantage in the maintenance of cellular osmotic conditions. A monomer joins with another monomer with the release of a water molecule, leading to the formation of a covalent bond. These types of reactions are known as dehydration or condensation reactions. When polymers are broken down into smaller units (monomers), a molecule of water is used for each bond broken by these reactions; such reactions are known as hydrolysis reactions. Dehydration and hydrolysis reactions are similar for all macromolecules, but each monomer and polymer reaction is specific to its class. Dehydration reactions typically require an investment of energy for new bond formation, while hydrolysis reactions typically release energy by breaking bonds.

Review Questions

Dehydration synthesis leads to formation of

1. monomers
2. polymers
3. water and polymers
4. none of the above

During the breakdown of polymers, which of the following reactions takes place?

1. hydrolysis
2. dehydration
3. condensation

4. covalent bond

Free Response

Why are biological macromolecules considered organic?

What role do electrons play in dehydration synthesis and hydrolysis?

Glossary

biological macromolecule large molecule necessary for life that is built from smaller organic molecules

dehydration synthesis (also, condensation) reaction that links monomer molecules together, releasing a molecule of water for each bond formed

hydrolysis reaction causes breakdown of larger molecules into smaller molecules with the utilization of water

monomer smallest unit of larger molecules called polymers

polymer chain of monomer residues that is linked by covalent bonds; polymerization is the process of polymer formation from monomers by condensation

Licensing & Attributions

CC licensed content, Shared previously

- synthesis of biological macromolecules. **Authored by**: open stax. **Provided by**: open stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:6kS4-uei@12/Synthesis-of-Biological-Macrom. **License**: *CC BY: Attribution*

CARBOHYDRATES

3.2 Carbohydrates

Summary

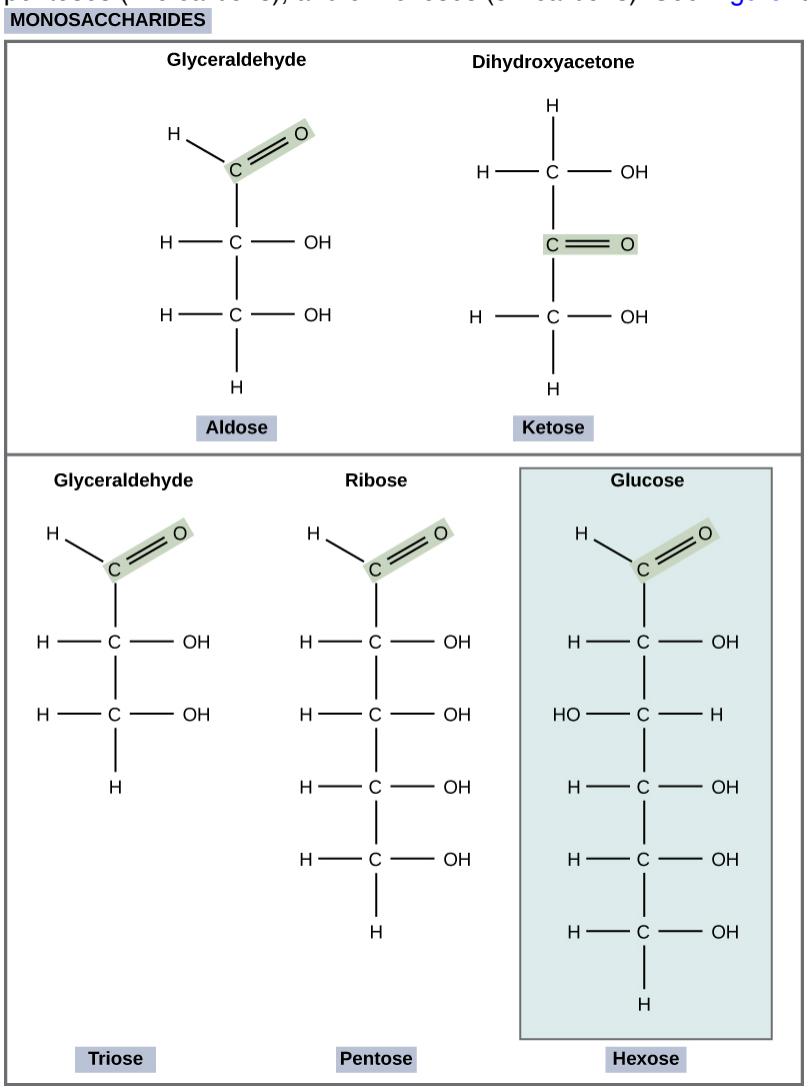
Most people are familiar with carbohydrates, one type of macromolecule, especially when it comes to what we eat. To lose weight, some individuals adhere to “low-carb” diets. Athletes, in contrast, often “carb-load” before important competitions to ensure that they have enough energy to compete at a high level. Carbohydrates are, in fact, an essential part of our diet; grains, fruits, and vegetables are all natural sources of carbohydrates. Carbohydrates provide energy to the body, particularly through glucose, a simple sugar that is a component of starch and an ingredient in many staple foods. Carbohydrates also have other important functions in humans, animals, and plants.

Molecular Structures

Carbohydrates can be represented by the stoichiometric formula $(CH_2O)_n$, where n is the number of carbons in the molecule. In other words, the ratio of carbon to hydrogen to oxygen is 1:2:1 in carbohydrate molecules. This formula also explains the origin of the term “carbohydrate”: the components are carbon (“carbo”) and the components of water (hence, “hydrate”). Carbohydrates are classified into three subtypes: monosaccharides, disaccharides, and polysaccharides.

Monosaccharides

Monosaccharides (mono- = “one”; sacchar- = “sweet”) are simple sugars, the most common of which is glucose. In monosaccharides, the number of carbons usually ranges from three to seven. Most monosaccharide names end with the suffix -ose. If the sugar has an aldehyde group (the functional group with the structure R-CHO), it is known as an aldose, and if it has a ketone group (the functional group with the structure RC(=O)R’), it is known as a ketose. Depending on the number of carbons in the sugar, they also may be known as trioses (three carbons), pentoses (five carbons), and or hexoses (six carbons). See Figure for an illustration of the monosaccharides.



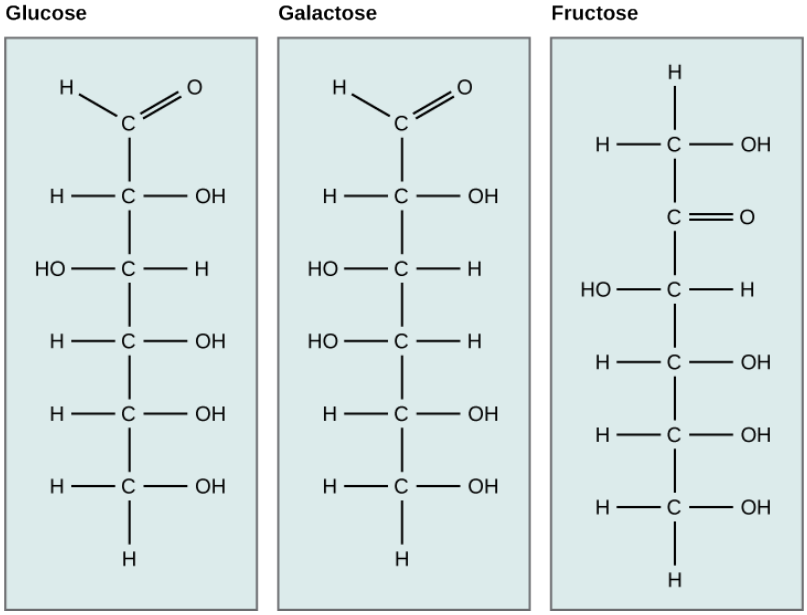
Monosaccharides are classified based on the position of their carbonyl group and the number of carbons in the backbone. Aldoses have a carbonyl group (indicated in green) at the end of the carbon chain, and ketoses have a carbonyl group in the middle of the carbon chain. Trioses, pentoses, and hexoses have three, five, and six carbon backbones, respectively.

The chemical formula for glucose is $C_6H_{12}O_6$. In humans, glucose is an important source of energy. During cellular respiration, energy is released from glucose, and that energy is used to help make adenosine triphosphate (ATP). Plants synthesize glucose using carbon dioxide and water, and glucose in turn is used for energy requirements for the plant. Excess glucose is often stored as starch that is catabolized (the breakdown of larger molecules by cells) by humans and other animals that feed on plants.

Galactose (part of lactose, or milk sugar) and fructose (found in sucrose, in fruit) are other common monosaccharides. Although glucose, galactose, and fructose all have the same chemical formula ($C_6H_{12}O_6$), they differ structurally and chemically (and are known as isomers) because of the different arrangement of functional

groups around the asymmetric carbon; all of these monosaccharides have more than one asymmetric carbon (Figure).

ART CONNECTION

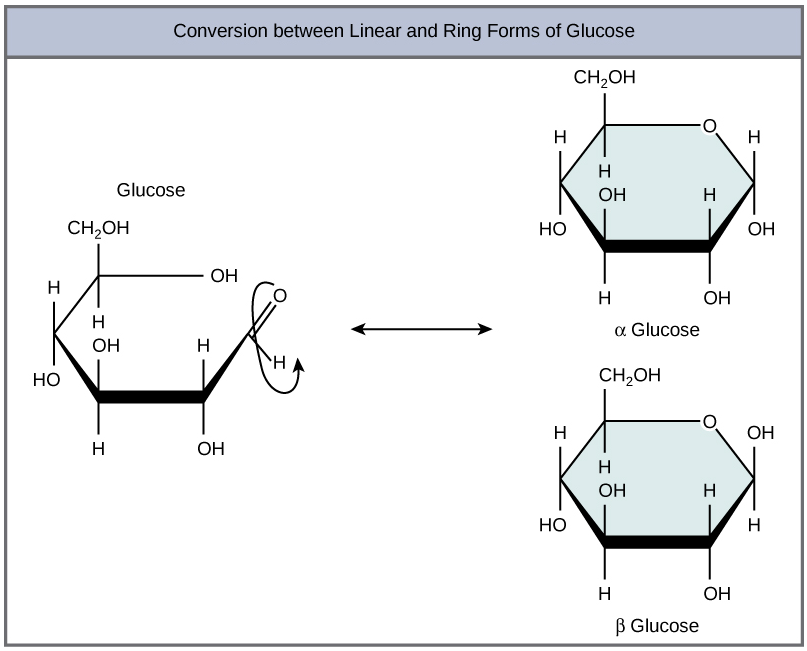


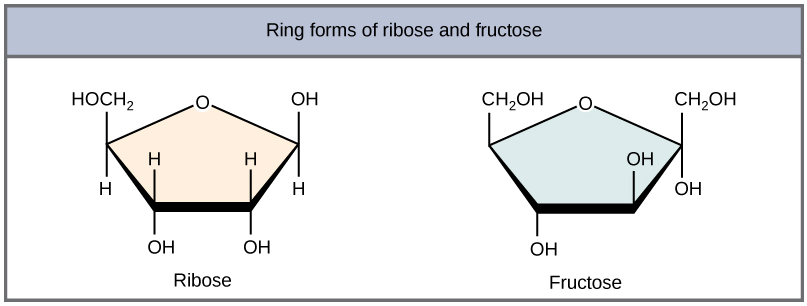
Glucose, galactose, and fructose are all hexoses. They are structural isomers, meaning they have the same chemical formula ($C_6H_{12}O_6$) but a different arrangement of atoms.

What kind of sugars are these, aldose or ketose?

Glucose, galactose, and fructose are isomeric monosaccharides (hexoses), meaning they have the same chemical formula but have slightly different structures. Glucose and galactose are aldoses, and fructose is a ketose.

Monosaccharides can exist as a linear chain or as ring-shaped molecules; in aqueous solutions they are usually found in ring forms (Figure). Glucose in a ring form can have two different arrangements of the hydroxyl group (OH) around the anomeric carbon (carbon 1 that becomes asymmetric in the process of ring formation). If the hydroxyl group is below carbon number 1 in the sugar, it is said to be in the alpha (α) position, and if it is above the plane, it is said to be in the beta (β) position.

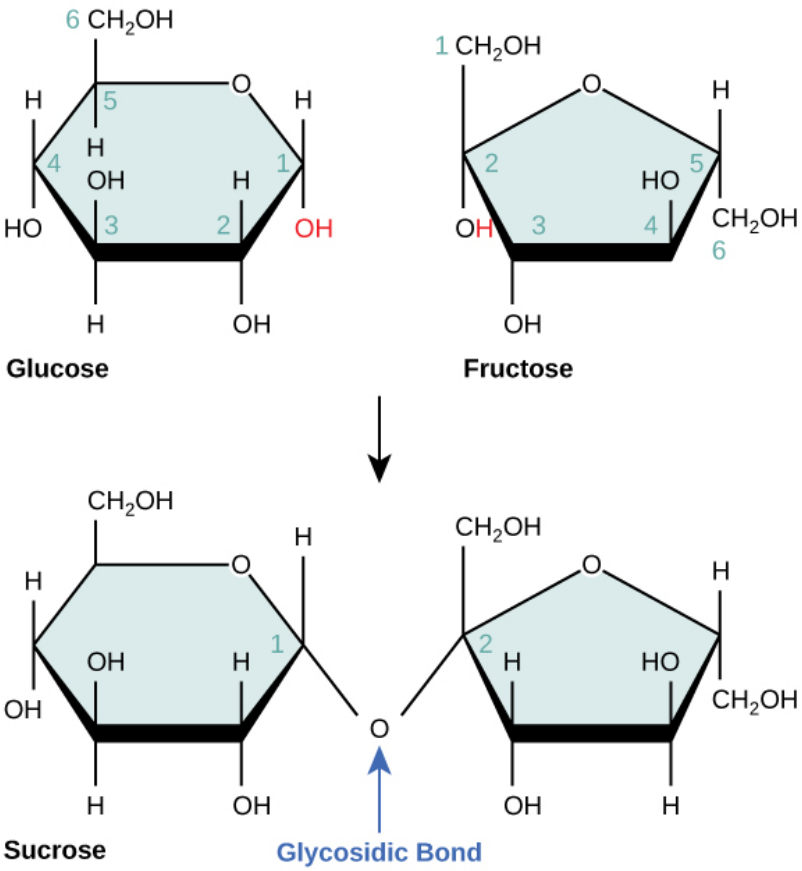




Five and six carbon monosaccharides exist in equilibrium between linear and ring forms. When the ring forms, the side chain it closes on is locked into an *α* or *β* position. Fructose and ribose also form rings, although they form five-membered rings as opposed to the six-membered ring of glucose.

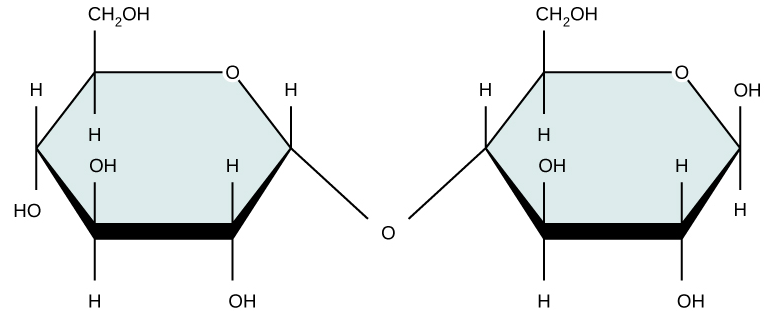
Disaccharides

Disaccharides (di- = “two”) form when two monosaccharides undergo a dehydration reaction (also known as a condensation reaction or dehydration synthesis). During this process, the hydroxyl group of one monosaccharide combines with the hydrogen of another monosaccharide, releasing a molecule of water and forming a covalent bond. A covalent bond formed between a carbohydrate molecule and another molecule (in this case, between two monosaccharides) is known as a glycosidic bond (Figure). Glycosidic bonds (also called glycosidic linkages) can be of the alpha or the beta type.

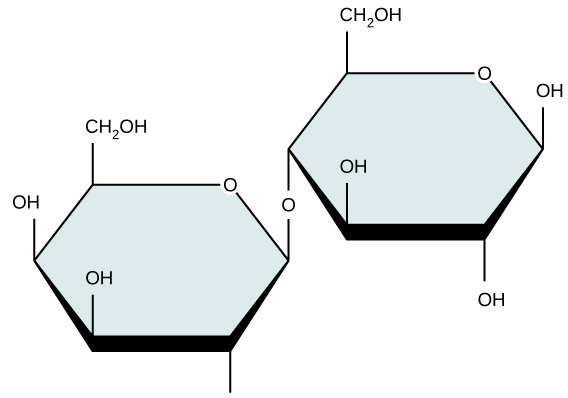


Sucrose is formed when a monomer of glucose and a monomer of fructose are joined in a dehydration reaction to form a glycosidic bond. In the process, a water molecule is lost. By convention, the carbon atoms in a monosaccharide are numbered from the terminal carbon closest to the carbonyl group. In sucrose, a glycosidic linkage is formed between carbon 1 in glucose and carbon 2 in fructose.

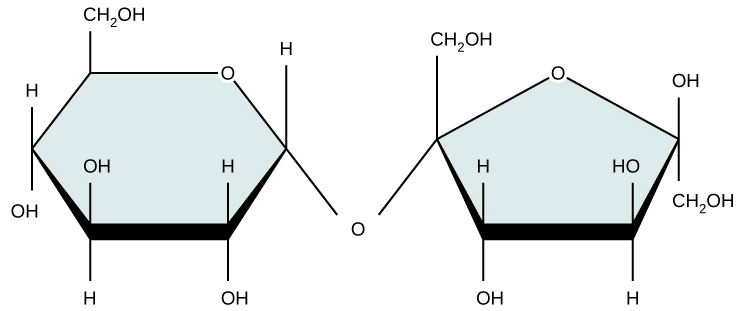
Common disaccharides include lactose, maltose, and sucrose (Figure). Lactose is a disaccharide consisting of the monomers glucose and galactose. It is found naturally in milk. Maltose, or malt sugar, is a disaccharide formed by a dehydration reaction between two glucose molecules. The most common disaccharide is sucrose, or table sugar, which is composed of the monomers glucose and fructose.



Maltose



Lactose



Sucrose

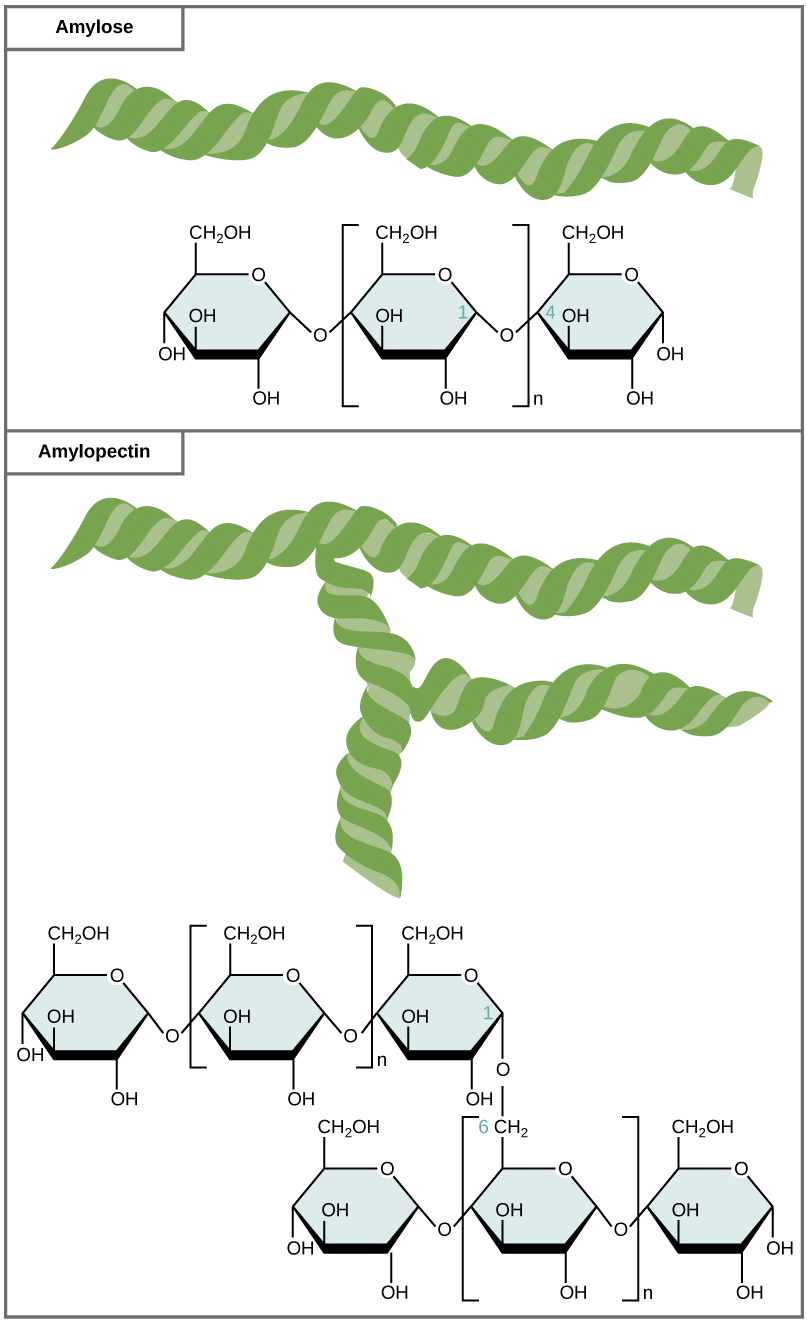
Common disaccharides include maltose (grain sugar), lactose (milk sugar), and sucrose (table sugar).

Polysaccharides

A long chain of monosaccharides linked by glycosidic bonds is known as a polysaccharide (poly- = “many”). The chain may be branched or unbranched, and it may contain different types of monosaccharides. The molecular weight may be 100,000 daltons or more depending on the number of monomers joined. Starch, glycogen, cellulose, and chitin are primary examples of polysaccharides.

Starch is the stored form of sugars in plants and is made up of a mixture of amylose and amylopectin (both polymers of glucose). Plants are able to synthesize glucose, and the excess glucose, beyond the plant’s immediate energy needs, is stored as starch in different plant parts, including roots and seeds. The starch in the seeds provides food for the embryo as it germinates and can also act as a source of food for humans and animals. The starch that is consumed by humans is broken down by enzymes, such as salivary amylases, into smaller molecules, such as maltose and glucose. The cells can then absorb the glucose.

Starch is made up of glucose monomers that are joined by *α* 1-4 or *α* 1-6 glycosidic bonds. The numbers 1-4 and 1-6 refer to the carbon number of the two residues that have joined to form the bond. As illustrated in Figure, amylose is starch formed by unbranched chains of glucose monomers (only *α* 1-4 linkages), whereas amylopectin is a branched polysaccharide (*α* 1-6 linkages at the branch points).



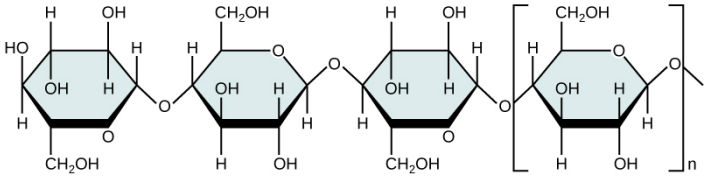
Amylose and amylopectin are two different forms of starch. Amylose is composed of unbranched chains of glucose monomers connected by *α* 1,4 glycosidic linkages. Amylopectin is composed of branched chains of glucose monomers connected by *α* 1,4 and *α* 1,6 glycosidic linkages. Because of the way the subunits are joined, the glucose chains have a helical structure. Glycogen (not shown) is similar in structure to amylopectin but more highly branched.

Glycogen is the storage form of glucose in humans and other vertebrates and is made up of monomers of glucose. Glycogen is the animal equivalent of starch and is a highly branched molecule usually stored in liver and muscle cells. Whenever blood glucose levels decrease, glycogen is broken down to release glucose in a process known as glycogenolysis.

Cellulose is the most abundant natural biopolymer. The cell wall of plants is mostly made of cellulose; this provides structural support to the cell. Wood and paper are mostly cellulosic in nature. Cellulose is made up of glucose monomers that are linked by *β*1-4 glycosidic bonds (Figure).

Cellulose fibers

Cellulose structure



In cellulose, glucose monomers are linked in unbranched chains by β 1-4 glycosidic linkages. Because of the way the glucose subunits are joined, every glucose monomer is flipped relative to the next one resulting in a linear, fibrous structure.

As shown in Figure, every other glucose monomer in cellulose is flipped over, and the monomers are packed tightly as extended long chains. This gives cellulose its rigidity and high tensile strength—which is so important to plant cells. While the β 1-4 linkage cannot be broken down by human digestive enzymes, herbivores such as cows, koalas, buffalos, and horses are able, with the help of the specialized flora in their stomach, to digest plant material that is rich in cellulose and use it as a food source. In these animals, certain species of bacteria and protists reside in the rumen (part of the digestive system of herbivores) and secrete the enzyme cellulase. The appendix of grazing animals also contains bacteria that digest cellulose, giving it an important role in the digestive systems of ruminants. Cellulases can break down cellulose into glucose monomers that can be used as an energy source by the animal. Termites are also able to break down cellulose because of the presence of other organisms in their bodies that secrete cellulases.

Carbohydrates serve various functions in different animals. Arthropods (insects, crustaceans, and others) have an outer skeleton, called the exoskeleton, which protects their internal body parts (as seen in the bee in Figure). This exoskeleton is made of the biological macromolecule chitin, which is a polysaccharide-containing nitrogen. It is made of repeating units of N-acetyl-β-d-glucosamine, a modified sugar. Chitin is also a major component of fungal cell walls; fungi are neither animals nor plants and form a kingdom of their own in the domain Eukarya.

Insects have a hard outer exoskeleton made of chitin, a type of polysaccharide. (credit: Louise Docker)
CAREER CONNECTIONS

Registered DietitianObesity is a worldwide health concern, and many diseases such as diabetes and heart disease are becoming more prevalent because of obesity. This is one of the reasons why registered dietitians are increasingly sought after for advice. Registered dietitians help plan nutrition programs for individuals in various settings. They often work with patients in health care facilities, designing nutrition plans to treat and prevent

diseases. For example, dietitians may teach a patient with diabetes how to manage blood sugar levels by eating the correct types and amounts of carbohydrates. Dietitians may also work in nursing homes, schools, and private practices.

To become a registered dietitian, one needs to earn at least a bachelor's degree in dietetics, nutrition, food technology, or a related field. In addition, registered dietitians must complete a supervised internship program and pass a national exam. Those who pursue careers in dietetics take courses in nutrition, chemistry, biochemistry, biology, microbiology, and human physiology. Dietitians must become experts in the chemistry and physiology (biological functions) of food (proteins, carbohydrates, and fats).

Benefits of Carbohydrates

Are carbohydrates good for you? People who wish to lose weight are often told that carbohydrates are bad for them and should be avoided. Some diets completely forbid carbohydrate consumption, claiming that a low-carbohydrate diet helps people to lose weight faster. However, carbohydrates have been an important part of the human diet for thousands of years; artifacts from ancient civilizations show the presence of wheat, rice, and corn in our ancestors' storage areas.

Carbohydrates should be supplemented with proteins, vitamins, and fats to be parts of a well-balanced diet. Calorie-wise, a gram of carbohydrate provides 4.3 Kcal. For comparison, fats provide 9 Kcal/g, a less desirable ratio. Carbohydrates contain soluble and insoluble elements; the insoluble part is known as fiber, which is mostly cellulose. Fiber has many uses; it promotes regular bowel movement by adding bulk, and it regulates the rate of consumption of blood glucose. Fiber also helps to remove excess cholesterol from the body: fiber binds to the cholesterol in the small intestine, then attaches to the cholesterol and prevents the cholesterol particles from entering the bloodstream, and then cholesterol exits the body via the feces. Fiber-rich diets also have a protective role in reducing the occurrence of colon cancer. In addition, a meal containing whole grains and vegetables gives a feeling of fullness. As an immediate source of energy, glucose is broken down during the process of cellular respiration, which produces ATP, the energy currency of the cell. Without the consumption of carbohydrates, the availability of "instant energy" would be reduced. Eliminating carbohydrates from the diet is not the best way to lose weight. A low-calorie diet that is rich in whole grains, fruits, vegetables, and lean meat, together with plenty of exercise and plenty of water, is the more sensible way to lose weight.

Link to Learning

For an additional perspective on carbohydrates, explore "Biomolecules: the Carbohydrates" through thisinteractive animation.

Section Summary

Carbohydrates are a group of macromolecules that are a vital energy source for the cell and provide structural support to plant cells, fungi, and all of the arthropods that include lobsters, crabs, shrimp, insects, and spiders. Carbohydrates are classified as monosaccharides, disaccharides, and polysaccharides depending on the number of monomers in the molecule. Monosaccharides are linked by glycosidic bonds that are formed as a result of dehydration reactions, forming disaccharides and polysaccharides with the elimination of a water molecule for each bond formed. Glucose, galactose, and fructose are common monosaccharides, whereas common disaccharides include lactose, maltose, and sucrose. Starch and glycogen, examples of polysaccharides, are the storage forms of glucose in plants and animals, respectively. The long polysaccharide chains may be branched or unbranched. Cellulose is an example of an unbranched polysaccharide, whereas amylopectin, a constituent of starch, is a highly branched molecule. Storage of glucose, in the form of polymers like starch of glycogen, makes it slightly less accessible for metabolism; however, this prevents it from leaking out of the cell or creating a high osmotic pressure that could cause excessive water uptake by the cell.

Art Connections

Figure What kind of sugars are these, aldose or ketose?

Review Questions

An example of a monosaccharide is _____.

1. fructose
2. glucose
3. galactose
4. all of the above

Cellulose and starch are examples of:

1. monosaccharides
2. disaccharides
3. lipids
4. polysaccharides

Plant cell walls contain which of the following in abundance?

1. starch
2. cellulose
3. glycogen
4. lactose

Lactose is a disaccharide formed by the formation of a _____ bond between glucose and _____.

1. glycosidic; lactose
2. glycosidic; galactose
3. hydrogen; sucrose
4. hydrogen; fructose

Free Response

Describe the similarities and differences between glycogen and starch.

Why is it impossible for humans to digest food that contains cellulose?

Glossary

carbohydrate biological macromolecule in which the ratio of carbon to hydrogen and to oxygen is 1:2:1; carbohydrates serve as energy sources and structural support in cells and form the a cellular exoskeleton of arthropods

cellulose polysaccharide that makes up the cell wall of plants; provides structural support to the cell

chitin type of carbohydrate that forms the outer skeleton of all arthropods that include crustaceans and insects; it also forms the cell walls of fungi

disaccharide two sugar monomers that are linked together by a glycosidic bond

glycogen storage carbohydrate in animals

glycosidic bond bond formed by a dehydration reaction between two monosaccharides with the elimination of a water molecule

monosaccharide single unit or monomer of carbohydrates

polysaccharide long chain of monosaccharides; may be branched or unbranched

starch storage carbohydrate in plants

Licensing & Attributions

CC licensed content, Shared previously

- carbohydrates. **Authored by**: open stax. **Provided by**: open stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:c7OsVIEc@6/Carbohydrates. **License**: *CC BY: Attribution*

LIPIDS

3.3 Lipids

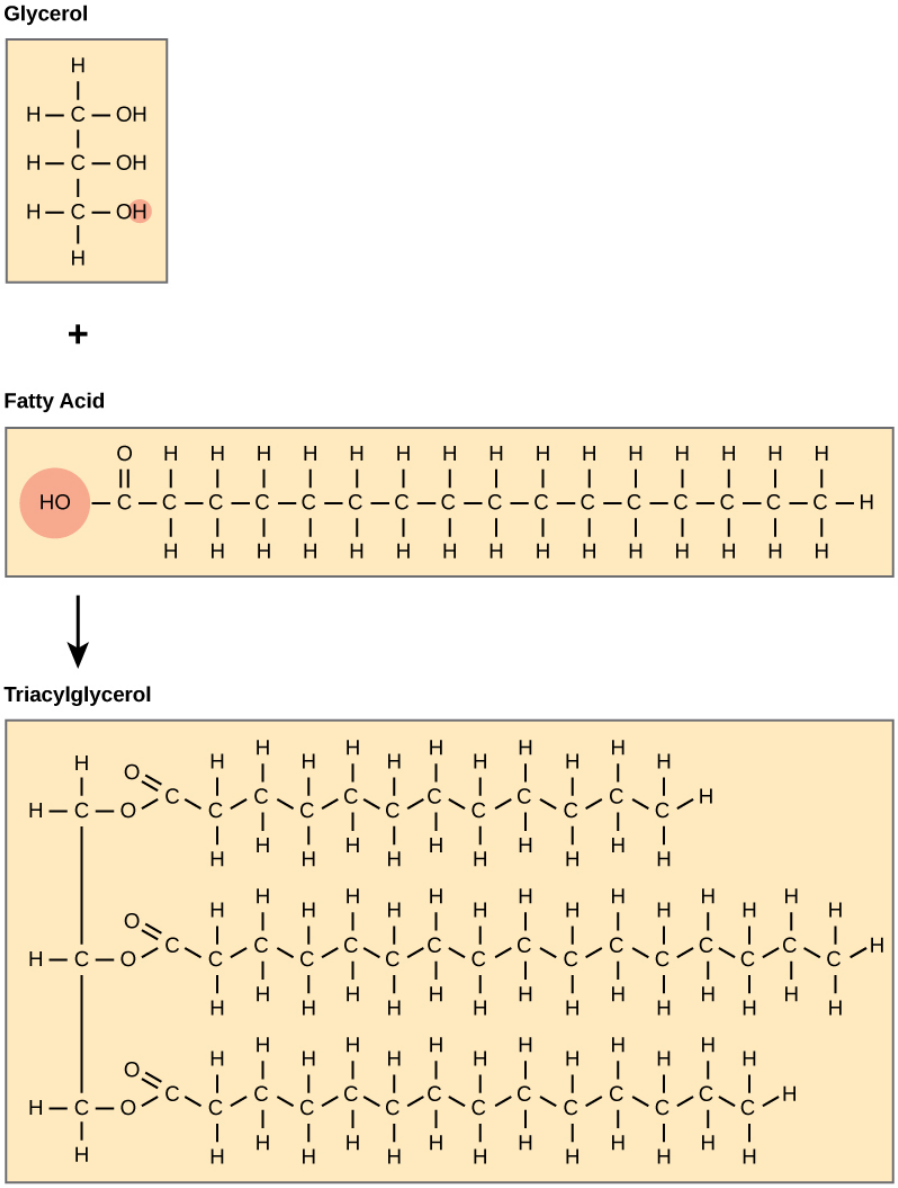
Summary

Lipids include a diverse group of compounds that are largely nonpolar in nature. This is because they are hydrocarbons that include mostly nonpolar carbon–carbon or carbon–hydrogen bonds. Non-polar molecules are hydrophobic (“water fearing”), or insoluble in water. Lipids perform many different functions in a cell. Cells store energy for long-term use in the form of fats. Lipids also provide insulation from the environment for plants and animals (Figure). For example, they help keep aquatic birds and mammals dry when forming a protective layer over fur or feathers because of their water-repellant hydrophobic nature. Lipids are also the building blocks of many hormones and are an important constituent of all cellular membranes. Lipids include fats, oils, waxes, phospholipids, and steroids.

Hydrophobic lipids in the fur of aquatic mammals, such as this river otter, protect them from the elements. (credit: Ken Bosma)

Fats and Oils

A fat molecule consists of two main components—glycerol and fatty acids. Glycerol is an organic compound (alcohol) with three carbons, five hydrogens, and three hydroxyl (OH) groups. Fatty acids have a long chain of hydrocarbons to which a carboxyl group is attached, hence the name “fatty acid.” The number of carbons in the fatty acid may range from 4 to 36; most common are those containing 12–18 carbons. In a fat molecule, the fatty acids are attached to each of the three carbons of the glycerol molecule with an ester bond through an oxygen atom (Figure).

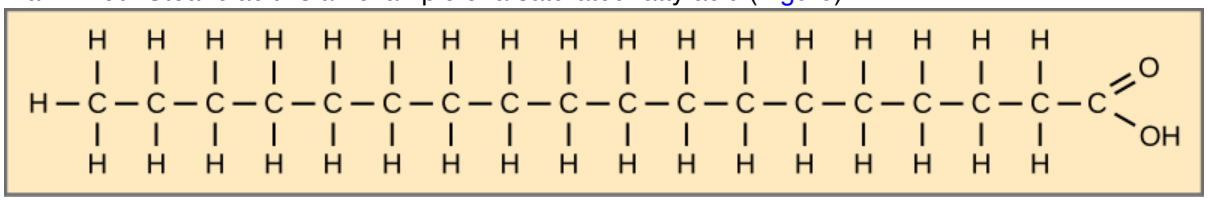


Triacylglycerol is formed by the joining of three fatty acids to a glycerol backbone in a dehydration reaction. Three molecules of water are released in the process.

During this ester bond formation, three water molecules are released. The three fatty acids in the triacylglycerol may be similar or dissimilar. Fats are also called triacylglycerols or triglycerides because of their chemical structure. Some fatty acids have common names that specify their origin. For example, palmitic acid, a saturated fatty acid, is derived from the palm tree. Arachidic acid is derived from *Arachis hypogea,* the scientific name for groundnuts or peanuts.

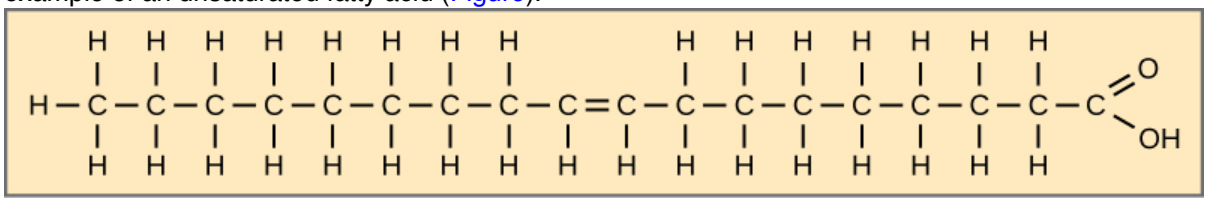
Fatty acids may be saturated or unsaturated. In a fatty acid chain, if there are only single bonds between neighboring carbons in the hydrocarbon chain, the fatty acid is said to be saturated. Saturated fatty acids are

saturated with hydrogen; in other words, the number of hydrogen atoms attached to the carbon skeleton is maximized. Stearic acid is an example of a saturated fatty acid (Figure)



Stearic acid is a common saturated fatty acid.

When the hydrocarbon chain contains a double bond, the fatty acid is said to be unsaturated. Oleic acid is an example of an unsaturated fatty acid (Figure).



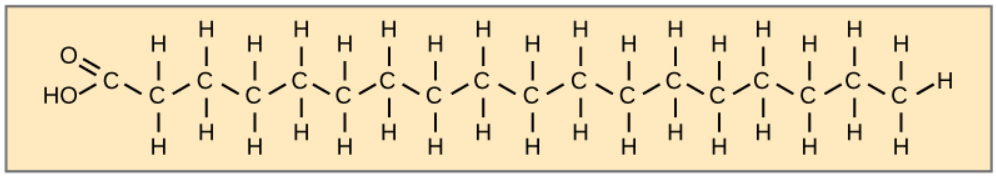
Oleic acid is a common unsaturated fatty acid.

Most unsaturated fats are liquid at room temperature and are called oils. If there is one double bond in the molecule, then it is known as a monounsaturated fat (e.g., olive oil), and if there is more than one double bond, then it is known as a polyunsaturated fat (e.g., canola oil).

When a fatty acid has no double bonds, it is known as a saturated fatty acid because no more hydrogen may be added to the carbon atoms of the chain. A fat may contain similar or different fatty acids attached to glycerol. Long straight fatty acids with single bonds tend to get packed tightly and are solid at room temperature. Animal fats with stearic acid and palmitic acid (common in meat) and the fat with butyric acid (common in butter) are examples of saturated fats. Mammals store fats in specialized cells called adipocytes, where globules of fat occupy most of the cell's volume. In plants, fat or oil is stored in many seeds and is used as a source of energy during seedling development. Unsaturated fats or oils are usually of plant origin and contain *cis* unsaturated fatty acids. *Cis* and *trans*indicate the configuration of the molecule around the double bond. If hydrogens are present in the same plane, it is referred to as a cis fat; if the hydrogen atoms are on two different planes, it is referred to as a trans fat. The *cis* double bond causes a bend or a "kink" that prevents the fatty acids from packing tightly, keeping them liquid at room temperature (Figure). Olive oil, corn oil, canola oil, and cod liver oil are examples of unsaturated fats. Unsaturated fats help to lower blood cholesterol levels whereas saturated fats contribute to plaque formation in the arteries.

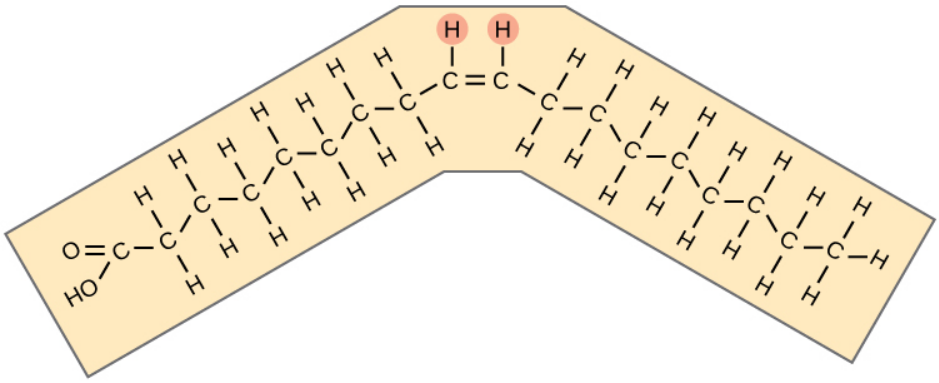
Saturated fatty acid

Stearic acid

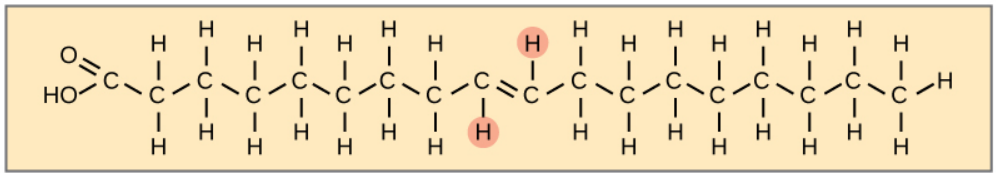


Unsaturated fatty acids

Cis oleic acid



Trans oleic acid



Saturated fatty acids have hydrocarbon chains connected by single bonds only. Unsaturated fatty acids have one or more double bonds. Each double bond may be in a *cis* or *trans* configuration. In the *cis* configuration, both hydrogens are on the same side of the hydrocarbon chain. In the *trans* configuration, the hydrogens are on opposite sides. A *cis* double bond causes a kink in the chain.

Trans Fats

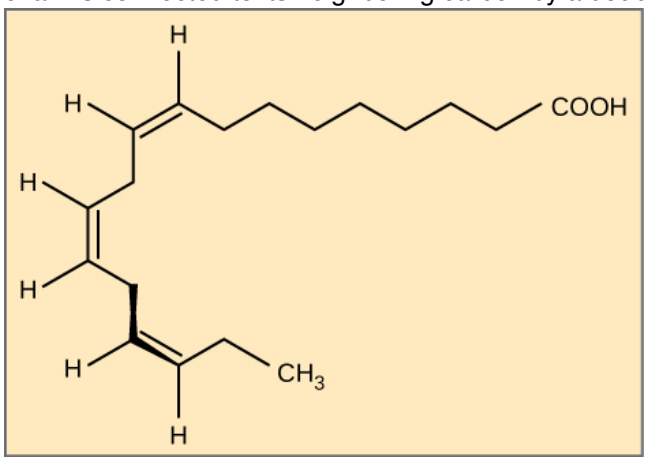
In the food industry, oils are artificially hydrogenated to make them semi-solid and of a consistency desirable for many processed food products. Simply speaking, hydrogen gas is bubbled through oils to solidify them. During this hydrogenation process, double bonds of the *cis*– conformation in the hydrocarbon chain may be converted to double bonds in the trans- conformation.

Margarine, some types of peanut butter, and shortening are examples of artificially hydrogenated trans fats. Recent studies have shown that an increase in trans fats in the human diet may lead to an increase in levels of low-density lipoproteins (LDL), or “bad” cholesterol, which in turn may lead to plaque deposition in the arteries, resulting in heart disease. Many fast food restaurants have recently banned the use of trans fats, and food labels are required to display the trans fat content.

Omega Fatty Acids

Essential fatty acids are fatty acids required but not synthesized by the human body. Consequently, they have to be supplemented through ingestion via the diet. Omega-3 fatty acids (like that shown in Figure) fall into this category and are one of only two known for humans (the other being omega-6 fatty acid). These are

polyunsaturated fatty acids and are called omega-3 because the third carbon from the end of the hydrocarbon chain is connected to its neighboring carbon by a double bond.



Alpha-linolenic acid is an example of an omega-3 fatty acid. It has three *cis* double bonds and, as a result, a curved shape. For clarity, the carbons are not shown. Each singly bonded carbon has two hydrogens associated with it, also not shown.

The farthest carbon away from the carboxyl group is numbered as the omega (ω) carbon, and if the double bond is between the third and fourth carbon from that end, it is known as an omega-3 fatty acid. Nutritionally important because the body does not make them, omega-3 fatty acids include alpha-linoleic acid (ALA), eicosapentaenoic acid (EPA), and docosahexaenoic acid (DHA), all of which are polyunsaturated. Salmon, trout, and tuna are good sources of omega-3 fatty acids. Research indicates that omega-3 fatty acids reduce the risk of sudden death from heart attacks, reduce triglycerides in the blood, lower blood pressure, and prevent thrombosis by inhibiting blood clotting. They also reduce inflammation, and may help reduce the risk of some cancers in animals.

Like carbohydrates, fats have received a lot of bad publicity. It is true that eating an excess of fried foods and other “fatty” foods leads to weight gain. However, fats do have important functions. Many vitamins are fat soluble, and fats serve as a long-term storage form of fatty acids: a source of energy. They also provide insulation for the body. Therefore, “healthy” fats in moderate amounts should be consumed on a regular basis.

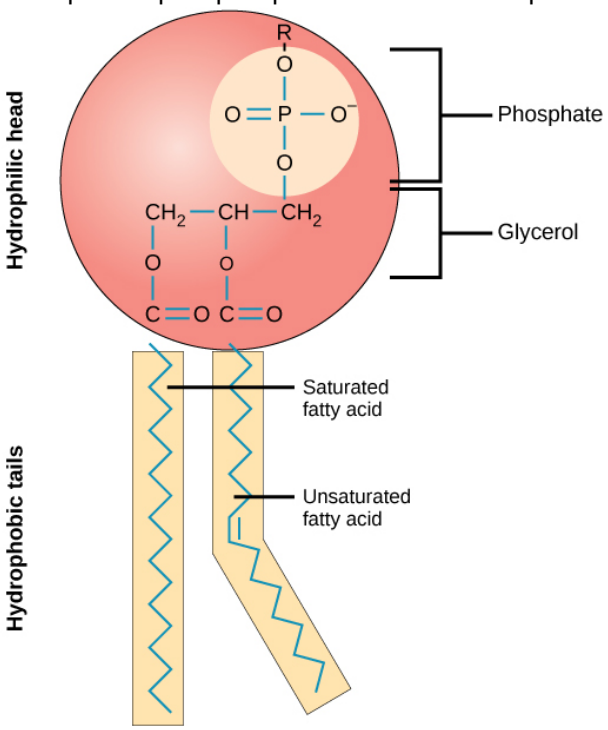
Waxes

Wax covers the feathers of some aquatic birds and the leaf surfaces of some plants. Because of the hydrophobic nature of waxes, they prevent water from sticking on the surface (Figure). Waxes are made up of long fatty acid chains esterified to long-chain alcohols.

Waxy coverings on some leaves are made of lipids. (credit: Roger Griffith)

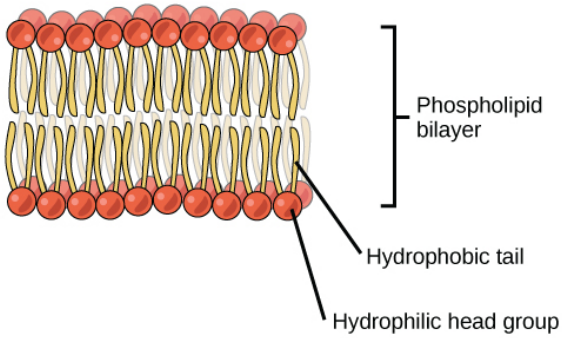
Phospholipids

Phospholipids are major constituents of the plasma membrane, the outermost layer of animal cells. Like fats, they are composed of fatty acid chains attached to a glycerol or sphingosine backbone. Instead of three fatty acids attached as in triglycerides, however, there are two fatty acids forming diacylglycerol, and the third carbon of the glycerol backbone is occupied by a modified phosphate group (Figure). A phosphate group alone attached to a diaglycerol does not qualify as a phospholipid; it is phosphatidate (diacylglycerol 3-phosphate), the precursor of phospholipids. The phosphate group is modified by an alcohol. Phosphatidylcholine and phosphatidylserine are two important phospholipids that are found in plasma membranes.



A phospholipid is a molecule with two fatty acids and a modified phosphate group attached to a glycerol backbone. The phosphate may be modified by the addition of charged or polar chemical groups. Two chemical groups that may modify the phosphate, choline and serine, are shown here. Both choline and serine attach to the phosphate group at the position labeled R via the hydroxyl group indicated in green.

A phospholipid is an amphipathic molecule, meaning it has a hydrophobic and a hydrophilic part. The fatty acid chains are hydrophobic and cannot interact with water, whereas the phosphate-containing group is hydrophilic and interacts with water (Figure).



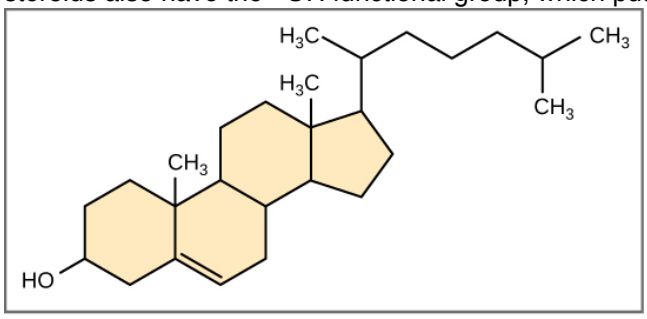
The phospholipid bilayer is the major component of all cellular membranes. The hydrophilic head groups of the phospholipids face the aqueous solution. The hydrophobic tails are sequestered in the middle of the bilayer.

The head is the hydrophilic part, and the tail contains the hydrophobic fatty acids. In a membrane, a bilayer of phospholipids forms the matrix of the structure, the fatty acid tails of phospholipids face inside, away from water, whereas the phosphate group faces the outside, aqueous side (Figure).

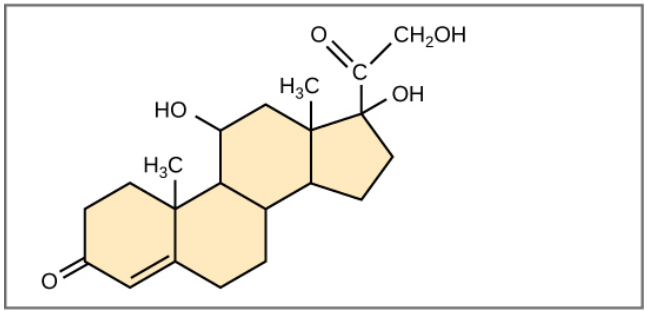
Phospholipids are responsible for the dynamic nature of the plasma membrane. If a drop of phospholipids is placed in water, it spontaneously forms a structure known as a micelle, where the hydrophilic phosphate heads face the outside and the fatty acids face the interior of this structure.

Steroids

Unlike the phospholipids and fats discussed earlier, steroids have a fused ring structure. Although they do not resemble the other lipids, they are grouped with them because they are also hydrophobicand insoluble in water. All steroids have four linked carbon rings and several of them, like cholesterol, have a short tail (Figure). Many steroids also have the –OH functional group, which puts them in the alcohol classification (sterols).



Cholesterol



Cortisol

Steroids such as cholesterol and cortisol are composed of four fused hydrocarbon rings.

Cholesterol is the most common steroid. Cholesterol is mainly synthesized in the liver and is the precursor to many steroid hormones such as testosterone and estradiol, which are secreted by the gonads and endocrine glands. It is also the precursor to Vitamin D. Cholesterol is also the precursor of bile salts, which help in the emulsification of fats and their subsequent absorption by cells. Although cholesterol is often spoken of in negative terms by lay people, it is necessary for proper functioning of the body. It is a component of the plasma membrane of animal cells and is found within the phospholipid bilayer. Being the outermost structure in animal cells, the plasma membrane is responsible for the transport of materials and cellular recognition and it is involved in cell-to-cell communication.

Link to Learning

For an additional perspective on lipids, explore the interactive animation “Biomolecules: The Lipids”

.

Section Summary

Lipids are a class of macromolecules that are nonpolar and hydrophobic in nature. Major types include fats and oils, waxes, phospholipids, and steroids. Fats are a stored form of energy and are also known as triacylglycerols

or triglycerides. Fats are made up of fatty acids and either glycerol or sphingosine. Fatty acids may be unsaturated or saturated, depending on the presence or absence of double bonds in the hydrocarbon chain. If only single bonds are present, they are known as saturated fatty acids. Unsaturated fatty acids may have one or more double bonds in the hydrocarbon chain. Phospholipids make up the matrix of membranes. They have a glycerol or sphingosine backbone to which two fatty acid chains and a phosphate-containing group are attached. Steroids are another class of lipids. Their basic structure has four fused carbon rings. Cholesterol is a type of steroid and is an important constituent of the plasma membrane, where it helps to maintain the fluid nature of the membrane. It is also the precursor of steroid hormones such as testosterone.

Review Questions

Saturated fats have all of the following characteristics except:

1. they are solid at room temperature
2. they have single bonds within the carbon chain
3. they are usually obtained from animal sources
4. they tend to dissolve in water easily

Phospholipids are important components of _____.

1. the plasma membrane of animal cells
2. the ring structure of steroids
3. the waxy covering on leaves
4. the double bond in hydrocarbon chains

Free Response

Explain at least three functions that lipids serve in plants and/or animals.

Why have trans fats been banned from some restaurants? How are they created?

Glossary

lipid macromolecule that is nonpolar and insoluble in water

omega fat type of polyunsaturated fat that is required by the body; the numbering of the carbon omega starts from the methyl end or the end that is farthest from the carboxylic end

phospholipid major constituent of the membranes; composed of two fatty acids and a phosphate-containing group attached to a glycerol backbone

saturated fatty acid long-chain of hydrocarbon with single covalent bonds in the carbon chain; the number of hydrogen atoms attached to the carbon skeleton is maximized

steroid type of lipid composed of four fused hydrocarbon rings forming a planar structure

trans fat fat formed artificially by hydrogenating oils, leading to a different arrangement of double bond(s) than those found in naturally occurring lipids

triacylglycerol (also, triglyceride) fat molecule; consists of three fatty acids linked to a glycerol molecule

unsaturated fatty acid long-chain hydrocarbon that has one or more double bonds in the hydrocarbon chain

wax lipid made of a long-chain fatty acid that is esterified to a long-chain alcohol; serves as a protective coating on some feathers, aquatic mammal fur, and leaves

Licensing & Attributions

CC licensed content, Shared previously

- lipids. **Authored by**: open stax. **Provided by**: open stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:l_8zQ2kk@8/Lipids. **License**: *CC BY: Attribution*

PROTEINS

3.4 Proteins

Summary

Proteins are one of the most abundant organic molecules in living systems and have the most diverse range of functions of all macromolecules. Proteins may be structural, regulatory, contractile, or protective; they may serve in transport, storage, or membranes; or they may be toxins or enzymes. Each cell in a living system may contain thousands of proteins, each with a unique function. Their structures, like their functions, vary greatly. They are all, however, polymers of amino acids, arranged in a linear sequence.

Types and Functions of Proteins

Enzymes, which are produced by living cells, are catalysts in biochemical reactions (like digestion) and are usually complex or conjugated proteins. Each enzyme is specific for the substrate (a reactant that binds to an enzyme) it acts on. The enzyme may help in breakdown, rearrangement, or synthesis reactions. Enzymes that break down their substrates are called catabolic enzymes, enzymes that build more complex molecules from their substrates are called anabolic enzymes, and enzymes that affect the rate of reaction are called catalytic enzymes. It should be noted that all enzymes increase the rate of reaction and, therefore, are considered to be organic catalysts. An example of an enzyme is salivary amylase, which hydrolyzes its substrate amylose, a component of starch.

Hormones are chemical-signaling molecules, usually small proteins or steroids, secreted by endocrine cells that act to control or regulate specific physiological processes, including growth, development, metabolism, and reproduction. For example, insulin is a protein hormone that helps to regulate the blood glucose level. The primary types and functions of proteins are listed in Table.

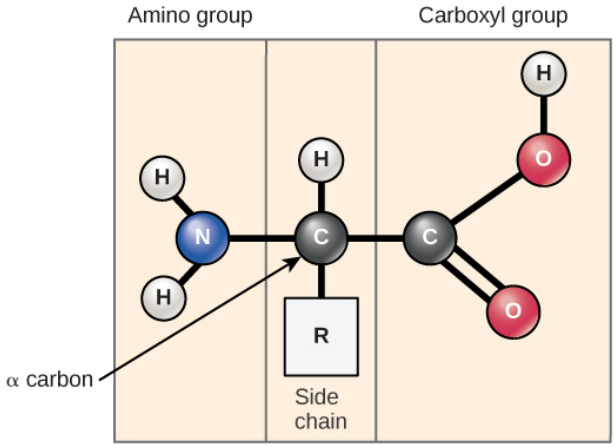
Protein Types and Functions		
Type	Examples	Functions
Digestive Enzymes	Amylase, lipase, pepsin, trypsin	Help in digestion of food by catabolizing nutrients into monomeric units
Transport	Hemoglobin, albumin	Carry substances in the blood or lymph throughout the body
Structural	Actin, tubulin, keratin	Construct different structures, like the cytoskeleton
Hormones	Insulin, thyroxine	Coordinate the activity of different body systems
Defense	Immunoglobulins	Protect the body from foreign pathogens
Contractile	Actin, myosin	Effect muscle contraction

Protein Types and Functions		
Type	Examples	Functions
Storage	Legume storage proteins, egg white (albumin)	Provide nourishment in early development of the embryo and the seedling

Proteins have different shapes and molecular weights; some proteins are globular in shape whereas others are fibrous in nature. For example, hemoglobin is a globular protein, but collagen, found in our skin, is a fibrous protein. Protein shape is critical to its function, and this shape is maintained by many different types of chemical bonds. Changes in temperature, pH, and exposure to chemicals may lead to permanent changes in the shape of the protein, leading to loss of function, known as denaturation. All proteins are made up of different arrangements of the same 20 types of amino acids.

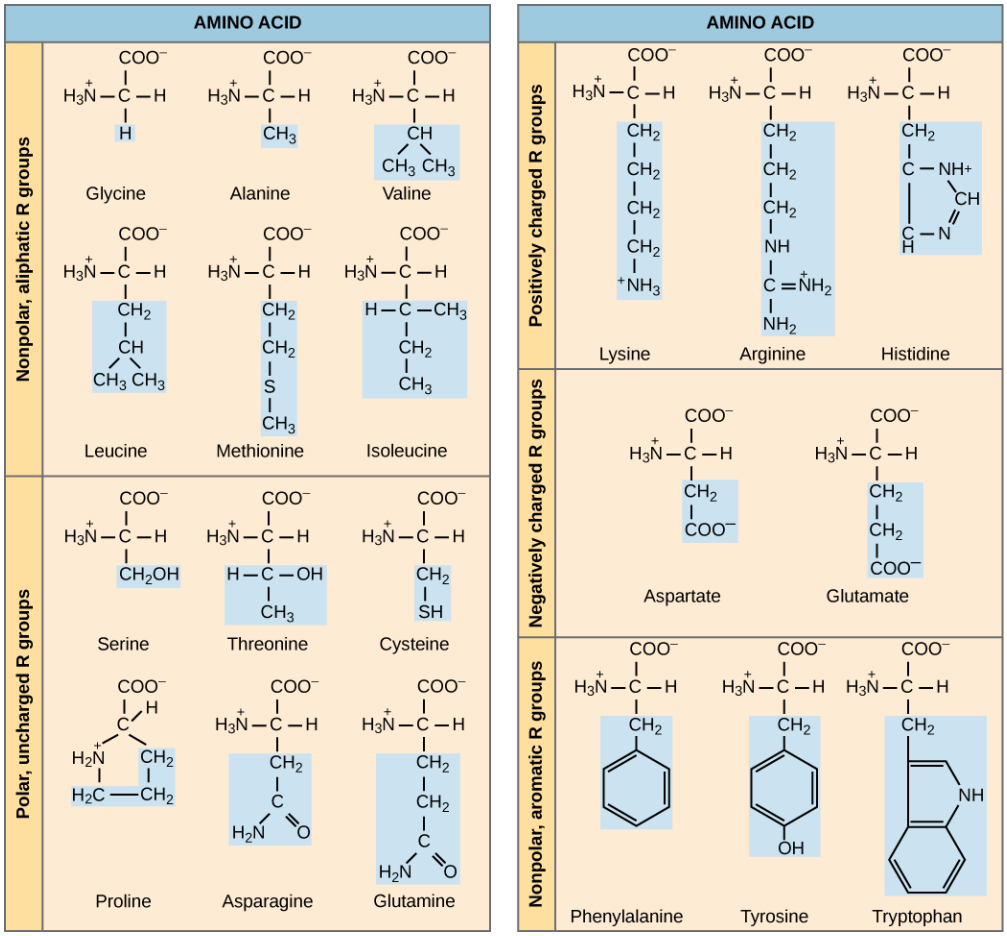
Amino Acids

Amino acids are the monomers that make up proteins. Each amino acid has the same fundamental structure, which consists of a central carbon atom, also known as the alpha (α) carbon, bonded to an amino group (NH_2), a carboxyl group (COOH), and to a hydrogen atom. Every amino acid also has another atom or group of atoms bonded to the central atom known as the R group (Figure).



Amino acids have a central asymmetric carbon to which an amino group, a carboxyl group, a hydrogen atom, and a side chain (R group) are attached.

The name “amino acid” is derived from the fact that they contain both amino group and carboxyl-acid-group in their basic structure. As mentioned, there are 20 amino acids present in proteins. Ten of these are considered essential amino acids in humans because the human body cannot produce them and they are obtained from the diet. For each amino acid, the R group (or side chain) is different (Figure).
ART CONNECTION



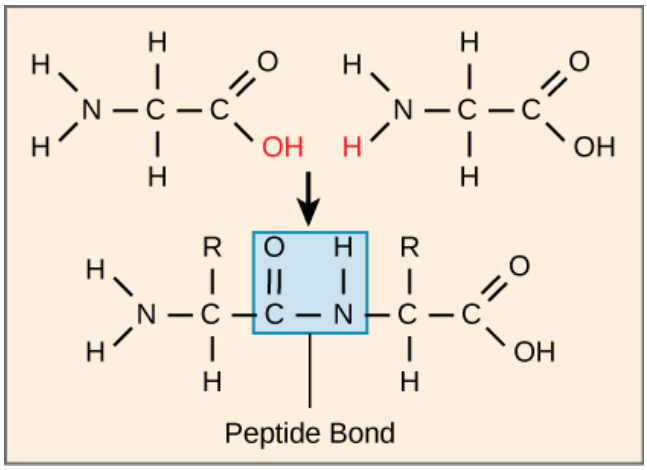
There are 20 common amino acids commonly found in proteins, each with a different R group (variant group) that determines its chemical nature.

Which categories of amino acid would you expect to find on the surface of a soluble protein, and which would you expect to find in the interior? What distribution of amino acids would you expect to find in a protein embedded in a lipid bilayer?

The chemical nature of the side chain determines the nature of the amino acid (that is, whether it is acidic, basic, polar, or nonpolar). For example, the amino acid glycine has a hydrogen atom as the R group. Amino acids such as valine, methionine, and alanine are nonpolar or hydrophobic in nature, while amino acids such as serine, threonine, and cysteine are polar and have hydrophilic side chains. The side chains of lysine and arginine are positively charged, and therefore these amino acids are also known as basic amino acids. Proline has an R group that is linked to the amino group, forming a ring-like structure. Proline is an exception to the standard structure of an animo acid since its amino group is not separate from the side chain (Figure).

Amino acids are represented by a single upper case letter or a three-letter abbreviation. For example, valine is known by the letter V or the three-letter symbol val. Just as some fatty acids are essential to a diet, some amino acids are necessary as well. They are known as essential amino acids, and in humans they include isoleucine, leucine, and cysteine. Essential amino acids refer to those necessary for construction of proteins in the body, although not produced by the body; which amino acids are essential varies from organism to organism.

The sequence and the number of amino acids ultimately determine the protein's shape, size, and function. Each amino acid is attached to another amino acid by a covalent bond, known as a peptide bond, which is formed by a dehydration reaction. The carboxyl group of one amino acid and the amino group of the incoming amino acid combine, releasing a molecule of water. The resulting bond is the peptide bond (Figure).



Peptide bond formation is a dehydration synthesis reaction. The carboxyl group of one amino acid is linked to the amino group of the incoming amino acid. In the process, a molecule of water is released.

The products formed by such linkages are called peptides. As more amino acids join to this growing chain, the resulting chain is known as a polypeptide. Each polypeptide has a free amino group at one end. This end is called the N terminal, or the amino terminal, and the other end has a free carboxyl group, also known as the C or carboxyl terminal. While the terms polypeptide and protein are sometimes used interchangeably, a polypeptide is technically a polymer of amino acids, whereas the term protein is used for a polypeptide or polypeptides that have combined together, often have bound non-peptide prosthetic groups, have a distinct shape, and have a unique function. After protein synthesis (translation), most proteins are modified. These are known as post-translational modifications. They may undergo cleavage, phosphorylation, or may require the addition of other chemical groups. Only after these modifications is the protein completely functional.

Link to Learning

Click through the steps of protein synthesis in this interactive tutorial.
EVOLUTION CONNECTION

The Evolutionary Significance of Cytochrome cCytochrome c is an important component of the electron transport chain, a part of cellular respiration, and it is normally found in the cellular organelle, the mitochondrion. This protein has a heme prosthetic group, and the central ion of the heme gets alternately reduced and oxidized during electron transfer. Because this essential protein's role in producing cellular energy is crucial, it has changed very little over millions of years. Protein sequencing has shown that there is a considerable amount of cytochrome c amino acid sequence homology among different species; in other words, evolutionary kinship can be assessed by measuring the similarities or differences among various species' DNA or protein sequences.

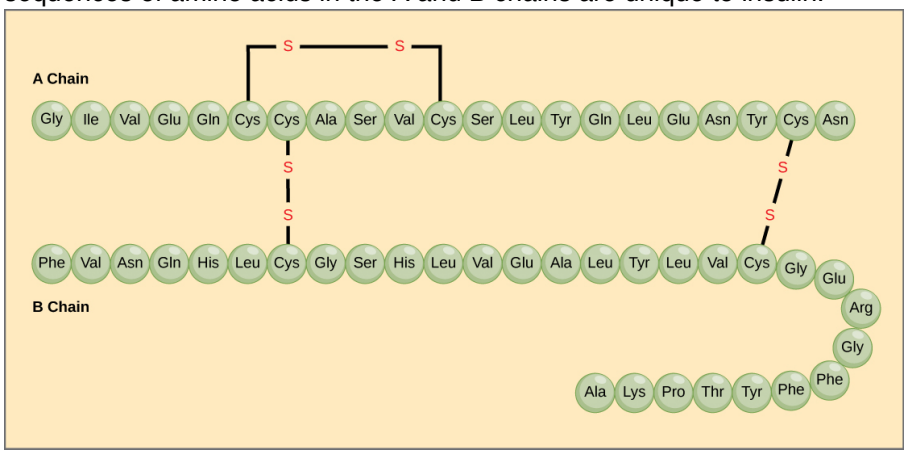
Scientists have determined that human cytochrome c contains 104 amino acids. For each cytochrome c molecule from different organisms that has been sequenced to date, 37 of these amino acids appear in the same position in all samples of cytochrome c. This indicates that there may have been a common ancestor. On comparing the human and chimpanzee protein sequences, no sequence difference was found. When human and rhesus monkey sequences were compared, the single difference found was in one amino acid. In another comparison, human to yeast sequencing shows a difference in the 44th position.

Protein Structure

As discussed earlier, the shape of a protein is critical to its function. For example, an enzyme can bind to a specific substrate at a site known as the active site. If this active site is altered because of local changes or changes in overall protein structure, the enzyme may be unable to bind to the substrate. To understand how the protein gets its final shape or conformation, we need to understand the four levels of protein structure: primary, secondary, tertiary, and quaternary.

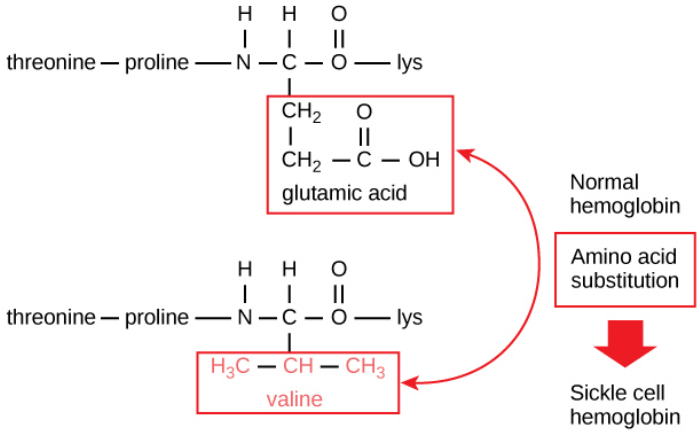
Primary Structure

The unique sequence of amino acids in a polypeptide chain is its primary structure. For example, the pancreatic hormone insulin has two polypeptide chains, A and B, and they are linked together by disulfide bonds. The N terminal amino acid of the A chain is glycine, whereas the C terminal amino acid is asparagine (Figure). The sequences of amino acids in the A and B chains are unique to insulin.



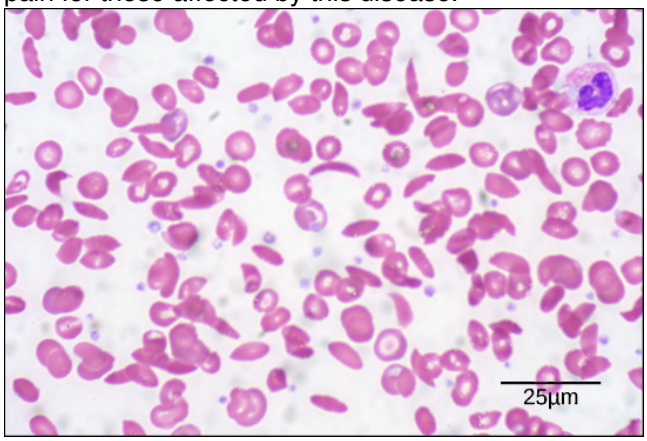
Bovine serum insulin is a protein hormone made of two peptide chains, A (21 amino acids long) and B (30 amino acids long). In each chain, primary structure is indicated by three-letter abbreviations that represent the names of the amino acids in the order they are present. The amino acid cysteine (cys) has a sulfhydryl (SH) group as a side chain. Two sulfhydryl groups can react in the presence of oxygen to form a disulfide (S-S) bond. Two disulfide bonds connect the A and B chains together, and a third helps the A chain fold into the correct shape. Note that all disulfide bonds are the same length, but are drawn different sizes for clarity.

The unique sequence for every protein is ultimately determined by the gene encoding the protein. A change in nucleotide sequence of the gene's coding region may lead to a different amino acid being added to the growing polypeptide chain, causing a change in protein structure and function. In sickle cell anemia, the hemoglobin β chain (a small portion of which is shown in Figure) has a single amino acid substitution, causing a change in protein structure and function. Specifically, the amino acid glutamic acid is substituted by valine in the β chain. What is most remarkable to consider is that a hemoglobin molecule is made up of two alpha chains and two beta chains that each consist of about 150 amino acids. The molecule, therefore, has about 600 amino acids. The structural difference between a normal hemoglobin molecule and a sickle cell molecule—which dramatically decreases life expectancy—is a single amino acid of the 600. What is even more remarkable is that those 600 amino acids are encoded by three nucleotides each, and the mutation is caused by a single base change (point mutation), 1 in 1800 bases.



The beta chain of hemoglobin is 147 residues in length, yet a single amino acid substitution leads to sickle cell anemia. In normal hemoglobin, the amino acid at position seven is glutamate. In sickle cell hemoglobin, this glutamate is replaced by a valine.

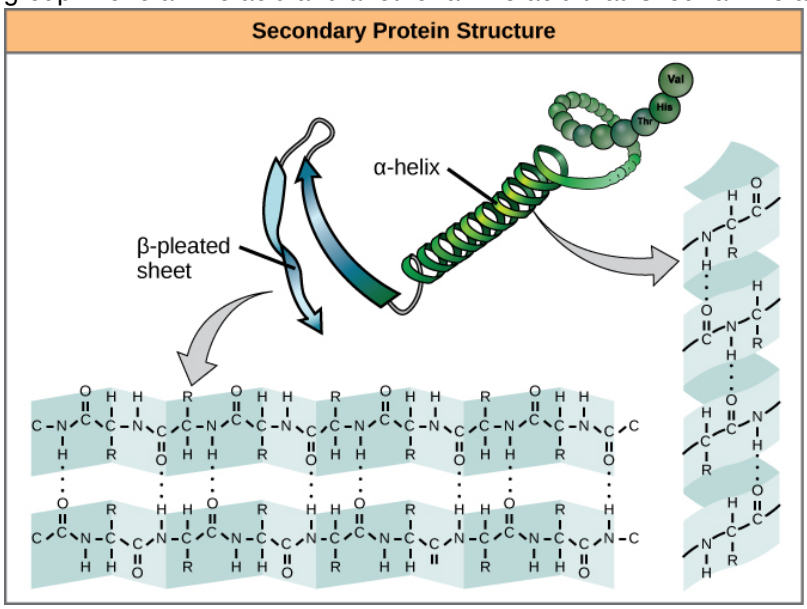
Because of this change of one amino acid in the chain, hemoglobin molecules form long fibers that distort the biconcave, or disc-shaped, red blood cells and assume a crescent or “sickle” shape, which clogs arteries (Figure). This can lead to myriad serious health problems such as breathlessness, dizziness, headaches, and abdominal pain for those affected by this disease.



In this blood smear, visualized at 535x magnification using bright field microscopy, sickle cells are crescent shaped, while normal cells are disc-shaped. (credit: modification of work by Ed Uthman; scale-bar data from Matt Russell)

Secondary Structure

The local folding of the polypeptide in some regions gives rise to the secondary structure of the protein. The most common are the*α*-helix and *β*-pleated sheet structures (Figure). Both structures are the *α*-helix structure—the helix held in shape by hydrogen bonds. The hydrogen bonds form between the oxygen atom in the carbonyl group in one amino acid and another amino acid that is four amino acids farther along the chain.



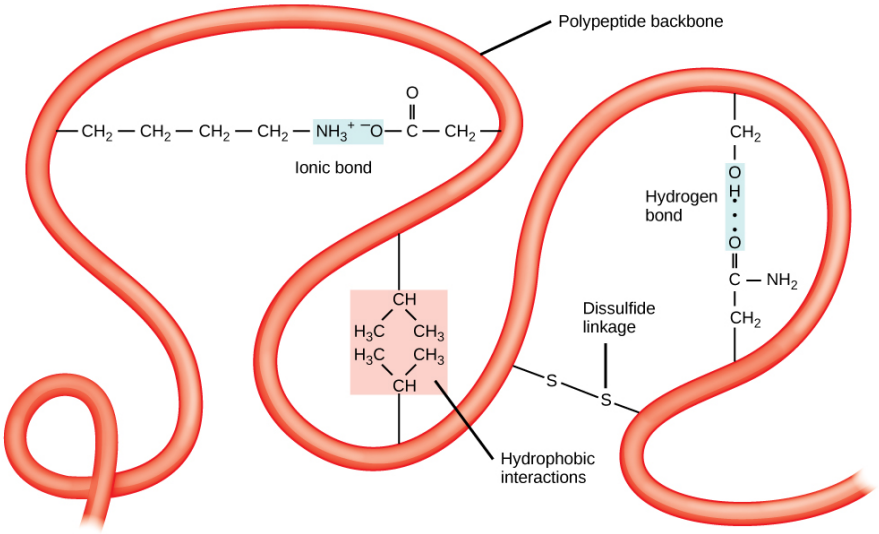
The *α*-helix and *β*-pleated sheet are secondary structures of proteins that form because of hydrogen bonding between carbonyl and amino groups in the peptide backbone. Certain amino acids have a propensity to form an *α*-helix, while others have a propensity to form a *β*-pleated sheet.

Every helical turn in an alpha helix has 3.6 amino acid residues. The R groups (the variant groups) of the polypeptide protrude out from the *α*-helix chain. In the *β*-pleated sheet, the “pleats” are formed by hydrogen bonding between atoms on the backbone of the polypeptide chain. The R groups are attached to the carbons and extend above and below the folds of the pleat. The pleated segments align parallel or antiparallel to each other, and hydrogen bonds form between the partially positive nitrogen atom in the amino group and the partially

negative oxygen atom in the carbonyl group of the peptide backbone. The α-helix and β-pleated sheet structures are found in most globular and fibrous proteins and they play an important structural role.

Tertiary Structure

The unique three-dimensional structure of a polypeptide is its tertiary structure (Figure). This structure is in part due to chemical interactions at work on the polypeptide chain. Primarily, the interactions among R groups creates the complex three-dimensional tertiary structure of a protein. The nature of the R groups found in the amino acids involved can counteract the formation of the hydrogen bonds described for standard secondary structures. For example, R groups with like charges are repelled by each other and those with unlike charges are attracted to each other (ionic bonds). When protein folding takes place, the hydrophobic R groups of nonpolar amino acids lay in the interior of the protein, whereas the hydrophilic R groups lay on the outside. The former types of interactions are also known as hydrophobic interactions. Interaction between cysteine side chains forms disulfide linkages in the presence of oxygen, the only covalent bond forming during protein folding.



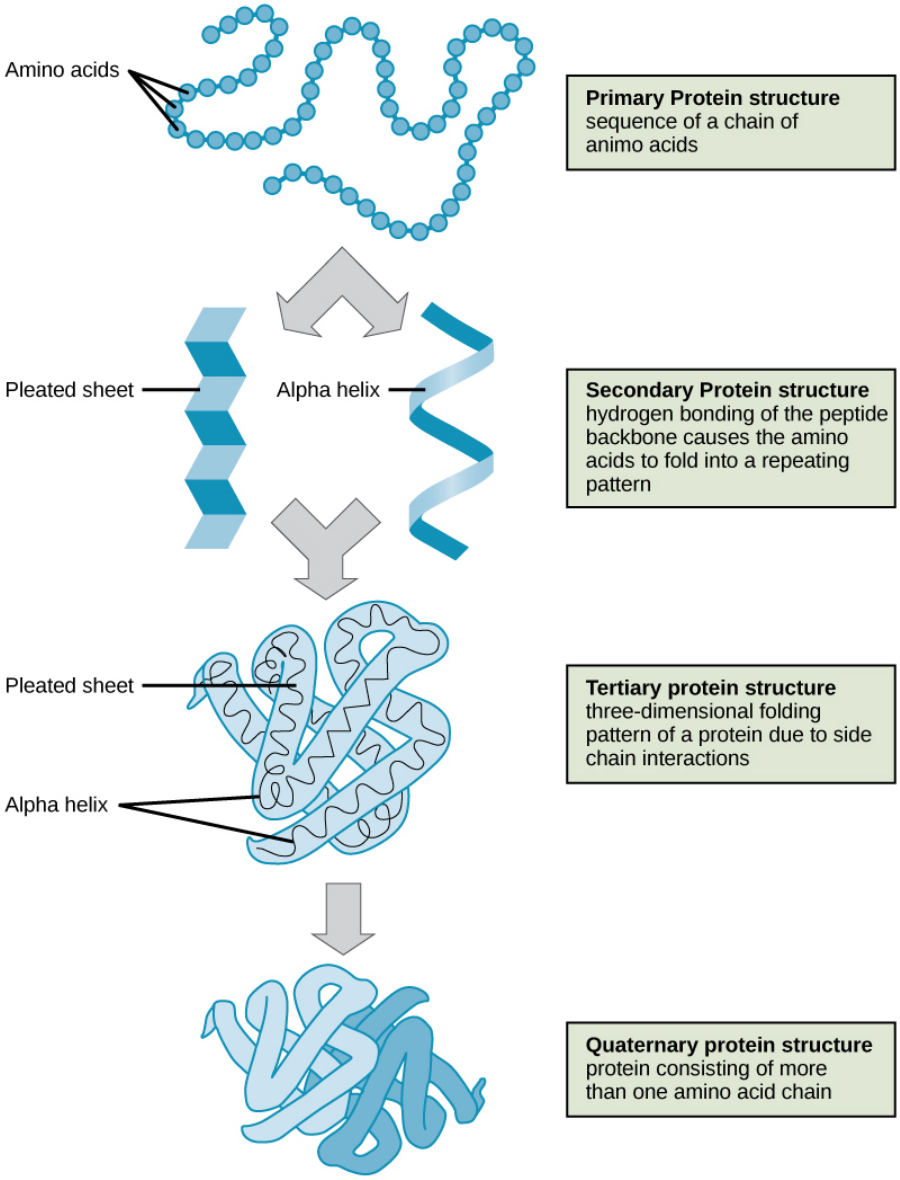
The tertiary structure of proteins is determined by a variety of chemical interactions. These include hydrophobic interactions, ionic bonding, hydrogen bonding and disulfide linkages.

All of these interactions, weak and strong, determine the final three-dimensional shape of the protein. When a protein loses its three-dimensional shape, it may no longer be functional.

Quaternary Structure

In nature, some proteins are formed from several polypeptides, also known as subunits, and the interaction of these subunits forms the quaternary structure. Weak interactions between the subunits help to stabilize the overall structure. For example, insulin (a globular protein) has a combination of hydrogen bonds and disulfide bonds that cause it to be mostly clumped into a ball shape. Insulin starts out as a single polypeptide and loses some internal sequences in the presence of post-translational modification after the formation of the disulfide linkages that hold the remaining chains together. Silk (a fibrous protein), however, has a β-pleated sheet structure that is the result of hydrogen bonding between different chains.

The four levels of protein structure (primary, secondary, tertiary, and quaternary) are illustrated in Figure.



The four levels of protein structure can be observed in these illustrations. (credit: modification of work by National Human Genome Research Institute)

Denaturation and Protein Folding

Each protein has its own unique sequence and shape that are held together by chemical interactions. If the protein is subject to changes in temperature, pH, or exposure to chemicals, the protein structure may change, losing its shape without losing its primary sequence in what is known as denaturation. Denaturation is often reversible because the primary structure of the polypeptide is conserved in the process if the denaturing agent is removed, allowing the protein to resume its function. Sometimes denaturation is irreversible, leading to loss of function. One example of irreversible protein denaturation is when an egg is fried. The albumin protein in the liquid egg white is denatured when placed in a hot pan. Not all proteins are denatured at high temperatures; for instance, bacteria that survive in hot springs have proteins that function at temperatures close to boiling. The stomach is also very acidic, has a low pH, and denatures proteins as part of the digestion process; however, the digestive enzymes of the stomach retain their activity under these conditions.

Protein folding is critical to its function. It was originally thought that the proteins themselves were responsible for the folding process. Only recently was it found that often they receive assistance in the folding process from protein helpers known as chaperones (or chaperonins) that associate with the target protein during the folding

process. They act by preventing aggregation of polypeptides that make up the complete protein structure, and they disassociate from the protein once the target protein is folded.

Link to Learning

For an additional perspective on proteins, view this animation called “Biomolecules: The Proteins.”

Section Summary

Proteins are a class of macromolecules that perform a diverse range of functions for the cell. They help in metabolism by providing structural support and by acting as enzymes, carriers, or hormones. The building blocks of proteins (monomers) are amino acids. Each amino acid has a central carbon that is linked to an amino group, a carboxyl group, a hydrogen atom, and an R group or side chain. There are 20 commonly occurring amino acids, each of which differs in the R group. Each amino acid is linked to its neighbors by a peptide bond. A long chain of amino acids is known as a polypeptide.

Proteins are organized at four levels: primary, secondary, tertiary, and (optional) quaternary. The primary structure is the unique sequence of amino acids. The local folding of the polypeptide to form structures such as the α helix and β-pleated sheet constitutes the secondary structure. The overall three-dimensional structure is the tertiary structure. When two or more polypeptides combine to form the complete protein structure, the configuration is known as the quaternary structure of a protein. Protein shape and function are intricately linked; any change in shape caused by changes in temperature or pH may lead to protein denaturation and a loss in function.

Art Connections

Figure Which categories of amino acid would you expect to find on the surface of a soluble protein, and which would you expect to find in the interior? What distribution of amino acids would you expect to find in a protein embedded in a lipid bilayer?

Review Questions

The monomers that make up proteins are called _____.

1. nucleotides
2. disaccharides
3. amino acids
4. chaperones

The α helix and the β-pleated sheet are part of which protein structure?

1. primary
2. secondary
3. tertiary
4. quaternary

Free Response

Explain what happens if even one amino acid is substituted for another in a polypeptide chain. Provide a specific example.

Describe the differences in the four protein structures.

Glossary

alpha-helix structure (*α*-helix) type of secondary structure of proteins formed by folding of the polypeptide into a helix shape with hydrogen bonds stabilizing the structure

amino acid monomer of a protein; has a central carbon or alpha carbon to which an amino group, a carboxyl group, a hydrogen, and an R group or side chain is attached; the R group is different for all 20 amino acids

beta-pleated sheet (*β*-pleated) secondary structure found in proteins in which “pleats” are formed by hydrogen bonding between atoms on the backbone of the polypeptide chain

chaperone (also, chaperonin) protein that helps nascent protein in the folding process

denaturation loss of shape in a protein as a result of changes in temperature, pH, or exposure to chemicals

enzyme catalyst in a biochemical reaction that is usually a complex or conjugated protein

hormone chemical signaling molecule, usually protein or steroid, secreted by endocrine cells that act to control or regulate specific physiological processes

peptide bond bond formed between two amino acids by a dehydration reaction

polypeptide long chain of amino acids linked by peptide bonds

primary structure linear sequence of amino acids in a protein

protein biological macromolecule composed of one or more chains of amino acids

quaternary structure association of discrete polypeptide subunits in a protein

secondary structure regular structure formed by proteins by intramolecular hydrogen bonding between the oxygen atom of one amino acid residue and the hydrogen attached to the nitrogen atom of another amino acid residue

tertiary structure three-dimensional conformation of a protein, including interactions between secondary structural elements; formed from interactions between amino acid side chains

Licensing & Attributions

CC licensed content, Shared previously

- proteins. **Authored by**: open stax. **Provided by**: o-pen stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:2zzm1QG9@7/Proteins. **License**: *CC BY: Attribution*

NUCLEIC ACIDS

3.5 Nucleic Acids

Summary

Nucleic acids are the most important macromolecules for the continuity of life. They carry the genetic blueprint of a cell and carry instructions for the functioning of the cell.

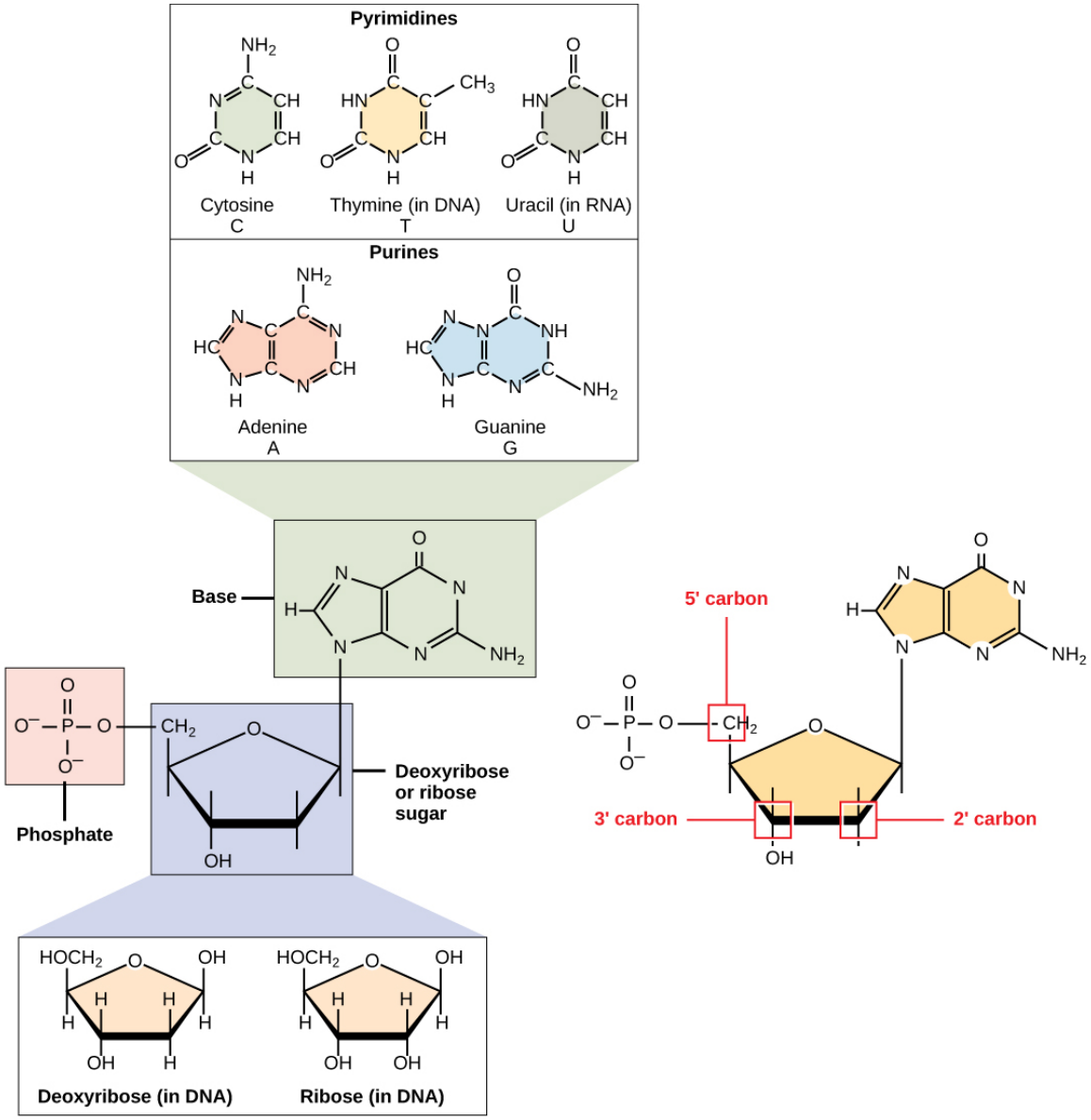
DNA and RNA

The two main types of nucleic acids are deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). DNA is the genetic material found in all living organisms, ranging from single-celled bacteria to multicellular mammals. It is found in the nucleus of eukaryotes and in the organelles, chloroplasts, and mitochondria. In prokaryotes, the DNA is not enclosed in a membranous envelope.

The entire genetic content of a cell is known as its genome, and the study of genomes is genomics. In eukaryotic cells but not in prokaryotes, DNA forms a complex with histone proteins to form chromatin, the substance of eukaryotic chromosomes. A chromosome may contain tens of thousands of genes. Many genes contain the information to make protein products; other genes code for RNA products. DNA controls all of the cellular activities by turning the genes “on” or “off.”

The other type of nucleic acid, RNA, is mostly involved in protein synthesis. The DNA molecules never leave the nucleus but instead use an intermediary to communicate with the rest of the cell. This intermediary is the messenger RNA (mRNA). Other types of RNA—like rRNA, tRNA, and microRNA—are involved in protein synthesis and its regulation.

DNA and RNA are made up of monomers known as nucleotides. The nucleotides combine with each other to form apolynucleotide, DNA or RNA. Each nucleotide is made up of three components: a nitrogenous base, a pentose (five-carbon) sugar, and a phosphate group (Figure). Each nitrogenous base in a nucleotide is attached to a sugar molecule, which is attached to one or more phosphate groups.



A nucleotide is made up of three components: a nitrogenous base, a pentose sugar, and one or more phosphate groups. Carbon residues in the pentose are numbered 1′ through 5′ (the prime distinguishes these residues from those in the base, which are numbered without using a prime notation). The base is attached to the 1′ position of the ribose, and the phosphate is attached to the 5′ position. When a polynucleotide is formed, the 5′ phosphate of the incoming nucleotide attaches to the 3′ hydroxyl group at the end of the growing chain. Two types of pentose are found in nucleotides, deoxyribose (found in DNA) and ribose (found in RNA). Deoxyribose is similar in structure to ribose, but it has an H instead of an OH at the 2′ position. Bases can be divided into two categories: purines and pyrimidines. Purines have a double ring structure, and pyrimidines have a single ring.

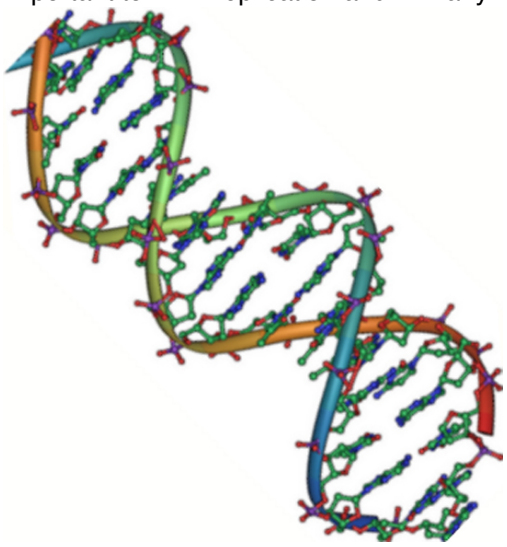
The nitrogenous bases, important components of nucleotides, are organic molecules and are so named because they contain carbon and nitrogen. They are bases because they contain an amino group that has the potential of binding an extra hydrogen, and thus, decreases the hydrogen ion concentration in its environment, making it more basic. Each nucleotide in DNA contains one of four possible nitrogenous bases: adenine (A), guanine (G) cytosine (C), and thymine (T).

Adenine and guanine are classified as purines. The primary structure of a purine is two carbon-nitrogen rings. Cytosine, thymine, and uracil are classified as pyrimidines which have a single carbon-nitrogen ring as their primary structure (Figure). Each of these basic carbon-nitrogen rings has different functional groups attached to it. In molecular biology shorthand, the nitrogenous bases are simply known by their symbols A, T, G, C, and U. DNA contains A, T, G, and C whereas RNA contains A, U, G, and C.

The pentose sugar in DNA is deoxyribose, and in RNA, the sugar is ribose (Figure). The difference between the sugars is the presence of the hydroxyl group on the second carbon of the ribose and hydrogen on the second carbon of the deoxyribose. The carbon atoms of the sugar molecule are numbered as 1′, 2′, 3′, 4′, and 5′ (1′ is read as “one prime”). The phosphate residue is attached to the hydroxyl group of the 5′ carbon of one sugar and the hydroxyl group of the 3′ carbon of the sugar of the next nucleotide, which forms a 5′–3′ phosphodiester linkage. The phosphodiester linkage is not formed by simple dehydration reaction like the other linkages connecting monomers in macromolecules: its formation involves the removal of two phosphate groups. A polynucleotide may have thousands of such phosphodiester linkages.

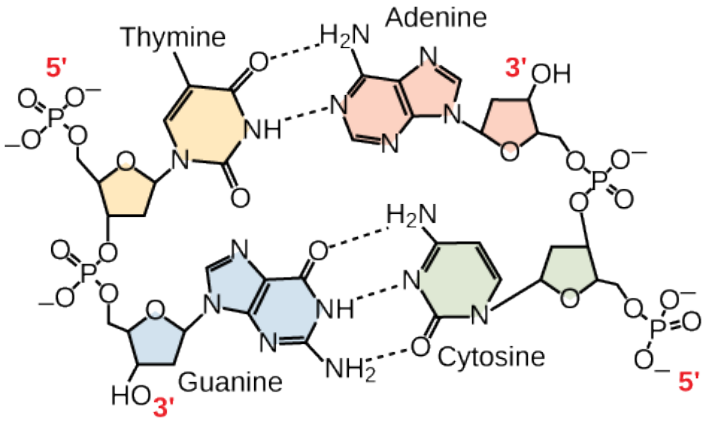
DNA Double-Helix Structure

DNA has a double-helix structure (Figure). The sugar and phosphate lie on the outside of the helix, forming the backbone of the DNA. The nitrogenous bases are stacked in the interior, like the steps of a staircase, in pairs; the pairs are bound to each other by hydrogen bonds. Every base pair in the double helivx is separated from the next base pair by 0.34 nm. The two strands of the helix run in opposite directions, meaning that the 5′ carbon end of one strand will face the 3′ carbon end of its matching strand. (This is referred to as antiparallel orientation and is important to DNA replication and in many nucleic acid interactions.)



Native DNA is an antiparallel double helix. The phosphate backbone (indicated by the curvy lines) is on the outside, and the bases are on the inside. Each base from one strand interacts via hydrogen bonding with a base from the opposing strand. (credit: Jerome Walker/Dennis Myts)

Only certain types of base pairing are allowed. For example, a certain purine can only pair with a certain pyrimidine. This means A can pair with T, and G can pair with C, as shown in Figure. This is known as the base complementary rule. In other words, the DNA strands are complementary to each other. If the sequence of one strand is AATTGGCC, the complementary strand would have the sequence TTAACCGG. During DNA replication, each strand is copied, resulting in a daughter DNA double helix containing one parental DNA strand and a newly synthesized strand.
ART CONNECTION



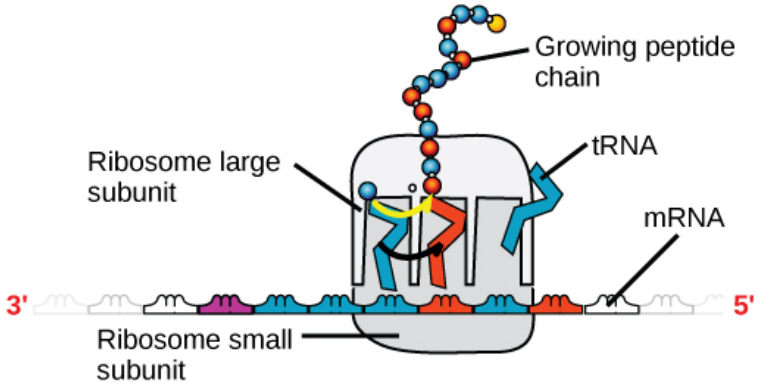
In a double stranded DNA molecule, the two strands run antiparallel to one another so that one strand runs 5′ to 3′ and the other 3′ to 5′. The phosphate backbone is located on the outside, and the bases are in the middle. Adenine forms hydrogen bonds (or base pairs) with thymine, and guanine base pairs with cytosine.

A mutation occurs, and cytosine is replaced with adenine. What impact do you think this will have on the DNA structure?

RNA

Ribonucleic acid, or RNA, is mainly involved in the process of protein synthesis under the direction of DNA. RNA is usually single-stranded and is made of ribonucleotides that are linked by phosphodiester bonds. A ribonucleotide in the RNA chain contains ribose (the pentose sugar), one of the four nitrogenous bases (A, U, G, and C), and the phosphate group.

There are four major types of RNA: messenger RNA (mRNA), ribosomal RNA (rRNA), transfer RNA (tRNA), and microRNA (miRNA). The first, mRNA, carries the message from DNA, which controls all of the cellular activities in a cell. If a cell requires a certain protein to be synthesized, the gene for this product is turned “on” and the messenger RNA is synthesized in the nucleus. The RNA base sequence is complementary to the coding sequence of the DNA from which it has been copied. However, in RNA, the base T is absent and U is present instead. If the DNA strand has a sequence AATTGCGC, the sequence of the complementary RNA is UUAACGCG. In the cytoplasm, the mRNA interacts with ribosomes and other cellular machinery (Figure).



A ribosome has two parts: a large subunit and a small subunit. The mRNA sits in between the two subunits. A tRNA molecule recognizes a codon on the mRNA, binds to it by complementary base pairing, and adds the correct amino acid to the growing peptide chain.

The mRNA is read in sets of three bases known as codons. Each codon codes for a single amino acid. In this way, the mRNA is read and the protein product is made. Ribosomal RNA (rRNA) is a major constituent of ribosomes on which the mRNA binds. The rRNA ensures the proper alignment of the mRNA and the ribosomes; the rRNA of the ribosome also has an enzymatic activity (peptidyl transferase) and catalyzes the formation of the peptide bonds between two aligned amino acids. Transfer RNA (tRNA) is one of the smallest of the four types of RNA, usually 70–90 nucleotides long. It carries the correct amino acid to the site of protein synthesis. It is the

base pairing between the tRNA and mRNA that allows for the correct amino acid to be inserted in the polypeptide chain. microRNAs are the smallest RNA molecules and their role involves the regulation of gene expression by interfering with the expression of certain mRNA messages. Table summarizes features of DNA and RNA.

Features of DNA and RNA		
	DNA	RNA
Function	Carries genetic information	Involved in protein synthesis
Location	Remains in the nucleus	Leaves the nucleus
Structure	Double helix	Usually single-stranded
Sugar	Deoxyribose	Ribose
Pyrimidines	Cytosine, thymine	Cytosine, uracil
Purines	Adenine, guanine	Adenine, guanine

Even though the RNA is single stranded, most RNA types show extensive intramolecular base pairing between complementary sequences, creating a predictable three-dimensional structure essential for their function.

As you have learned, information flow in an organism takes place from DNA to RNA to protein. DNA dictates the structure of mRNA in a process known as transcription, and RNA dictates the structure of protein in a process known as translation. This is known as the Central Dogma of Life, which holds true for all organisms; however, exceptions to the rule occur in connection with viral infections.

Link to Learning

To learn more about DNA, explore the Howard Hughes Medical Institute BioInteractive animations on the topic of DNA.

Section Summary

Nucleic acids are molecules made up of nucleotides that direct cellular activities such as cell division and protein synthesis. Each nucleotide is made up of a pentose sugar, a nitrogenous base, and a phosphate group. There are two types of nucleic acids: DNA and RNA. DNA carries the genetic blueprint of the cell and is passed on from parents to offspring (in the form of chromosomes). It has a double-helical structure with the two strands running in opposite directions, connected by hydrogen bonds, and complementary to each other. RNA is single-stranded and is made of a pentose sugar (ribose), a nitrogenous base, and a phosphate group. RNA is involved in protein synthesis and its regulation. Messenger RNA (mRNA) is copied from the DNA, is exported from the nucleus to the cytoplasm, and contains information for the construction of proteins. Ribosomal RNA (rRNA) is a part of the ribosomes at the site of protein synthesis, whereas transfer RNA (tRNA) carries the amino acid to the site of protein synthesis. microRNA regulates the use of mRNA for protein synthesis.

Art Connections

Figure A mutation occurs, and cytosine is replaced with adenine. What impact do you think this will have on the DNA structure?

Review Questions

A nucleotide of DNA may contain _____.

1. ribose, uracil, and a phosphate group
2. deoxyribose, uracil, and a phosphate group
3. deoxyribose, thymine, and a phosphate group
4. ribose, thymine, and a phosphate group

The building blocks of nucleic acids are _____.

1. sugars
2. nitrogenous bases
3. peptides
4. nucleotides

Free Response

What are the structural differences between RNA and DNA?

What are the four types of RNA and how do they function?

Glossary

deoxyribonucleic acid (DNA) double-helical molecule that carries the hereditary information of the cell

messenger RNA (mRNA) RNA that carries information from DNA to ribosomes during protein synthesis

nucleic acid biological macromolecule that carries the genetic blueprint of a cell and carries instructions for the functioning of the cell

nucleotide monomer of nucleic acids; contains a pentose sugar, one or more phosphate groups, and a nitrogenous base

phosphodiester linkage covalent chemical bond that holds together the polynucleotide chains with a phosphate group linking two pentose sugars of neighboring nucleotides

polynucleotide long chain of nucleotides

purine type of nitrogenous base in DNA and RNA; adenine and guanine are purines

pyrimidine type of nitrogenous base in DNA and RNA; cytosine, thymine, and uracil are pyrimidines

ribonucleic acid (RNA) single-stranded, often internally base paired, molecule that is involved in protein synthesis

ribosomal RNA (rRNA) RNA that ensures the proper alignment of the mRNA and the ribosomes during protein synthesis and catalyzes the formation of the peptide linkage

transcription process through which messenger RNA forms on a template of DNA

transfer RNA (tRNA) RNA that carries activated amino acids to the site of protein synthesis on the ribosome

translation process through which RNA directs the formation of protein

Licensing & Attributions

CC licensed content, Shared previously

- nucleic acids. **Authored by**: open stax. **Provided by**: open stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:yxeAKc4X@8/Nucleic-Acids. **License**: *CC BY: Attribution*

STUDY GUIDE: MACROMOLECULES

Study Questions

Objective: Describe the molecular basis of life.

Use this page to check your understanding of the content.

Vocabulary:

1. Monosaccharaides
2. Disaccharaides
3. Polysaccharaides
4. amino acid
5. triglyceride
6. hydrolysis
7. dehydration synthesis
8. nucleic acid
9. primary, secondary, teritary and quatnary strucutre
10. enzyme
11. isomers
12. functional groups
13. ATP
14. purine and pyramidine
15. Hydrophilic
16. Hydrophobic

Study Guide Questions

1. Give examples of monosaccharaides, disaccaraides, and polysaccharaides.
2. Compare and contrast the following polysaccharaides: Glycogen, starch, chitin, cellulose
3. What is the monomer that makes up each of the following classes of bio molecules what type of bonds link the mononers?
    1. Carbohydrates
    2. Lipids
    3. Nucleic acids
    4. Proteins
4. Compare and contrast saturated, unsaturated, and trans-fatty acids.
5. Describe the 4 different levels of protein structure.
6. What happens when a protein is denatured?
7. What factors can result in protein denaturation?
8. List some important functions of proteins.
9. Describe the structure and function of the 4 major classes of biomolecules. Answer this question to the detail covered in lecture.

Licensing & Attributions

CC licensed content, Shared previously

- Study Guide: Chemistry of Life. **Authored by**: Wendy Riggs. **Provided by**: College of the Redwoods. **Project**: Kaleidoscope. **License**: *CC BY: Attribution*

CELL STRUCTURE AND FUNCTION

INTRODUCTION

Close your eyes and picture a brick wall. What is the basic building block of that wall? It is a single brick, of course. Like a brick wall, your body is composed of basic building blocks, and the building blocks of your body are cells.

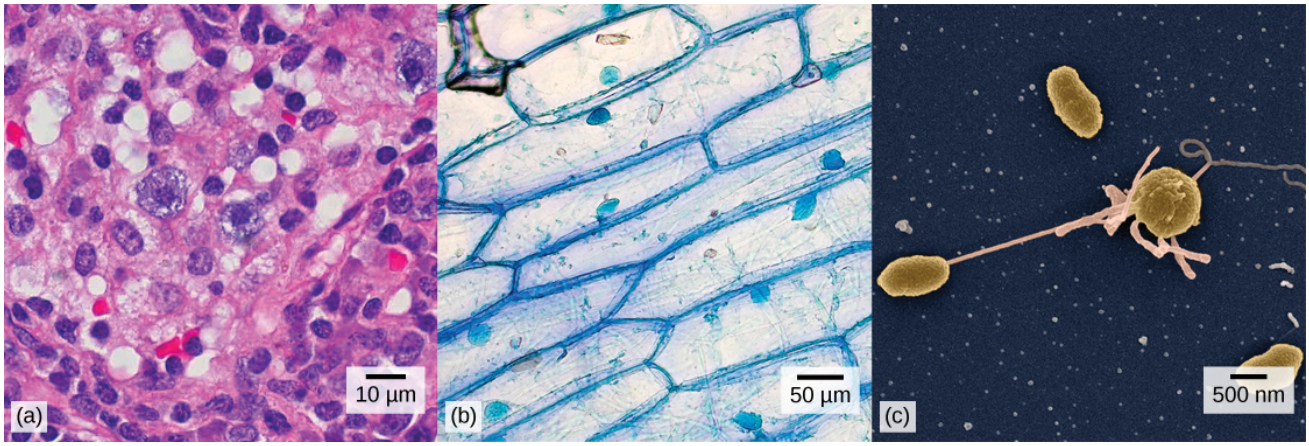


Figure 1. (a) Nasal sinus cells (viewed with a light microscope), (b) onion cells (viewed with a light microscope), and (c) Vibrio tasmaniensis bacterial cells (viewed using a scanning electron microscope) are from very different organism, yet all share certain characteristics of basic cell structure. (credit a: modification of work by Ed Uthman, MD; credit b: modification of work by Umberto Salvagnin; credit c: modification of work by Anthony D'Onofrio; scale-bar data from Matt Russell)

Your body has many kinds of cells, each specialized for a specific purpose. Just as a home is made from a variety of building materials, the human body is constructed from many cell types. For example, epithelial cells protect the surface of the body and cover the organs and body cavities within. Bone cells help to support and protect the body. Cells of the immune system fight invading bacteria. Additionally, red blood cells carry oxygen throughout the body. Each of these cell types plays a vital role during the growth, development, and day-to-day maintenance of the body. In spite of their enormous variety, however, all cells share certain fundamental characteristics.

Licensing & Attributions

CC licensed content, Shared previously

- Concepts of Biology. **Authored by**: Open Stax. **Located at**: http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@8.10:1/Concepts_of_Biology. **License**: *CC BY: Attribution*

STUDYING CELLS

4.1 Studying Cells

Summary

A cell is the smallest unit of a living thing. A living thing, whether made of one cell (like bacteria) or many cells (like a human), is called an organism. Thus, cells are the basic building blocks of all organisms.

Several cells of one kind that interconnect with each other and perform a shared function form tissues, several tissues combine to form an organ (your stomach, heart, or brain), and several organs make up an organ system (such as the digestive system, circulatory system, or nervous system). Several systems that function together form an organism (like a human being). Here, we will examine the structure and function of cells.

There are many types of cells, all grouped into one of two broad categories: prokaryotic and eukaryotic. For example, both animal and plant cells are classified as eukaryotic cells, whereas bacterial cells are classified as prokaryotic. Before discussing the criteria for determining whether a cell is prokaryotic or eukaryotic, let's first examine how biologists study cells.

Microscopy

Cells vary in size. With few exceptions, individual cells cannot be seen with the naked eye, so scientists use microscopes (micro- = "small"; -scope = "to look at") to study them. A microscope is an instrument that magnifies an object. Most photographs of cells are taken with a microscope, and these images can also be called micrographs.

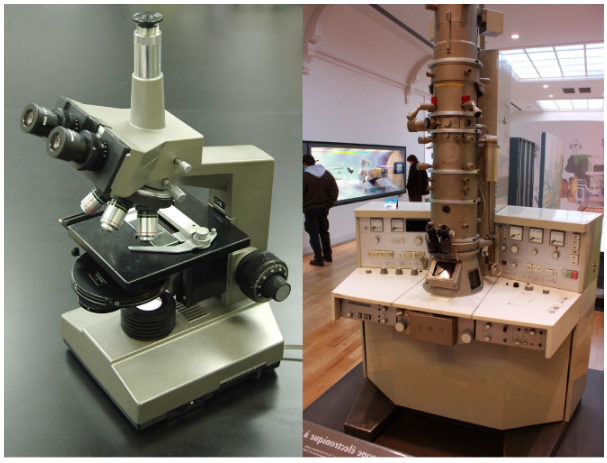
The optics of a microscope's lenses change the orientation of the image that the user sees. A specimen that is right-side up and facing right on the microscope slide will appear upside-down and facing left when viewed through a microscope, and vice versa. Similarly, if the slide is moved left while looking through the microscope, it will appear to move right, and if moved down, it will seem to move up. This occurs because microscopes use two sets of lenses to magnify the image. Because of the manner by which light travels through the lenses, this system of two lenses produces an inverted image (binocular, or dissecting microscopes, work in a similar manner, but include an additional magnification system that makes the final image appear to be upright).

Light Microscopes

To give you a sense of cell size, a typical human red blood cell is about eight millionths of a meter or eight micrometers (abbreviated as eight μm) in diameter; the head of a pin of is about two thousandths of a meter (two mm) in diameter. That means about 250 red blood cells could fit on the head of a pin.

Most student microscopes are classified as light microscopes (Figure**a**). Visible light passes and is bent through the lens system to enable the user to see the specimen. Light microscopes are advantageous for viewing living organisms, but since individual cells are generally transparent, their components are not distinguishable unless they are colored with special stains. Staining, however, usually kills the cells.

Light microscopes commonly used in the undergraduate college laboratory magnify up to approximately 400 times. Two parameters that are important in microscopy are magnification and resolving power. Magnification is the process of enlarging an object in appearance. Resolving power is the ability of a microscope to distinguish two adjacent structures as separate: the higher the resolution, the better the clarity and detail of the image. When oil immersion lenses are used for the study of small objects, magnification is usually increased to 1,000 times. In order to gain a better understanding of cellular structure and function, scientists typically use electron microscopes.



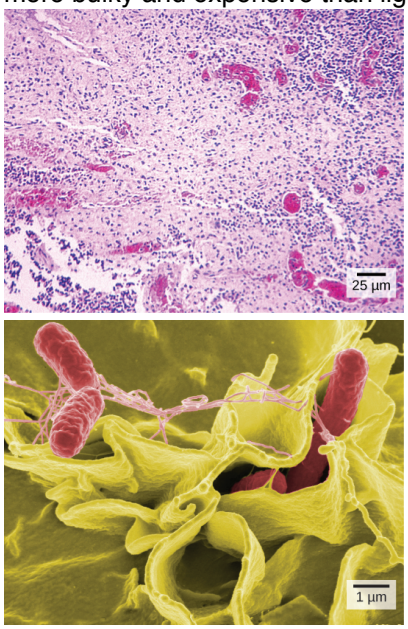
(a) (b)

(a) Most light microscopes used in a college biology lab can magnify cells up to approximately 400 times and have a resolution of about 200 nanometers. (b) Electron microscopes provide a much higher magnification, 100,000x, and a have a resolution of 50 picometers. (credit a: modification of work by “GcG”/Wikimedia Commons; credit b: modification of work by Evan Bench)

Electron Microscopes

In contrast to light microscopes, electron microscopes (Figureb) use a beam of electrons instead of a beam of light. Not only does this allow for higher magnification and, thus, more detail (Figure), it also provides higher resolving power. The method used to prepare the specimen for viewing with an electron microscope kills the specimen. Electrons have short wavelengths (shorter than photons) that move best in a vacuum, so living cells cannot be viewed with an electron microscope.

In a scanning electron microscope, a beam of electrons moves back and forth across a cell’s surface, creating details of cell surface characteristics. In a transmission electron microscope, the electron beam penetrates the cell and provides details of a cell’s internal structures. As you might imagine, electron microscopes are significantly more bulky and expensive than light microscopes.



(a) These *Salmonella* bacteria appear as tiny purple dots when viewed with a light microscope. (b) This scanning electron microscope micrograph shows *Salmonella* bacteria (in red) invading human cells (yellow). Even though subfigure (b) shows a different*Salmonella* specimen than subfigure (a), you can still observe the comparative increase in magnification and detail. (credit a: modification of work by CDC/Armed Forces Institute of Pathology,

Charles N. Farmer, Rocky Mountain Laboratories; credit b: modification of work by NIAID, NIH; scale-bar data from Matt Russell)

Link to Learning

For another perspective on cell size, try the HowBig interactive at this site.

Cell Theory

The microscopes we use today are far more complex than those used in the 1600s by Antony van Leeuwenhoek, a Dutch shopkeeper who had great skill in crafting lenses. Despite the limitations of his now-ancient lenses, van Leeuwenhoek observed the movements of protista (a type of single-celled organism) and sperm, which he collectively termed “animalcules.”

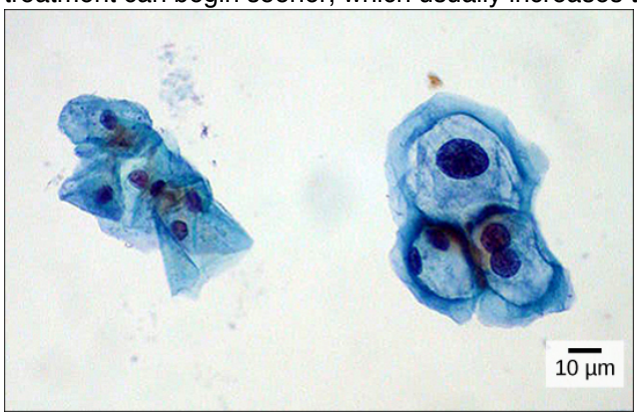
In a 1665 publication called *Micrographia*, experimental scientist Robert Hooke coined the term “cell” for the box-like structures he observed when viewing cork tissue through a lens. In the 1670s, van Leeuwenhoek discovered bacteria and protozoa. Later advances in lenses, microscope construction, and staining techniques enabled other scientists to see some components inside cells.

By the late 1830s, botanist Matthias Schleiden and zoologist Theodor Schwann were studying tissues and proposed the unified cell theory, which states that all living things are composed of one or more cells, the cell is the basic unit of life, and new cells arise from existing cells. Rudolf Virchow later made important contributions to this theory.
CAREER CONNECTION

CytotechnologistHave you ever heard of a medical test called a Pap smear (Figure)? In this test, a doctor takes a small sample of cells from the uterine cervix of a patient and sends it to a medical lab where a cytotechnologist stains the cells and examines them for any changes that could indicate cervical cancer or a microbial infection.

Cytotechnologists (cyto- = “cell”) are professionals who study cells via microscopic examinations and other laboratory tests. They are trained to determine which cellular changes are within normal limits and which are abnormal. Their focus is not limited to cervical cells; they study cellular specimens that come from all organs. When they notice abnormalities, they consult a pathologist, who is a medical doctor who can make a clinical diagnosis.

Cytotechnologists play a vital role in saving people’s lives. When abnormalities are discovered early, a patient’s treatment can begin sooner, which usually increases the chances of a successful outcome.



These uterine cervix cells, viewed through a light microscope, were obtained from a Pap smear. Normal cells are on the left. The cells on the right are infected with human papillomavirus (HPV). Notice that the infected cells are larger; also, two of these cells each have two nuclei instead of one, the normal number. (credit: modification of work by Ed Uthman, MD; scale-bar data from Matt Russell)

Section Summary

A cell is the smallest unit of life. Most cells are so tiny that they cannot be seen with the naked eye. Therefore, scientists use microscopes to study cells. Electron microscopes provide higher magnification, higher resolution, and more detail than light microscopes. The unified cell theory states that all organisms are composed of one or more cells, the cell is the basic unit of life, and new cells arise from existing cells.

Review Questions

When viewing a specimen through a light microscope, scientists use _____ to distinguish the individual components of cells.

1. a beam of electrons
2. radioactive isotopes
3. special stains
4. high temperatures

The _____ is the basic unit of life.

1. organism
2. cell
3. tissue
4. organ

Free Response

In your everyday life, you have probably noticed that certain instruments are ideal for certain situations. For example, you would use a spoon rather than a fork to eat soup because a spoon is shaped for scooping, while soup would slip between the tines of a fork. The use of ideal instruments also applies in science. In what situation(s) would the use of a light microscope be ideal, and why?

In what situation(s) would the use of a scanning electron microscope be ideal, and why?

In what situation(s) would a transmission electron microscope be ideal, and why?

What are the advantages and disadvantages of each of these types of microscopes?

Glossary

cell theory see unified cell theory

electron microscope an instrument that magnifies an object using a beam of electrons passed and bent through a lens system to visualize a specimen

light microscope an instrument that magnifies an object using a beam visible light passed and bent through a lens system to visualize a specimen

microscope an instrument that magnifies an object

unified cell theory a biological concept that states that all organisms are composed of one or more cells; the cell is the basic unit of life; and new cells arise from existing cells

Licensing & Attributions

CC licensed content, Shared previously

- studying cells. **Authored by**: open stax. **Provided by**: open stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:ADrtuvNU@8/Studying-Cells. **License**: *CC BY: Attribution*

PROKARYOTIC CELLS

4.2 Prokaryotic Cells

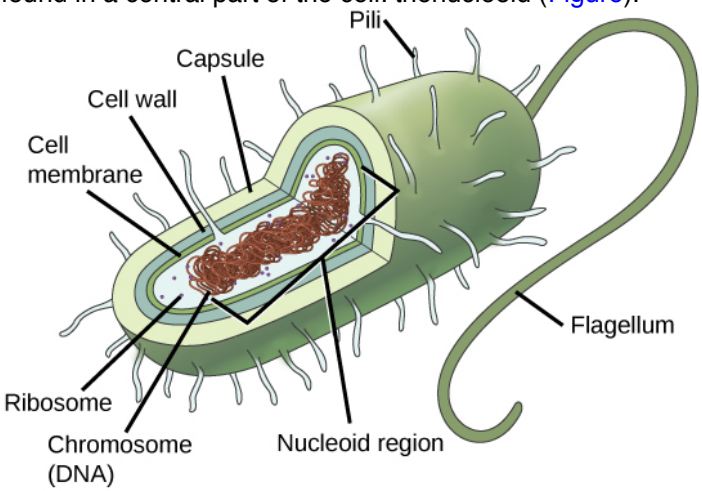
Summary

Cells fall into one of two broad categories: prokaryotic and eukaryotic. Only the predominantly single-celled organisms of the domains Bacteria and Archaea are classified as prokaryotes (pro- = “before”; -kary- = “nucleus”). Cells of animals, plants, fungi, and protists are all eukaryotes (ceu- = “true”) and are made up of eukaryotic cells.

Components of Prokaryotic Cells

All cells share four common components: 1) a plasma membrane, an outer covering that separates the cell’s interior from its surrounding environment; 2) cytoplasm, consisting of a jelly-like cytosol within the cell in which other cellular components are found; 3) DNA, the genetic material of the cell; and 4) ribosomes, which synthesize proteins. However, prokaryotes differ from eukaryotic cells in several ways.

A prokaryote is a simple, mostly single-celled (unicellular) organism that lacks a nucleus, or any other membrane-bound organelle. We will shortly come to see that this is significantly different in eukaryotes. Prokaryotic DNA is found in a central part of the cell: thenucleoid (Figure).



This figure shows the generalized structure of a prokaryotic cell. All prokaryotes have chromosomal DNA localized in a nucleoid, ribosomes, a cell membrane, and a cell wall. The other structures shown are present in some, but not all, bacteria.

Most prokaryotes have a peptidoglycan cell wall and many have a polysaccharide capsule (Figure). The cell wall acts as an extra layer of protection, helps the cell maintain its shape, and prevents dehydration. The capsule enables the cell to attach to surfaces in its environment. Some prokaryotes have flagella, pili, or fimbriae. Flagella are used for locomotion. Pili are used to exchange genetic material during a type of reproduction called conjugation. Fimbriae are used by bacteria to attach to a host cell.
CAREER CONNECTION

MicrobiologistThe most effective action anyone can take to prevent the spread of contagious illnesses is to wash his or her hands. Why? Because microbes (organisms so tiny that they can only be seen with microscopes) are ubiquitous. They live on doorknobs, money, your hands, and many other surfaces. If someone sneezes into his hand and touches a doorknob, and afterwards you touch that same doorknob, the microbes from the sneezer’s

mucus are now on your hands. If you touch your hands to your mouth, nose, or eyes, those microbes can enter your body and could make you sick.

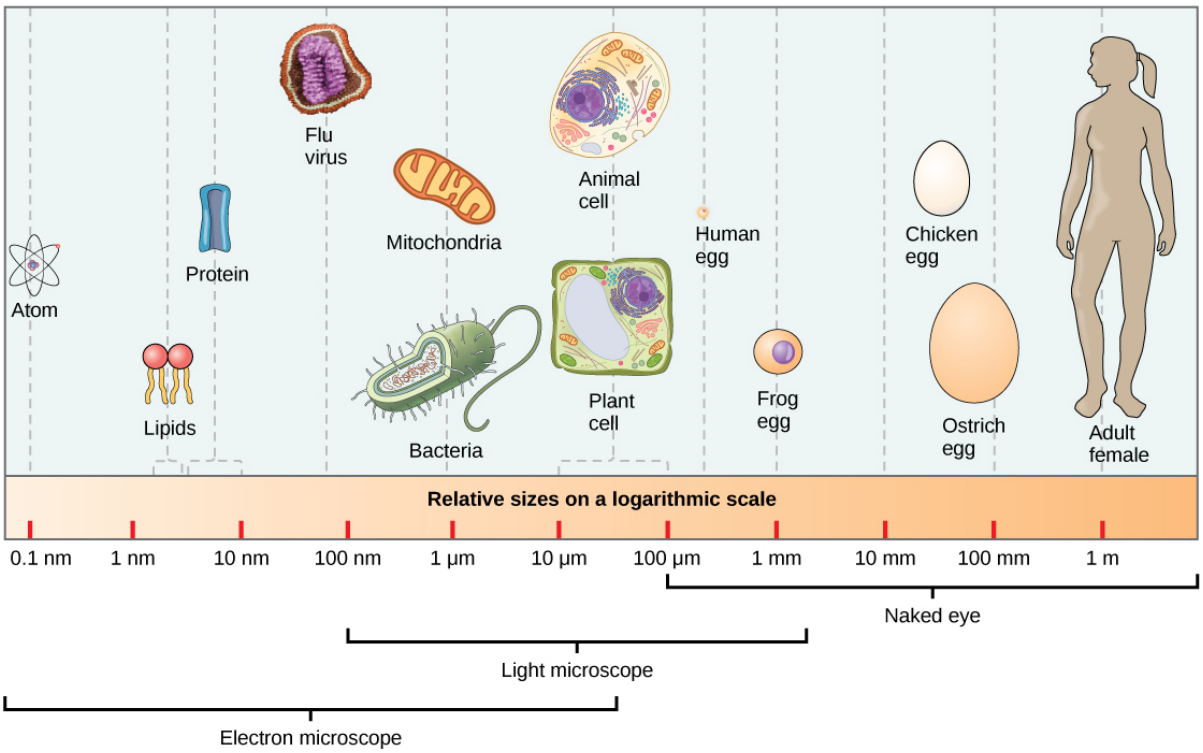
However, not all microbes (also called microorganisms) cause disease; most are actually beneficial. You have microbes in your gut that make vitamin K. Other microorganisms are used to ferment beer and wine.

Microbiologists are scientists who study microbes. Microbiologists can pursue a number of careers. Not only do they work in the food industry, they are also employed in the veterinary and medical fields. They can work in the pharmaceutical sector, serving key roles in research and development by identifying new sources of antibiotics that could be used to treat bacterial infections.

Environmental microbiologists may look for new ways to use specially selected or genetically engineered microbes for the removal of pollutants from soil or groundwater, as well as hazardous elements from contaminated sites. These uses of microbes are called bioremediation technologies. Microbiologists can also work in the field of bioinformatics, providing specialized knowledge and insight for the design, development, and specificity of computer models of, for example, bacterial epidemics.

Cell Size

At 0.1 to 5.0 µm in diameter, prokaryotic cells are significantly smaller than eukaryotic cells, which have diameters ranging from 10 to 100 µm (Figure). The small size of prokaryotes allows ions and organic molecules that enter them to quickly diffuse to other parts of the cell. Similarly, any wastes produced within a prokaryotic cell can quickly diffuse out. This is not the case in eukaryotic cells, which have developed different structural adaptations to enhance intracellular transport.

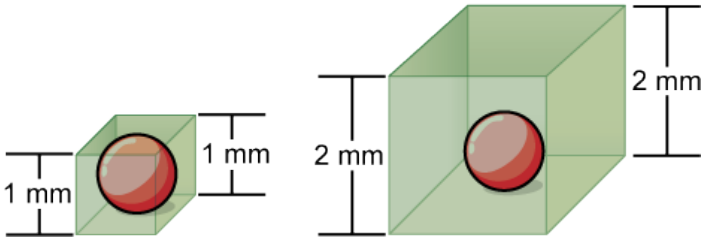


This figure shows relative sizes of microbes on a logarithmic scale (recall that each unit of increase in a logarithmic scale represents a 10-fold increase in the quantity being measured).

Small size, in general, is necessary for all cells, whether prokaryotic or eukaryotic. Let's examine why that is so. First, we'll consider the area and volume of a typical cell. Not all cells are spherical in shape, but most tend to approximate a sphere. You may remember from your high school geometry course that the formula for the surface area of a sphere is $4\pi r^2$, while the formula for its volume is $4\pi r^3/3$. Thus, as the radius of a cell increases, its surface area increases as the square of its radius, but its volume increases as the cube of its radius (much more rapidly). Therefore, as a cell increases in size, its surface area-to-volume ratio decreases. This same principle would apply if the cell had the shape of a cube (Figure). If the cell grows too large, the plasma

membrane will not have sufficient surface area to support the rate of diffusion required for the increased volume. In other words, as a cell grows, it becomes less efficient. One way to become more efficient is to divide; another way is to develop organelles that perform specific tasks. These adaptations lead to the development of more sophisticated cells called eukaryotic cells.

ART CONNECTION



Notice that as a cell increases in size, its surface area-to-volume ratio decreases. When there is insufficient surface area to support a cell's increasing volume, a cell will either divide or die. The cell on the left has a volume of 1 mm^3 and a surface area of 6 mm^2, with a surface area-to-volume ratio of 6 to 1, whereas the cell on the right has a volume of 8 mm^3 and a surface area of 24 mm^2, with a surface area-to-volume ratio of 3 to 1.

Prokaryotic cells are much smaller than eukaryotic cells. What advantages might small cell size confer on a cell? What advantages might large cell size have?

Section Summary

Prokaryotes are predominantly single-celled organisms of the domains Bacteria and Archaea. All prokaryotes have plasma membranes, cytoplasm, ribosomes, and DNA that is not membrane-bound. Most have peptidoglycan cell walls and many have polysaccharide capsules. Prokaryotic cells range in diameter from 0.1 to 5.0 µm.

As a cell increases in size, its surface area-to-volume ratio decreases. If the cell grows too large, the plasma membrane will not have sufficient surface area to support the rate of diffusion required for the increased volume.

Art Connections

Figure Prokaryotic cells are much smaller than eukaryotic cells. What advantages might small cell size confer on a cell? What advantages might large cell size have?

Review Questions

Prokaryotes depend on _____ to obtain some materials and to get rid of wastes.

1. ribosomes
2. flagella
3. cell division
4. diffusion

Bacteria that lack fimbriae are less likely to _____.

1. adhere to cell surfaces
2. swim through bodily fluids
3. synthesize proteins
4. retain the ability to divide

Free Response

Antibiotics are medicines that are used to fight bacterial infections. These medicines kill prokaryotic cells without harming human cells. What part or parts of the bacterial cell do you think antibiotics target? Why?

Explain why not all microbes are harmful.

Glossary

nucleoid central part of a prokaryotic cell in which the chromosome is found

prokaryote unicellular organism that lacks a nucleus or any other membrane-bound organelle

Licensing & Attributions

CC licensed content, Shared previously

- prokaryotic cells. **Authored by**: open stax. **Provided by**: open stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:pOpVdIwp@11/Prokaryotic-Cells. **License**: *CC BY: Attribution*

EUKARYOTIC CELLS

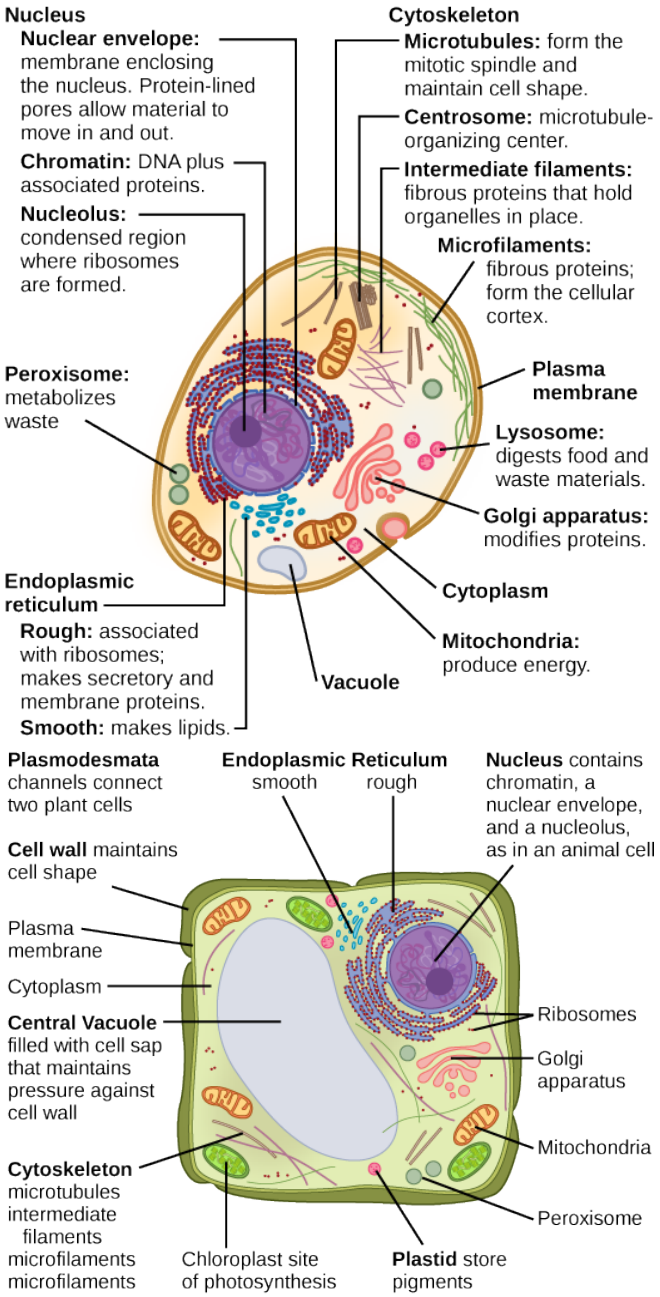
4.3 Eukaryotic Cells

Summary

Have you ever heard the phrase “form follows function?” It’s a philosophy practiced in many industries. In architecture, this means that buildings should be constructed to support the activities that will be carried out inside them. For example, a skyscraper should be built with several elevator banks; a hospital should be built so that its emergency room is easily accessible.

Our natural world also utilizes the principle of form following function, especially in cell biology, and this will become clear as we explore eukaryotic cells (Figure). Unlike prokaryotic cells, eukaryotic cells have: 1) a membrane-bound nucleus; 2) numerous membrane-bound organelles such as the endoplasmic reticulum, Golgi apparatus, chloroplasts, mitochondria, and others; and 3) several, rod-shaped chromosomes. Because a eukaryotic cell’s nucleus is surrounded by a membrane, it is often said to have a “true nucleus.” The word “organelle” means “little organ,” and, as already mentioned, organelles have specialized cellular functions, just as the organs of your body have specialized functions.

At this point, it should be clear to you that eukaryotic cells have a more complex structure than prokaryotic cells. Organelles allow different functions to be compartmentalized in different areas of the cell. Before turning to organelles, let’s first examine two important components of the cell: the plasma membrane and the cytoplasm.
ART CONNECTION

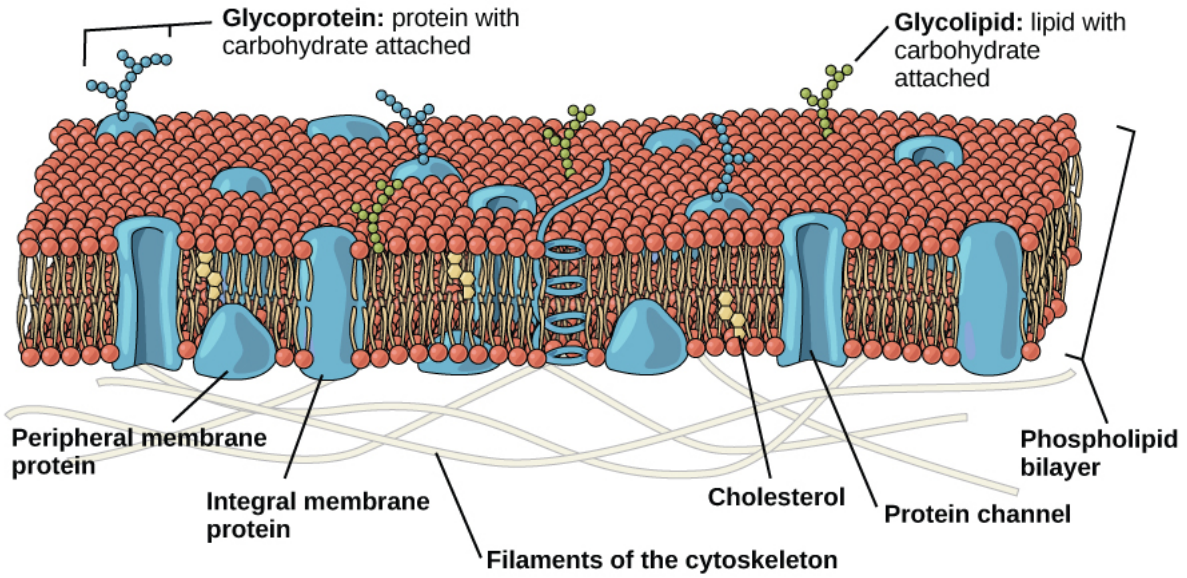


These figures show the major organelles and other cell components of (a) a typical animal cell and (b) a typical eukaryotic plant cell. The plant cell has a cell wall, chloroplasts, plastids, and a central vacuole—structures not found in animal cells. Plant cells do not have lysosomes or centrosomes.

If the nucleolus were not able to carry out its function, what other cellular organelles would be affected?

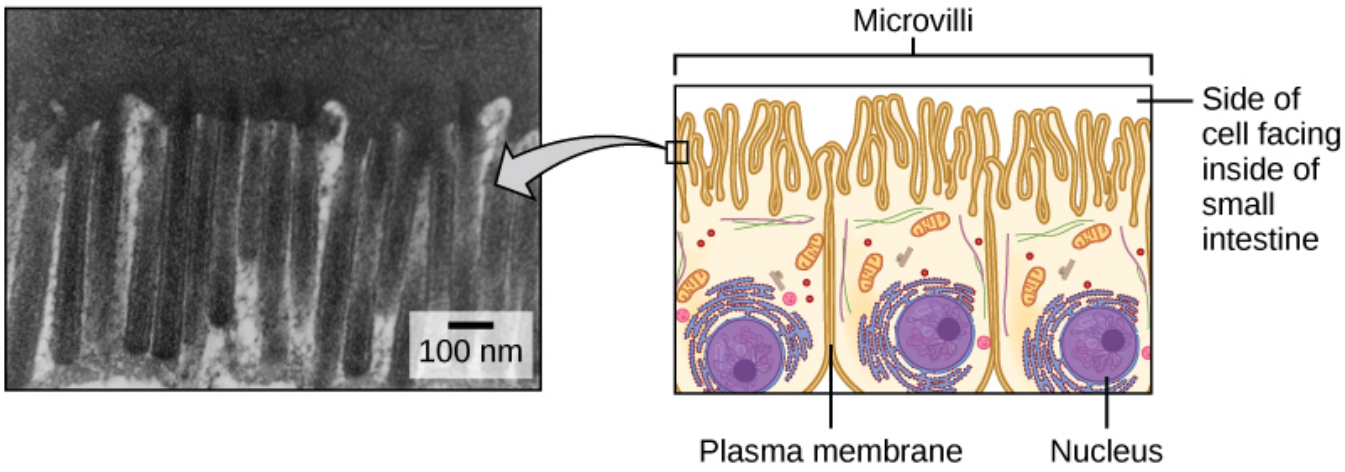
The Plasma Membrane

Like prokaryotes, eukaryotic cells have a plasma membrane (Figure), a phospholipid bilayer with embedded proteins that separates the internal contents of the cell from its surrounding environment. A phospholipid is a lipid molecule with two fatty acid chains and a phosphate-containing group. The plasma membrane controls the passage of organic molecules, ions, water, and oxygen into and out of the cell. Wastes (such as carbon dioxide and ammonia) also leave the cell by passing through the plasma membrane.



The eukaryotic plasma membrane is a phospholipid bilayer with proteins and cholesterol embedded in it.

The plasma membranes of cells that specialize in absorption are folded into fingerlike projections called microvilli (singular = microvillus); (Figure). Such cells are typically found lining the small intestine, the organ that absorbs nutrients from digested food. This is an excellent example of form following function. People with celiac disease have an immune response to gluten, which is a protein found in wheat, barley, and rye. The immune response damages microvilli, and thus, afflicted individuals cannot absorb nutrients. This leads to malnutrition, cramping, and diarrhea. Patients suffering from celiac disease must follow a gluten-free diet.



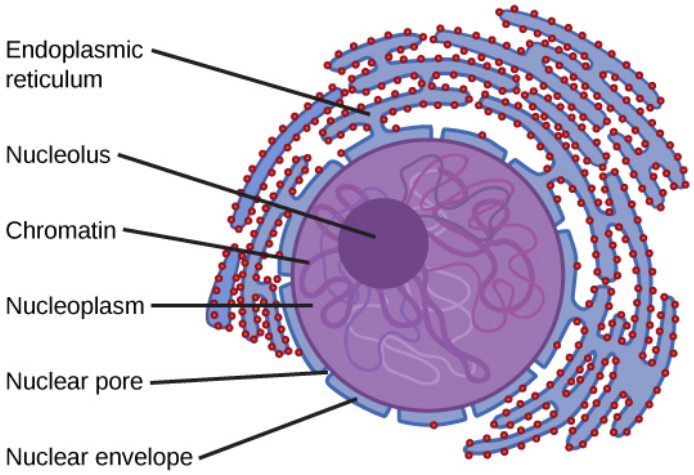
Microvilli, shown here as they appear on cells lining the small intestine, increase the surface area available for absorption. These microvilli are only found on the area of the plasma membrane that faces the cavity from which substances will be absorbed. (credit “micrograph”: modification of work by Louisa Howard)

The Cytoplasm

The cytoplasm is the entire region of a cell between the plasma membrane and the nuclear envelope (a structure to be discussed shortly). It is made up of organelles suspended in the gel-like cytosol, the cytoskeleton, and various chemicals (Figure). Even though the cytoplasm consists of 70 to 80 percent water, it has a semi-solid consistency, which comes from the proteins within it. However, proteins are not the only organic molecules found in the cytoplasm. Glucose and other simple sugars, polysaccharides, amino acids, nucleic acids, fatty acids, and derivatives of glycerol are found there, too. Ions of sodium, potassium, calcium, and many other elements are also dissolved in the cytoplasm. Many metabolic reactions, including protein synthesis, take place in the cytoplasm.

The Nucleus

Typically, the nucleus is the most prominent organelle in a cell (Figure). The nucleus (plural = nuclei) houses the cell's DNA and directs the synthesis of ribosomes and proteins. Let's look at it in more detail (Figure).



The nucleus stores chromatin (DNA plus proteins) in a gel-like substance called the nucleoplasm. The nucleolus is a condensed region of chromatin where ribosome synthesis occurs. The boundary of the nucleus is called the nuclear envelope. It consists of two phospholipid bilayers: an outer membrane and an inner membrane. The nuclear membrane is continuous with the endoplasmic reticulum. Nuclear pores allow substances to enter and exit the nucleus.

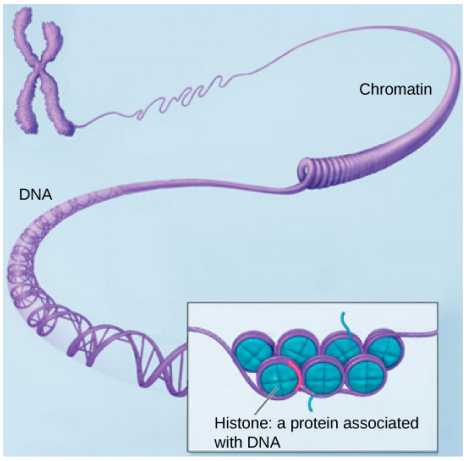
The Nuclear Envelope

The nuclear envelope is a double-membrane structure that constitutes the outermost portion of the nucleus (Figure). Both the inner and outer membranes of the nuclear envelope are phospholipid bilayers.

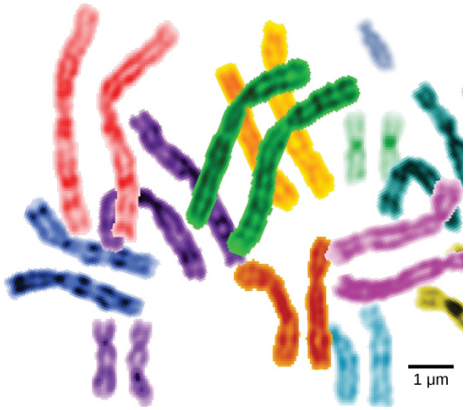
The nuclear envelope is punctuated with pores that control the passage of ions, molecules, and RNA between the nucleoplasm and cytoplasm. The nucleoplasm is the semi-solid fluid inside the nucleus, where we find the chromatin and the nucleolus.

Chromatin and Chromosomes

To understand chromatin, it is helpful to first consider chromosomes. Chromosomes are structures within the nucleus that are made up of DNA, the hereditary material. You may remember that in prokaryotes, DNA is organized into a single circular chromosome. In eukaryotes, chromosomes are linear structures. Every eukaryotic species has a specific number of chromosomes in the nuclei of its body's cells. For example, in humans, the chromosome number is 46, while in fruit flies, it is eight. Chromosomes are only visible and distinguishable from one another when the cell is getting ready to divide. When the cell is in the growth and maintenance phases of its life cycle, proteins are attached to chromosomes, and they resemble an unwound, jumbled bunch of threads. These unwound protein-chromosome complexes are called chromatin (Figure); chromatin describes the material that makes up the chromosomes both when condensed and decondensed.



Chromosomes



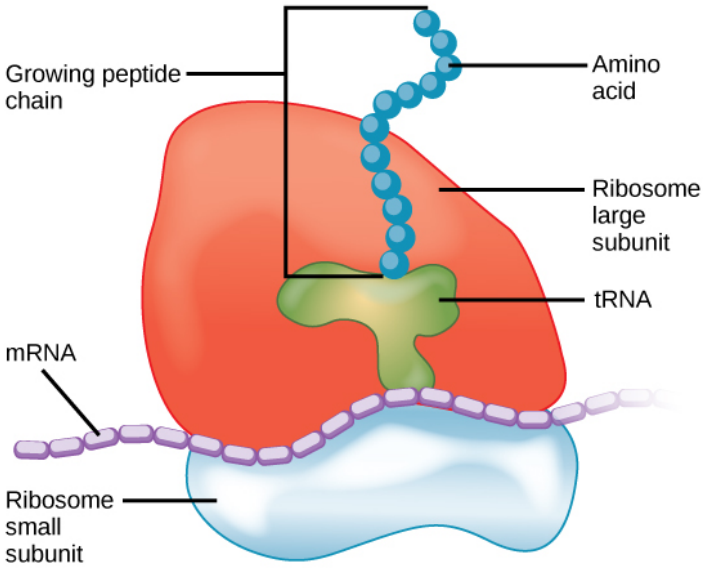
(a) This image shows various levels of the organization of chromatin (DNA and protein). (b) This image shows paired chromosomes. (credit b: modification of work by NIH; scale-bar data from Matt Russell)

The Nucleolus

We already know that the nucleus directs the synthesis of ribosomes, but how does it do this? Some chromosomes have sections of DNA that encode ribosomal RNA. A darkly staining area within the nucleus called the nucleolus (plural = nucleoli) aggregates the ribosomal RNA with associated proteins to assemble the ribosomal subunits that are then transported out through the pores in the nuclear envelope to the cytoplasm.

Ribosomes

Ribosomes are the cellular structures responsible for protein synthesis. When viewed through an electron microscope, ribosomes appear either as clusters (polyribosomes) or single, tiny dots that float freely in the cytoplasm. They may be attached to the cytoplasmic side of the plasma membrane or the cytoplasmic side of the endoplasmic reticulum and the outer membrane of the nuclear envelope (Figure). Electron microscopy has shown us that ribosomes, which are large complexes of protein and RNA, consist of two subunits, aptly called large and small (Figure). Ribosomes receive their “orders” for protein synthesis from the nucleus where the DNA is transcribed into messenger RNA (mRNA). The mRNA travels to the ribosomes, which translate the code provided by the sequence of the nitrogenous bases in the mRNA into a specific order of amino acids in a protein. Amino acids are the building blocks of proteins.



Ribosomes are made up of a large subunit (top) and a small subunit (bottom). During protein synthesis, ribosomes assemble amino acids into proteins.

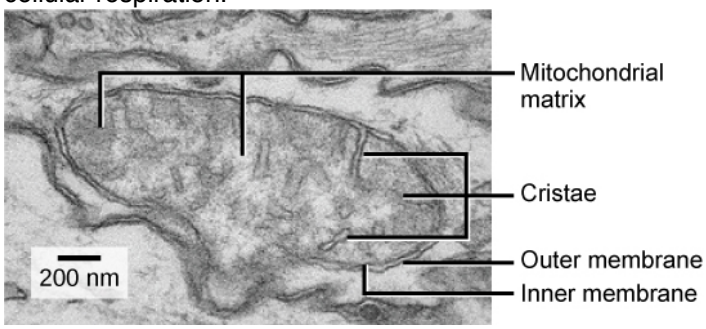
Because proteins synthesis is an essential function of all cells (including enzymes, hormones, antibodies, pigments, structural components, and surface receptors), ribosomes are found in practically every cell. Ribosomes are particularly abundant in cells that synthesize large amounts of protein. For example, the pancreas is responsible for creating several digestive enzymes and the cells that produce these enzymes contain many ribosomes. Thus, we see another example of form following function.

Mitochondria

Mitochondria (singular = mitochondrion) are often called the “powerhouses” or “energy factories” of a cell because they are responsible for making adenosine triphosphate (ATP), the cell’s main energy-carrying molecule. ATP represents the short-term stored energy of the cell. Cellular respiration is the process of making ATP using the chemical energy found in glucose and other nutrients. In mitochondria, this process uses oxygen and produces carbon dioxide as a waste product. In fact, the carbon dioxide that you exhale with every breath comes from the cellular reactions that produce carbon dioxide as a byproduct.

In keeping with our theme of form following function, it is important to point out that muscle cells have a very high concentration of mitochondria that produce ATP. Your muscle cells need a lot of energy to keep your body moving. When your cells don’t get enough oxygen, they do not make a lot of ATP. Instead, the small amount of ATP they make in the absence of oxygen is accompanied by the production of lactic acid.

Mitochondria are oval-shaped, double membrane organelles (Figure) that have their own ribosomes and DNA. Each membrane is a phospholipid bilayer embedded with proteins. The inner layer has folds called cristae. The area surrounded by the folds is called the mitochondrial matrix. The cristae and the matrix have different roles in cellular respiration.



This electron micrograph shows a mitochondrion as viewed with a transmission electron microscope. This organelle has an outer membrane and an inner membrane. The inner membrane contains folds, called cristae,

which increase its surface area. The space between the two membranes is called the intermembrane space, and the space inside the inner membrane is called the mitochondrial matrix. ATP synthesis takes place on the inner membrane. (credit: modification of work by Matthew Britton; scale-bar data from Matt Russell)

Peroxisomes

Peroxisomes are small, round organelles enclosed by single membranes. They carry out oxidation reactions that break down fatty acids and amino acids. They also detoxify many poisons that may enter the body. (Many of these oxidation reactions release hydrogen peroxide, H_2O_2, which would be damaging to cells; however, when these reactions are confined to peroxisomes, enzymes safely break down the H_2O_2 into oxygen and water.) For example, alcohol is detoxified by peroxisomes in liver cells. Glyoxysomes, which are specialized peroxisomes in plants, are responsible for converting stored fats into sugars.

Vesicles and Vacuoles

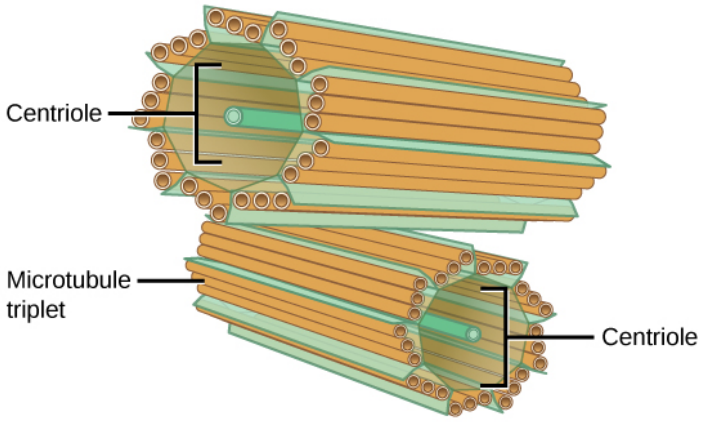
Vesicles and vacuoles are membrane-bound sacs that function in storage and transport. Other than the fact that vacuoles are somewhat larger than vesicles, there is a very subtle distinction between them: The membranes of vesicles can fuse with either the plasma membrane or other membrane systems within the cell. Additionally, some agents such as enzymes within plant vacuoles break down macromolecules. The membrane of a vacuole does not fuse with the membranes of other cellular components.

Animal Cells versus Plant Cells

At this point, you know that each eukaryotic cell has a plasma membrane, cytoplasm, a nucleus, ribosomes, mitochondria, peroxisomes, and in some, vacuoles, but there are some striking differences between animal and plant cells. While both animal and plant cells have microtubule organizing centers (MTOCs), animal cells also have centrioles associated with the MTOC: a complex called the centrosome. Animal cells each have a centrosome and lysosomes, whereas plant cells do not. Plant cells have a cell wall, chloroplasts and other specialized plastids, and a large central vacuole, whereas animal cells do not.

The Centrosome

The centrosome is a microtubule-organizing center found near the nuclei of animal cells. It contains a pair of centrioles, two structures that lie perpendicular to each other (Figure). Each centriole is a cylinder of nine triplets of microtubules.



The centrosome consists of two centrioles that lie at right angles to each other. Each centriole is a cylinder made up of nine triplets of microtubules. Nontubulin proteins (indicated by the green lines) hold the microtubule triplets together.

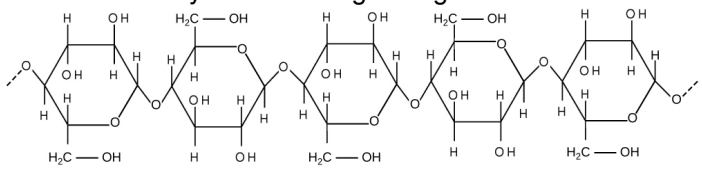
The centrosome (the organelle where all microtubules originate) replicates itself before a cell divides, and the centrioles appear to have some role in pulling the duplicated chromosomes to opposite ends of the dividing cell. However, the exact function of the centrioles in cell division isn't clear, because cells that have had the centrosome removed can still divide, and plant cells, which lack centrosomes, are capable of cell division.

Lysosomes

Animal cells have another set of organelles not found in plant cells: lysosomes. The lysosomes are the cell's "garbage disposal." In plant cells, the digestive processes take place in vacuoles. Enzymes within the lysosomes aid the breakdown of proteins, polysaccharides, lipids, nucleic acids, and even worn-out organelles. These enzymes are active at a much lower pH than that of the cytoplasm. Therefore, the pH within lysosomes is more acidic than the pH of the cytoplasm. Many reactions that take place in the cytoplasm could not occur at a low pH, so again, the advantage of compartmentalizing the eukaryotic cell into organelles is apparent.

The Cell Wall

If you examine Figureb, the diagram of a plant cell, you will see a structure external to the plasma membrane called the cell wall. The cell wall is a rigid covering that protects the cell, provides structural support, and gives shape to the cell. Fungal and protistan cells also have cell walls. While the chief component of prokaryotic cell walls is peptidoglycan, the major organic molecule in the plant cell wall is cellulose (Figure), a polysaccharide made up of glucose units. Have you ever noticed that when you bite into a raw vegetable, like celery, it crunches? That's because you are tearing the rigid cell walls of the celery cells with your teeth.

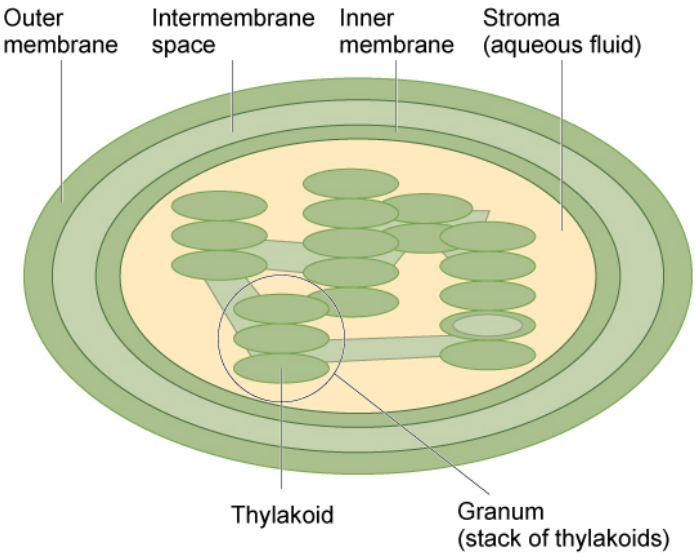


Cellulose is a long chain of β-glucose molecules connected by a 1-4 linkage. The dashed lines at each end of the figure indicate a series of many more glucose units. The size of the page makes it impossible to portray an entire cellulose molecule.

Chloroplasts

Like the mitochondria, chloroplasts have their own DNA and ribosomes, but chloroplasts have an entirely different function.Chloroplasts are plant cell organelles that carry out photosynthesis. Photosynthesis is the series of reactions that use carbon dioxide, water, and light energy to make glucose and oxygen. This is a major difference between plants and animals; plants (autotrophs) are able to make their own food, like sugars, while animals (heterotrophs) must ingest their food.

Like mitochondria, chloroplasts have outer and inner membranes, but within the space enclosed by a chloroplast's inner membrane is a set of interconnected and stacked fluid-filled membrane sacs called thylakoids (Figure). Each stack of thylakoids is called a granum (plural = grana). The fluid enclosed by the inner membrane that surrounds the grana is called the stroma.



The chloroplast has an outer membrane, an inner membrane, and membrane structures called thylakoids that are stacked into grana. The space inside the thylakoid membranes is called the thylakoid space. The light harvesting reactions take place in the thylakoid membranes, and the synthesis of sugar takes place in the fluid inside the inner membrane, which is called the stroma. Chloroplasts also have their own genome, which is contained on a single circular chromosome.

The chloroplasts contain a green pigment called chlorophyll, which captures the light energy that drives the reactions of photosynthesis. Like plant cells, photosynthetic protists also have chloroplasts. Some bacteria perform photosynthesis, but their chlorophyll is not relegated to an organelle.
EVOLUTION CONNECTION

EndosymbiosisWe have mentioned that both mitochondria and chloroplasts contain DNA and ribosomes. Have you wondered why? Strong evidence points to endosymbiosis as the explanation.

Symbiosis is a relationship in which organisms from two separate species depend on each other for their survival. Endosymbiosis (endo- = “within”) is a mutually beneficial relationship in which one organism lives inside the other. Endosymbiotic relationships abound in nature. We have already mentioned that microbes that produce vitamin K live inside the human gut. This relationship is beneficial for us because we are unable to synthesize vitamin K. It is also beneficial for the microbes because they are protected from other organisms and from drying out, and they receive abundant food from the environment of the large intestine.

Scientists have long noticed that bacteria, mitochondria, and chloroplasts are similar in size. We also know that bacteria have DNA and ribosomes, just as mitochondria and chloroplasts do. Scientists believe that host cells and bacteria formed an endosymbiotic relationship when the host cells ingested both aerobic and autotrophic bacteria (cyanobacteria) but did not destroy them. Through many millions of years of evolution, these ingested bacteria became more specialized in their functions, with the aerobic bacteria becoming mitochondria and the autotrophic bacteria becoming chloroplasts.

The Central Vacuole

Previously, we mentioned vacuoles as essential components of plant cells. If you look at Figureb, you will see that plant cells each have a large central vacuole that occupies most of the area of the cell. The central vacuole plays a key role in regulating the cell’s concentration of water in changing environmental conditions. Have you ever noticed that if you forget to water a plant for a few days, it wilts? That’s because as the water concentration in the soil becomes lower than the water concentration in the plant, water moves out of the central vacuoles and cytoplasm. As the central vacuole shrinks, it leaves the cell wall unsupported. This loss of support to the cell walls of plant cells results in the wilted appearance of the plant.

The central vacuole also supports the expansion of the cell. When the central vacuole holds more water, the cell gets larger without having to invest a lot of energy in synthesizing new cytoplasm.

Section Summary

Like a prokaryotic cell, a eukaryotic cell has a plasma membrane, cytoplasm, and ribosomes, but a eukaryotic cell is typically larger than a prokaryotic cell, has a true nucleus (meaning its DNA is surrounded by a membrane), and has other membrane-bound organelles that allow for compartmentalization of functions. The plasma membrane is a phospholipid bilayer embedded with proteins. The nucleus's nucleolus is the site of ribosome assembly. Ribosomes are either found in the cytoplasm or attached to the cytoplasmic side of the plasma membrane or endoplasmic reticulum. They perform protein synthesis. Mitochondria participate in cellular respiration; they are responsible for the majority of ATP produced in the cell. Peroxisomes hydrolyze fatty acids, amino acids, and some toxins. Vesicles and vacuoles are storage and transport compartments. In plant cells, vacuoles also help break down macromolecules.

Animal cells also have a centrosome and lysosomes. The centrosome has two bodies perpendicular to each other, the centrioles, and has an unknown purpose in cell division. Lysosomes are the digestive organelles of animal cells.

Plant cells and plant-like cells each have a cell wall, chloroplasts, and a central vacuole. The plant cell wall, whose primary component is cellulose, protects the cell, provides structural support, and gives shape to the cell. Photosynthesis takes place in chloroplasts. The central vacuole can expand without having to produce more cytoplasm.

Art Connections

Figure If the nucleolus were not able to carry out its function, what other cellular organelles would be affected?

Review Questions

Which of the following is surrounded by two phospholipid bilayers?

1. the ribosomes
2. the vesicles
3. the cytoplasm
4. the nucleoplasm

Peroxisomes got their name because hydrogen peroxide is:

1. used in their detoxification reactions
2. produced during their oxidation reactions
3. incorporated into their membranes
4. a cofactor for the organelles' enzymes

In plant cells, the function of the lysosomes is carried out by _____.

1. vacuoles
2. peroxisomes
3. ribosomes
4. nuclei

Which of the following is found both in eukaryotic and prokaryotic cells?

1. nucleus
2. mitochondrion
3. vacuole
4. ribosomes

Free Response

You already know that ribosomes are abundant in red blood cells. In what other cells of the body would you find them in great abundance? Why?

What are the structural and functional similarities and differences between mitochondria and chloroplasts?

Glossary

cell wall rigid cell covering made of various molecules that protects the cell, provides structural support, and gives shape to the cell

central vacuole large plant cell organelle that regulates the cell's storage compartment, holds water, and plays a significant role in cell growth as the site of macromolecule degradation

centrosome region in animal cells made of two centrioles

chlorophyll green pigment that captures the light energy that drives the light reactions of photosynthesis

chloroplast plant cell organelle that carries out photosynthesis

chromatin protein-DNA complex that serves as the building material of chromosomes

chromosome structure within the nucleus that is made up of chromatin that contains DNA, the hereditary material

cytoplasm entire region between the plasma membrane and the nuclear envelope, consisting of organelles suspended in the gel-like cytosol, the cytoskeleton, and various chemicals

cytosol gel-like material of the cytoplasm in which cell structures are suspended

eukaryotic cell cell that has a membrane-bound nucleus and several other membrane-bound compartments or sacs

lysosome organelle in an animal cell that functions as the cell's digestive component; it breaks down proteins, polysaccharides, lipids, nucleic acids, and even worn-out organelles

mitochondria (singular = mitochondrion) cellular organelles responsible for carrying out cellular respiration, resulting in the production of ATP, the cell's main energy-carrying molecule

nuclear envelope double-membrane structure that constitutes the outermost portion of the nucleus

nucleolus darkly staining body within the nucleus that is responsible for assembling the subunits of the ribosomes

nucleoplasm semi-solid fluid inside the nucleus that contains the chromatin and nucleolus

nucleus cell organelle that houses the cell's DNA and directs the synthesis of ribosomes and proteins

organelle compartment or sac within a cell

peroxisome small, round organelle that contains hydrogen peroxide, oxidizes fatty acids and amino acids, and detoxifies many poisons

plasma membrane phospholipid bilayer with embedded (integral) or attached (peripheral) proteins, and separates the internal content of the cell from its surrounding environment

ribosome cellular structure that carries out protein synthesis

vacuole membrane-bound sac, somewhat larger than a vesicle, which functions in cellular storage and transport

vesicle small, membrane-bound sac that functions in cellular storage and transport; its membrane is capable of fusing with the plasma membrane and the membranes of the endoplasmic reticulum and Golgi apparatus

Licensing & Attributions

CC licensed content, Shared previously

- eukaryotic cells. **Authored by**: open stax. **Provided by**: open stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:FPF-phhT@14/Eukaryotic-Cells. **License**: *CC BY: Attribution*

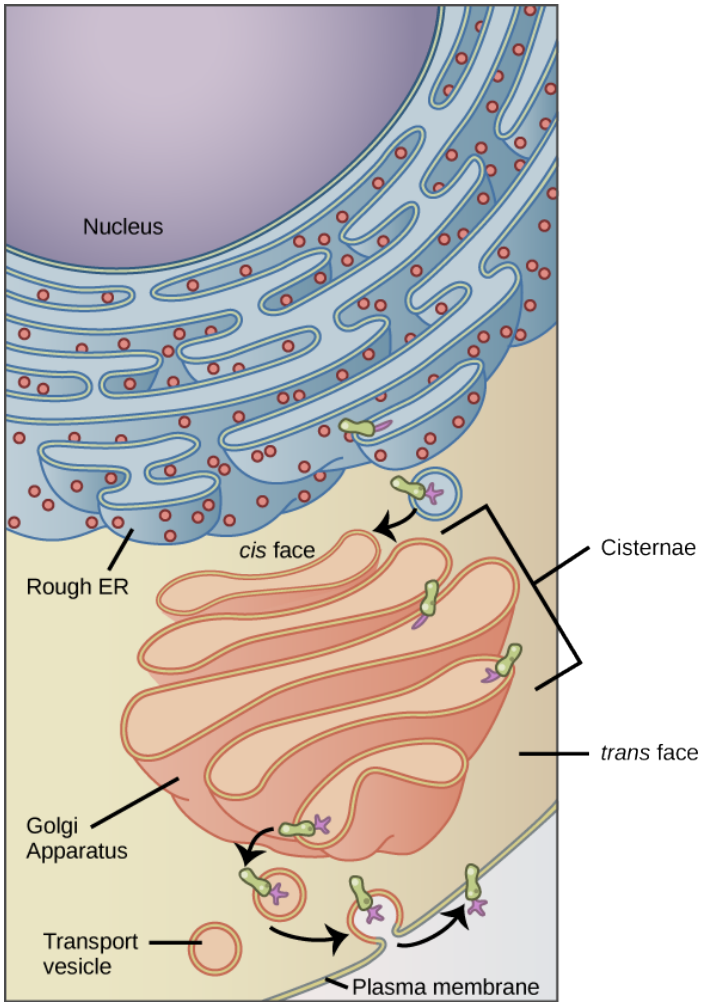
THE ENDOMEMBRANE SYSTEM

4.4 The Endomembrane System and Proteins

Summary

The endomembrane system (endo = “within”) is a group of membranes and organelles (Figure) in eukaryotic cells that works together to modify, package, and transport lipids and proteins. It includes the nuclear envelope, lysosomes, and vesicles, which we’ve already mentioned, and the endoplasmic reticulum and Golgi apparatus, which we will cover shortly. Although not technically *within* the cell, the plasma membrane is included in the endomembrane system because, as you will see, it interacts with the other endomembranous organelles. The endomembrane system does not include the membranes of either mitochondria or chloroplasts.

ART CONNECTION



“Membrane and secretory proteins are synthesized in the rough endoplasmic reticulum (RER). The RER also sometimes modifies proteins. In this illustration, a (green) integral membrane protein in the ER is modified by attachment of a (purple) carbohydrate. Vesicles with the integral protein bud from the ER and fuse with the cis

face of the Golgi apparatus. As the protein passes along the Golgi's cisternae, it is further modified by the addition of more carbohydrates. After its synthesis is complete, it exits as integral membrane protein of the vesicle that bud from the Golgi's trans face and when the vesicle fuses with the cell membrane the protein becomes integral portion of that cell membrane. (credit: modification of work by Magnus Manske)

If a peripheral membrane protein were synthesized in the lumen (inside) of the ER, would it end up on the inside or outside of the plasma membrane?

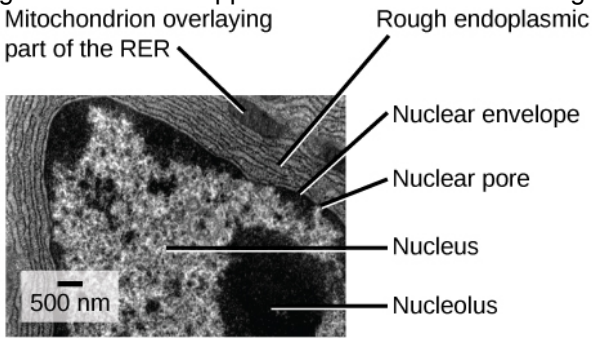
The Endoplasmic Reticulum

The endoplasmic reticulum (ER) (Figure) is a series of interconnected membranous sacs and tubules that collectively modifies proteins and synthesizes lipids. However, these two functions are performed in separate areas of the ER: the rough ER and the smooth ER, respectively.

The hollow portion of the ER tubules is called the lumen or cisternal space. The membrane of the ER, which is a phospholipid bilayer embedded with proteins, is continuous with the nuclear envelope.

Rough ER

The rough endoplasmic reticulum (RER) is so named because the ribosomes attached to its cytoplasmic surface give it a studded appearance when viewed through an electron microscope (Figure).



This transmission electron micrograph shows the rough endoplasmic reticulum and other organelles in a pancreatic cell. (credit: modification of work by Louisa Howard)

Ribosomes transfer their newly synthesized proteins into the lumen of the RER where they undergo structural modifications, such as folding or the acquisition of side chains. These modified proteins will be incorporated into cellular membranes—the membrane of the ER or those of other organelles—or secreted from the cell (such as protein hormones, enzymes). The RER also makes phospholipids for cellular membranes.

If the phospholipids or modified proteins are not destined to stay in the RER, they will reach their destinations via transport vesicles that bud from the RER's membrane (Figure).

Since the RER is engaged in modifying proteins (such as enzymes, for example) that will be secreted from the cell, you would be correct in assuming that the RER is abundant in cells that secrete proteins. This is the case with cells of the liver, for example.

Smooth ER

The smooth endoplasmic reticulum (SER) is continuous with the RER but has few or no ribosomes on its cytoplasmic surface (Figure). Functions of the SER include synthesis of carbohydrates, lipids, and steroid hormones; detoxification of medications and poisons; and storage of calcium ions.

In muscle cells, a specialized SER called the sarcoplasmic reticulum is responsible for storage of the calcium ions that are needed to trigger the coordinated contractions of the muscle cells.

Link to Learning

You can watch an excellent animation of the endomembrane system here. At the end of the animation, there is a short self-assessment.
CAREER CONNECTION

CardiologistHeart disease is the leading cause of death in the United States. This is primarily due to our sedentary lifestyle and our high trans-fat diets.

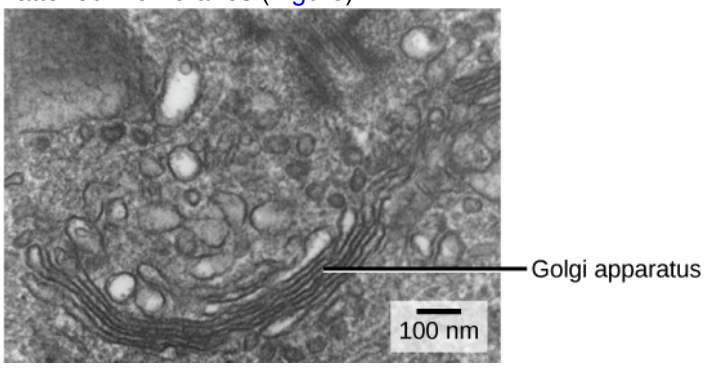
Heart failure is just one of many disabling heart conditions. Heart failure does not mean that the heart has stopped working. Rather, it means that the heart can't pump with sufficient force to transport oxygenated blood to all the vital organs. Left untreated, heart failure can lead to kidney failure and failure of other organs.

The wall of the heart is composed of cardiac muscle tissue. Heart failure occurs when the endoplasmic reticula of cardiac muscle cells do not function properly. As a result, an insufficient number of calcium ions are available to trigger a sufficient contractile force.

Cardiologists (cardi- = "heart"; -ologist = "one who studies") are doctors who specialize in treating heart diseases, including heart failure. Cardiologists can make a diagnosis of heart failure via physical examination, results from an electrocardiogram (ECG, a test that measures the electrical activity of the heart), a chest X-ray to see whether the heart is enlarged, and other tests. If heart failure is diagnosed, the cardiologist will typically prescribe appropriate medications and recommend a reduction in table salt intake and a supervised exercise program.

The Golgi Apparatus

We have already mentioned that vesicles can bud from the ER and transport their contents elsewhere, but where do the vesicles go? Before reaching their final destination, the lipids or proteins within the transport vesicles still need to be sorted, packaged, and tagged so that they wind up in the right place. Sorting, tagging, packaging, and distribution of lipids and proteins takes place in theGolgi apparatus (also called the Golgi body), a series of flattened membranes (Figure).



The Golgi apparatus in this white blood cell is visible as a stack of semicircular, flattened rings in the lower portion of the image. Several vesicles can be seen near the Golgi apparatus. (credit: modification of work by Louisa Howard)

The receiving side of the Golgi apparatus is called the *cis* face. The opposite side is called the *trans* face. The transport vesicles that formed from the ER travel to the *cis* face, fuse with it, and empty their contents into the lumen of the Golgi apparatus. As the proteins and lipids travel through the Golgi, they undergo further modifications that allow them to be sorted. The most frequent modification is the addition of short chains of sugar molecules. These newly modified proteins and lipids are then tagged with phosphate groups or other small molecules so that they can be routed to their proper destinations.

Finally, the modified and tagged proteins are packaged into secretory vesicles that bud from the *trans* face of the Golgi. While some of these vesicles deposit their contents into other parts of the cell where they will be used, other secretory vesicles fuse with the plasma membrane and release their contents outside the cell.

In another example of form following function, cells that engage in a great deal of secretory activity (such as cells of the salivary glands that secrete digestive enzymes or cells of the immune system that secrete antibodies) have an abundance of Golgi.

In plant cells, the Golgi apparatus has the additional role of synthesizing polysaccharides, some of which are incorporated into the cell wall and some of which are used in other parts of the cell.
CAREER CONNECTION

GeneticistMany diseases arise from genetic mutations that prevent the synthesis of critical proteins. One such disease is Lowe disease (also called oculocerebrorenal syndrome, because it affects the eyes, brain, and kidneys). In Lowe disease, there is a deficiency in an enzyme localized to the Golgi apparatus. Children with Lowe disease are born with cataracts, typically develop kidney disease after the first year of life, and may have impaired mental abilities.

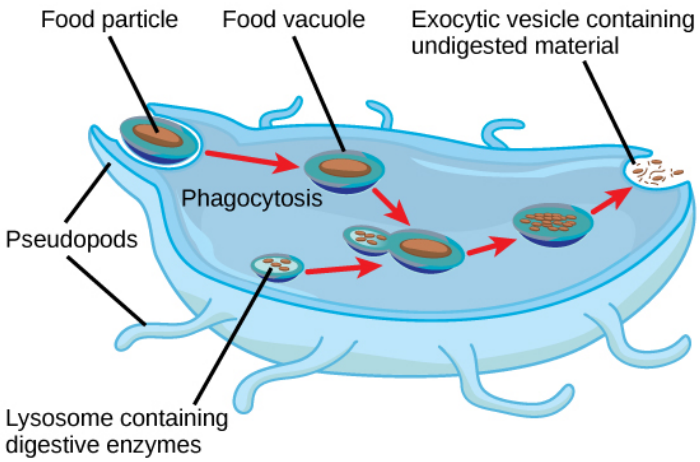
Lowe disease is a genetic disease caused by a mutation on the X chromosome. The X chromosome is one of the two human sex chromosome, as these chromosomes determine a person's sex. Females possess two X chromosomes while males possess one X and one Y chromosome. In females, the genes on only one of the two X chromosomes are expressed. Therefore, females who carry the Lowe disease gene on one of their X chromosomes have a 50/50 chance of having the disease. However, males only have one X chromosome and the genes on this chromosome are always expressed. Therefore, males will always have Lowe disease if their X chromosome carries the Lowe disease gene. The location of the mutated gene, as well as the locations of many other mutations that cause genetic diseases, has now been identified. Through prenatal testing, a woman can find out if the fetus she is carrying may be afflicted with one of several genetic diseases.

Geneticists analyze the results of prenatal genetic tests and may counsel pregnant women on available options. They may also conduct genetic research that leads to new drugs or foods, or perform DNA analyses that are used in forensic investigations.

Lysosomes

In addition to their role as the digestive component and organelle-recycling facility of animal cells, lysosomes are considered to be parts of the endomembrane system. Lysosomes also use their hydrolytic enzymes to destroy pathogens (disease-causing organisms) that might enter the cell. A good example of this occurs in a group of white blood cells called macrophages, which are part of your body's immune system. In a process known as phagocytosis or endocytosis, a section of the plasma membrane of the macrophage invaginates (folds in) and engulfs a pathogen. The invaginated section, with the pathogen inside, then pinches itself off from the plasma membrane and becomes a vesicle. The vesicle fuses with a lysosome. The lysosome's hydrolytic enzymes then destroy the pathogen (Figure).

Phagocytosis



A macrophage has engulfed (phagocytized) a potentially pathogenic bacterium and then fuses with a lysosomes within the cell to destroy the pathogen. Other organelles are present in the cell but for simplicity are not shown.

Section Summary

The endomembrane system includes the nuclear envelope, lysosomes, vesicles, the ER, and Golgi apparatus, as well as the plasma membrane. These cellular components work together to modify, package, tag, and transport proteins and lipids that form the membranes.

The RER modifies proteins and synthesizes phospholipids used in cell membranes. The SER synthesizes carbohydrates, lipids, and steroid hormones; engages in the detoxification of medications and poisons; and stores calcium ions. Sorting, tagging, packaging, and distribution of lipids and proteins take place in the Golgi apparatus. Lysosomes are created by the budding of the membranes of the RER and Golgi. Lysosomes digest macromolecules, recycle worn-out organelles, and destroy pathogens.

Art Connections

Figure If a peripheral membrane protein were synthesized in the lumen (inside) of the ER, would it end up on the inside or outside of the plasma membrane?

Review Questions

Which of the following is not a component of the endomembrane system?

1. mitochondrion
2. Golgi apparatus
3. endoplasmic reticulum
4. lysosome

The process by which a cell engulfs a foreign particle is known as:

1. endosymbiosis
2. phagocytosis
3. hydrolysis
4. membrane synthesis

Which of the following is most likely to have the greatest concentration of smooth endoplasmic reticulum?

1. a cell that secretes enzymes
2. a cell that destroys pathogens
3. a cell that makes steroid hormones
4. a cell that engages in photosynthesis

Which of the following sequences correctly lists in order the steps involved in the incorporation of a proteinaceous molecule within a cell?

1. synthesis of the protein on the ribosome; modification in the Golgi apparatus; packaging in the endoplasmic reticulum; tagging in the vesicle
2. synthesis of the protein on the lysosome; tagging in the Golgi; packaging in the vesicle; distribution in the endoplasmic reticulum
3. synthesis of the protein on the ribosome; modification in the endoplasmic reticulum; tagging in the Golgi; distribution via the vesicle
4. synthesis of the protein on the lysosome; packaging in the vesicle; distribution via the Golgi; tagging in the endoplasmic reticulum

Free Response

In the context of cell biology, what do we mean by form follows function? What are at least two examples of this concept?

In your opinion, is the nuclear membrane part of the endomembrane system? Why or why not? Defend your answer.

Glossary

endomembrane system group of organelles and membranes in eukaryotic cells that work together modifying, packaging, and transporting lipids and proteins

endoplasmic reticulum (ER) series of interconnected membranous structures within eukaryotic cells that collectively modify proteins and synthesize lipids

Golgi apparatus eukaryotic organelle made up of a series of stacked membranes that sorts, tags, and packages lipids and proteins for distribution

rough endoplasmic reticulum (RER) region of the endoplasmic reticulum that is studded with ribosomes and engages in protein modification and phospholipid synthesis

smooth endoplasmic reticulum (SER) region of the endoplasmic reticulum that has few or no ribosomes on its cytoplasmic surface and synthesizes carbohydrates, lipids, and steroid hormones; detoxifies certain chemicals (like pesticides, preservatives, medications, and environmental pollutants), and stores calcium ions

- Downloads
- History
- Attribution
- More Information

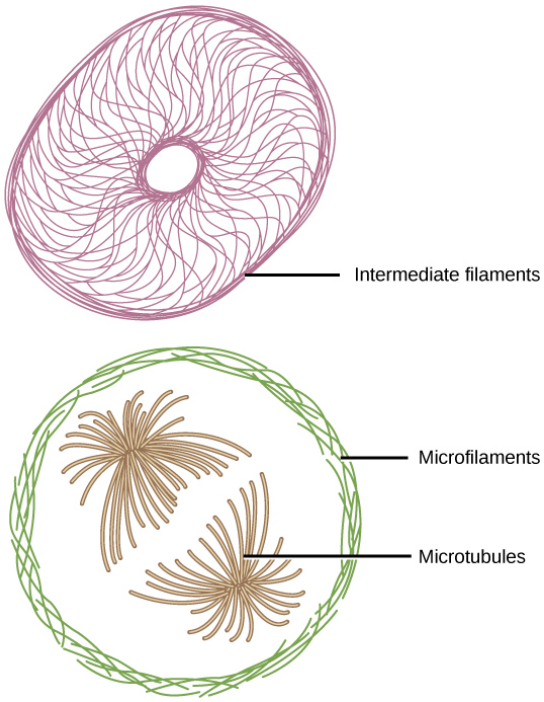
Licensing & Attributions

CC licensed content, Original

- the endomembrane system. **Authored by**: open stax. **Provided by**: open stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:ELU3gCIJ@12/The-Endomembrane-System-and-Pr. **License**: *CC BY: Attribution*

THE CYTOSKELETON

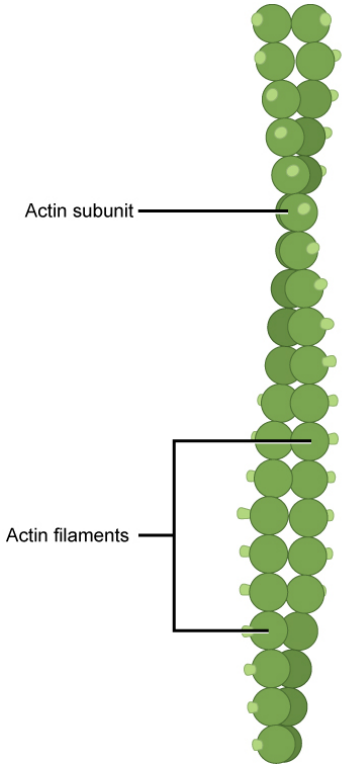
If you were to remove all the organelles from a cell, would the plasma membrane and the cytoplasm be the only components left? No. Within the cytoplasm, there would still be ions and organic molecules, plus a network of protein fibers that help maintain the shape of the cell, secure some organelles in specific positions, allow cytoplasm and vesicles to move within the cell, and enable cells within multicellular organisms to move. Collectively, this network of protein fibers is known as the cytoskeleton. There are three types of fibers within the cytoskeleton: microfilaments, intermediate filaments, and microtubules (Figure). Here, we will examine each.



Microfilaments thicken the cortex around the inner edge of a cell; like rubber bands, they resist tension. Microtubules are found in the interior of the cell where they maintain cell shape by resisting compressive forces. Intermediate filaments are found throughout the cell and hold organelles in place.

Microfilaments

Of the three types of protein fibers in the cytoskeleton, microfilaments are the narrowest. They function in cellular movement, have a diameter of about 7 nm, and are made of two intertwined strands of a globular protein called actin (Figure). For this reason, microfilaments are also known as actin filaments.



Microfilaments are made of two intertwined strands of actin.

Actin is powered by ATP to assemble its filamentous form, which serves as a track for the movement of a motor protein called myosin. This enables actin to engage in cellular events requiring motion, such as cell division in animal cells and cytoplasmic streaming, which is the circular movement of the cell cytoplasm in plant cells. Actin and myosin are plentiful in muscle cells. When your actin and myosin filaments slide past each other, your muscles contract.

Microfilaments also provide some rigidity and shape to the cell. They can depolymerize (disassemble) and reform quickly, thus enabling a cell to change its shape and move. White blood cells (your body's infection-fighting cells) make good use of this ability. They can move to the site of an infection and phagocytize the pathogen.

Link to Learning

To see an example of a white blood cell in action, watch a short time-lapse video of the cell capturing two bacteria. It engulfs one and then moves on to the other.

Visit this page in your course online to use this activity.

Intermediate Filaments

Intermediate filaments are made of several strands of fibrous proteins that are wound together (Figure). These elements of the cytoskeleton get their name from the fact that their diameter, 8 to 10 nm, is between those of microfilaments and microtubules.

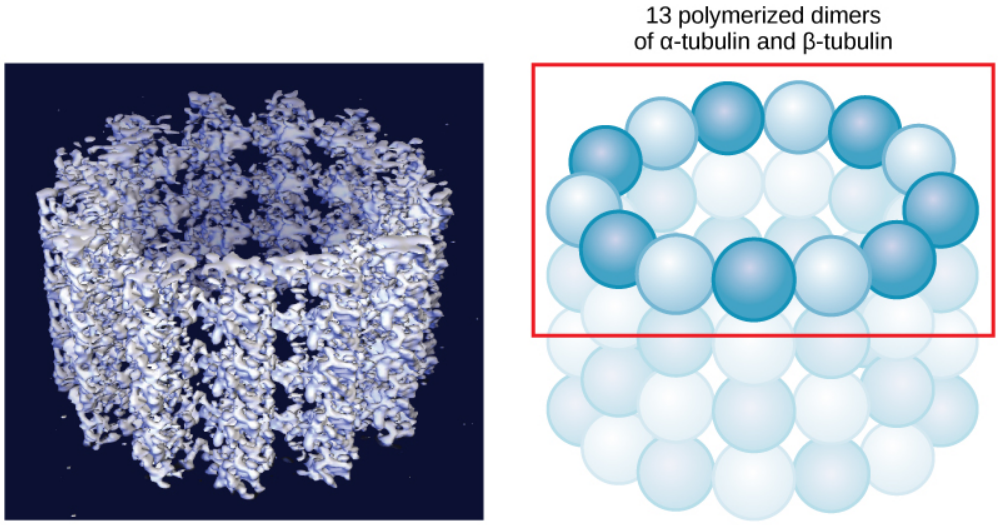
Intermediate filaments consist of several intertwined strands of fibrous proteins.

Intermediate filaments have no role in cell movement. Their function is purely structural. They bear tension, thus maintaining the shape of the cell, and anchor the nucleus and other organelles in place. Figure shows how intermediate filaments create a supportive scaffolding inside the cell.

The intermediate filaments are the most diverse group of cytoskeletal elements. Several types of fibrous proteins are found in the intermediate filaments. You are probably most familiar with keratin, the fibrous protein that strengthens your hair, nails, and the epidermis of the skin.

Microtubules

As their name implies, microtubules are small hollow tubes. The walls of the microtubule are made of polymerized dimers of α-tubulin and β-tubulin, two globular proteins (Figure). With a diameter of about 25 nm, microtubules are the widest components of the cytoskeleton. They help the cell resist compression, provide a track along which vesicles move through the cell, and pull replicated chromosomes to opposite ends of a dividing cell. Like microfilaments, microtubules can dissolve and reform quickly.



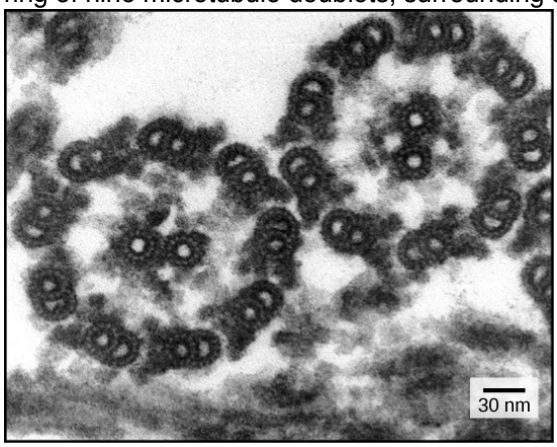
Microtubules are hollow. Their walls consist of 13 polymerized dimers of α-tubulin and β-tubulin (right image). The left image shows the molecular structure of the tube.

Microtubules are also the structural elements of flagella, cilia, and centrioles (the latter are the two perpendicular bodies of the centrosome). In fact, in animal cells, the centrosome is the microtubule-organizing center. In eukaryotic cells, flagella and cilia are quite different structurally from their counterparts in prokaryotes, as discussed below.

Flagella and Cilia

To refresh your memory, flagella (singular = flagellum) are long, hair-like structures that extend from the plasma membrane and are used to move an entire cell (for example, sperm, *Euglena*). When present, the cell has just one flagellum or a few flagella. Whencilia (singular = cilium) are present, however, many of them extend along the entire surface of the plasma membrane. They are short, hair-like structures that are used to move entire cells (such as paramecia) or substances along the outer surface of the cell (for example, the cilia of cells lining the Fallopian tubes that move the ovum toward the uterus, or cilia lining the cells of the respiratory tract that trap particulate matter and move it toward your nostrils.)

Despite their differences in length and number, flagella and cilia share a common structural arrangement of microtubules called a “9 + 2 array.” This is an appropriate name because a single flagellum or cilium is made of a ring of nine microtubule doublets, surrounding a single microtubule doublet in the center (Figure).



This transmission electron micrograph of two flagella shows the 9 + 2 array of microtubules: nine microtubule doublets surround a single microtubule doublet. (credit: modification of work by Dartmouth Electron Microscope Facility, Dartmouth College; scale-bar data from Matt Russell)

You have now completed a broad survey of the components of prokaryotic and eukaryotic cells. For a summary of cellular components in prokaryotic and eukaryotic cells, see Table.

Components of Prokaryotic and Eukaryotic Cells				
Cell Component	Function	Present in Prokaryotes?	Present in Animal Cells?	Present in Plant Cells?
Plasma membrane	Separates cell from external environment; controls passage of organic molecules, ions, water, oxygen, and wastes into and out of cell	Yes	Yes	Yes
Cytoplasm	Provides turgor pressure to plant cells as fluid inside the central vacuole; site of many metabolic reactions; medium in which organelles are found	Yes	Yes	Yes
Nucleolus	Darkened area within the nucleus where ribosomal subunits are synthesized.	No	Yes	Yes
Nucleus	Cell organelle that houses DNA and directs synthesis of ribosomes and proteins	No	Yes	Yes
Ribosomes	Protein synthesis	Yes	Yes	Yes
Mitochondria	ATP production/cellular respiration	No	Yes	Yes
Peroxisomes	Oxidizes and thus breaks down fatty acids and amino acids, and detoxifies poisons	No	Yes	Yes
Vesicles and vacuoles	Storage and transport; digestive function in plant cells	No	Yes	Yes
Centrosome	Unspecified role in cell division in animal cells; source of microtubules in animal cells	No	Yes	No
Lysosomes	Digestion of macromolecules; recycling of worn-out organelles	No	Yes	No
Cell wall	Protection, structural support and maintenance of cell shape	Yes, primarily peptidoglycan	No	Yes, primarily cellulose
Chloroplasts	Photosynthesis	No	No	Yes
Endoplasmic reticulum	Modifies proteins and synthesizes lipids	No	Yes	Yes
Golgi apparatus	Modifies, sorts, tags, packages, and distributes lipids and proteins	No	Yes	Yes
Cytoskeleton	Maintains cell's shape, secures organelles in specific positions, allows cytoplasm and vesicles to move within cell, and enables unicellular organisms to move independently	Yes	Yes	Yes

Components of Prokaryotic and Eukaryotic Cells				
Cell Component	Function	Present in Prokaryotes?	Present in Animal Cells?	Present in Plant Cells?
Flagella	Cellular locomotion	Some	Some	No, except for some plant sperm cells.
Cilia	Cellular locomotion, movement of particles along extracellular surface of plasma membrane, and filtration	Some	Some	No

Section Summary

The cytoskeleton has three different types of protein elements. From narrowest to widest, they are the microfilaments (actin filaments), intermediate filaments, and microtubules. Microfilaments are often associated with myosin. They provide rigidity and shape to the cell and facilitate cellular movements. Intermediate filaments bear tension and anchor the nucleus and other organelles in place. Microtubules help the cell resist compression, serve as tracks for motor proteins that move vesicles through the cell, and pull replicated chromosomes to opposite ends of a dividing cell. They are also the structural element of centrioles, flagella, and cilia.

Review Questions

Which of the following have the ability to disassemble and reform quickly?

1. microfilaments and intermediate filaments
2. microfilaments and microtubules
3. intermediate filaments and microtubules
4. only intermediate filaments

Which of the following do not play a role in intracellular movement?

1. microfilaments and intermediate filaments
2. microfilaments and microtubules
3. intermediate filaments and microtubules
4. only intermediate filaments

Free Response

What are the similarities and differences between the structures of centrioles and flagella?

How do cilia and flagella differ?

Glossary

cilium (plural = cilia) short, hair-like structure that extends from the plasma membrane in large numbers and is used to move an entire cell or move substances along the outer surface of the cell

cytoskeleton network of protein fibers that collectively maintain the shape of the cell, secure some organelles in specific positions, allow cytoplasm and vesicles to move within the cell, and enable unicellular organisms to move independently

flagellum (plural = flagella) long, hair-like structure that extends from the plasma membrane and is used to move the cell

intermediate filament cytoskeletal component, composed of several intertwined strands of fibrous protein, that bears tension, supports cell-cell junctions, and anchors cells to extracellular structures

microfilament narrowest element of the cytoskeleton system; it provides rigidity and shape to the cell and enables cellular movements

microtubule widest element of the cytoskeleton system; it helps the cell resist compression, provides a track along which vesicles move through the cell, pulls replicated chromosomes to opposite ends of a dividing cell, and is the structural element of centrioles, flagella, and cilia

- Downloads
- History
- Attribution
- More Information

Licensing & Attributions

CC licensed content, Shared previously

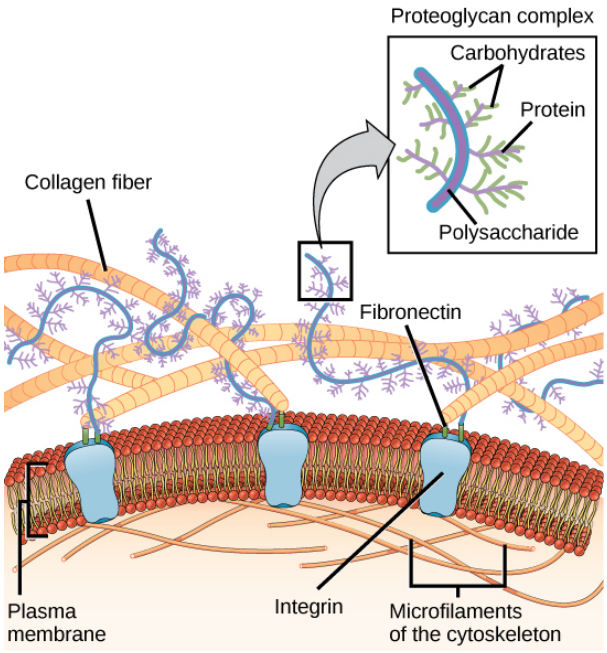
- the cytoskeleton. **Authored by**: open stax. **Provided by**: open stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:c20HSC1x@9/The-Cytoskeleton. **License**: *CC BY: Attribution*

CONNECTIONS BETWEEN CELLS AND CELLULAR ACTIVITIES

You already know that a group of similar cells working together is called a tissue. As you might expect, if cells are to work together, they must communicate with each other, just as you need to communicate with others if you work on a group project. Let's take a look at how cells communicate with each other.

Extracellular Matrix of Animal Cells

Most animal cells release materials into the extracellular space. The primary components of these materials are proteins, and the most abundant protein is collagen. Collagen fibers are interwoven with carbohydrate-containing protein molecules called proteoglycans. Collectively, these materials are called the extracellular matrix (Figure). Not only does the extracellular matrix hold the cells together to form a tissue, but it also allows the cells within the tissue to communicate with each other. How can this happen?



The extracellular matrix consists of a network of proteins and carbohydrates.

Cells have protein receptors on the extracellular surfaces of their plasma membranes. When a molecule within the matrix binds to the receptor, it changes the molecular structure of the receptor. The receptor, in turn, changes the conformation of the microfilaments positioned just inside the plasma membrane. These conformational changes induce chemical signals inside the cell that reach the nucleus and turn “on” or “off” the transcription of specific sections of DNA, which affects the production of associated proteins, thus changing the activities within the cell.

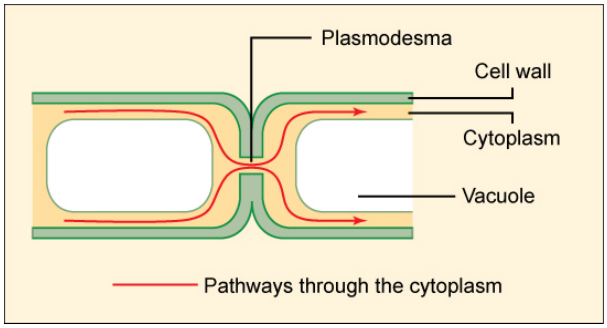
Blood clotting provides an example of the role of the extracellular matrix in cell communication. When the cells lining a blood vessel are damaged, they display a protein receptor called tissue factor. When tissue factor binds with another factor in the extracellular matrix, it causes platelets to adhere to the wall of the damaged blood vessel, stimulates the adjacent smooth muscle cells in the blood vessel to contract (thus constricting the blood vessel), and initiates a series of steps that stimulate the platelets to produce clotting factors.

Intercellular Junctions

Cells can also communicate with each other via direct contact, referred to as intercellular junctions. There are some differences in the ways that plant and animal cells do this. Plasmodesmata are junctions between plant cells, whereas animal cell contacts include tight junctions, gap junctions, and desmosomes.

Plasmodesmata

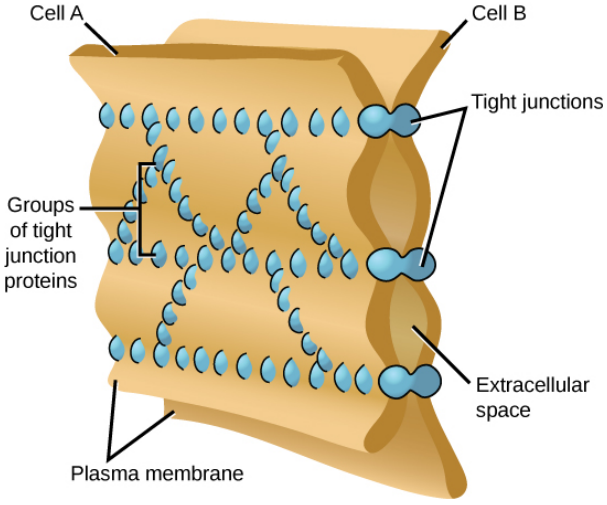
In general, long stretches of the plasma membranes of neighboring plant cells cannot touch one another because they are separated by the cell wall that surrounds each cell ([link]**b**). How then, can a plant transfer water and other soil nutrients from its roots, through its stems, and to its leaves? Such transport uses the vascular tissues (xylem and phloem) primarily. There also exist structural modifications called plasmodesmata (singular = plasmodesma), numerous channels that pass between cell walls of adjacent plant cells, connect their cytoplasm, and enable materials to be transported from cell to cell, and thus throughout the plant (Figure).



A plasmodesma is a channel between the cell walls of two adjacent plant cells. Plasmodesmata allow materials to pass from the cytoplasm of one plant cell to the cytoplasm of an adjacent cell.

Tight Junctions

A tight junction is a watertight seal between two adjacent animal cells (Figure). The cells are held tightly against each other by proteins (predominantly two proteins called claudins and occludins).

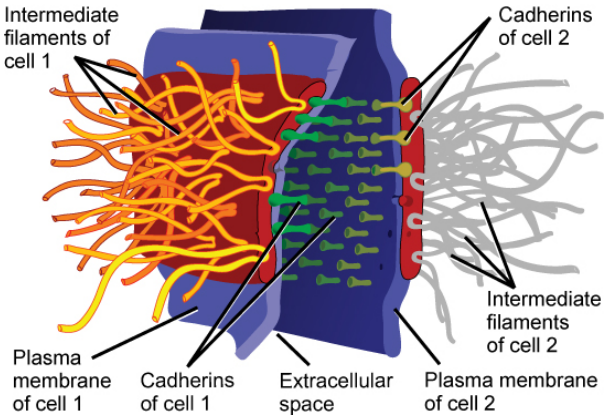


Tight junctions form watertight connections between adjacent animal cells. Proteins create tight junction adherence. (credit: modification of work by Mariana Ruiz Villareal)

This tight adherence prevents materials from leaking between the cells; tight junctions are typically found in epithelial tissues that line internal organs and cavities, and comprise most of the skin. For example, the tight junctions of the epithelial cells lining your urinary bladder prevent urine from leaking out into the extracellular space.

Desmosomes

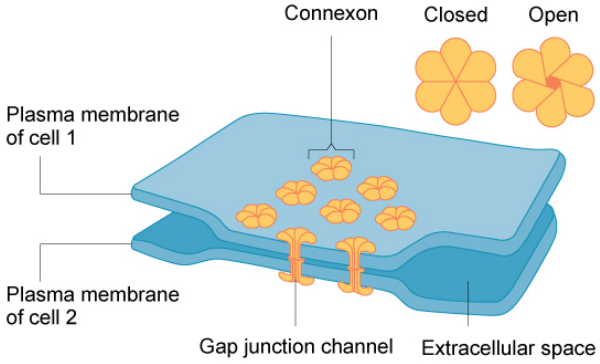
Also found only in animal cells are desmosomes, which act like spot welds between adjacent epithelial cells (Figure). Short proteins called cadherins in the plasma membrane connect to intermediate filaments to create desmosomes. The cadherins join two adjacent cells together and maintain the cells in a sheet-like formation in organs and tissues that stretch, like the skin, heart, and muscles.



A desmosome forms a very strong spot weld between cells. It is created by the linkage of cadherins and intermediate filaments. (credit: modification of work by Mariana Ruiz Villareal)

Gap Junctions

Gap junctions in animal cells are like plasmodesmata in plant cells in that they are channels between adjacent cells that allow for the transport of ions, nutrients, and other substances that enable cells to communicate (Figure). Structurally, however, gap junctions and plasmodesmata differ.



A gap junction is a protein-lined pore that allows water and small molecules to pass between adjacent animal cells. (credit: modification of work by Mariana Ruiz Villareal)

Gap junctions develop when a set of six proteins (called connexins) in the plasma membrane arrange themselves in an elongated donut-like configuration called a connexon. When the pores (“doughnut holes”) of connexons in adjacent animal cells align, a channel between the two cells forms. Gap junctions are particularly important in cardiac muscle: The electrical signal for the muscle to contract is passed efficiently through gap junctions, allowing the heart muscle cells to contract in tandem.

Link to Learning

To conduct a virtual microscopy lab and review the parts of a cell, work through the steps of thisinteractive assignment.

Section Summary

Animal cells communicate via their extracellular matrices and are connected to each other via tight junctions, desmosomes, and gap junctions. Plant cells are connected and communicate with each other via plasmodesmata.

When protein receptors on the surface of the plasma membrane of an animal cell bind to a substance in the extracellular matrix, a chain of reactions begins that changes activities taking place within the cell.

Plasmodesmata are channels between adjacent plant cells, while gap junctions are channels between adjacent animal cells. However, their structures are quite different. A tight junction is a watertight seal between two adjacent cells, while a desmosome acts like a spot weld.

Review Questions

Which of the following are found only in plant cells?

1. gap junctions
2. desmosomes
3. plasmodesmata
4. tight junctions

The key components of desmosomes are cadherins and _____.

1. actin
2. microfilaments
3. intermediate filaments
4. microtubules

Free Response

How does the structure of a plasmodesma differ from that of a gap junction?

Explain how the extracellular matrix functions.

Glossary

desmosome linkages between adjacent epithelial cells that form when cadherins in the plasma membrane attach to intermediate filaments

extracellular matrix material (primarily collagen, glycoproteins, and proteoglycans) secreted from animal cells that provides mechanical protection and anchoring for the cells in the tissue

gap junction channel between two adjacent animal cells that allows ions, nutrients, and low molecular weight substances to pass between cells, enabling the cells to communicate

plasmodesma (plural = plasmodesmata) channel that passes between the cell walls of adjacent plant cells, connects their cytoplasm, and allows materials to be transported from cell to cell

tight junction firm seal between two adjacent animal cells created by protein adherence

- Downloads
- History
- Attribution
- More Information

Licensing & Attributions

CC licensed content, Shared previously

- connections between cells and cellular activities. **Authored by**: open stax. **Provided by**: open stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:8Uypx7vu@7/Connections-between-Cells-and-. **License**: *CC BY: Attribution*

STUDY GUIDE: THE CELL

Study Questions

Objective: Describe the structure and function of a cell.

Use this page to check your understanding of the content.

Vocabulary- Know the function of these cell organelles and be able to state what types of cells these parts are found in.

1. Plasma membrane
2. Cytoplasm
3. Nucleus
4. Nucleoplasm
5. Nuclear membrane
6. Nuclear pores
7. Nucleolus
8. Ribosomes
9. Mitochondria
10. Cytoskeleton
11. Rough ER
12. Smooth ER
13. Golgi complex
14. Vesicles
15. Lysosomes
16. Cell wall
17. Chloroplast
18. Central vacuole

Study Guide Questions

1. Be able to describe characteristics that are shared by ALL CELLS. (In other words, what is The Cell Theory?)
2. Why are there upper limits on cell size? how can a cell increase its surface area?
3. Compare and contrast prokaryotes and eukaryotes.
4. Compare and contrast plant cells, fungal cells and animal cells.
5. Label the cell parts shown on this animal cell. Contrast this with what you'd expect to see on a plant cell.

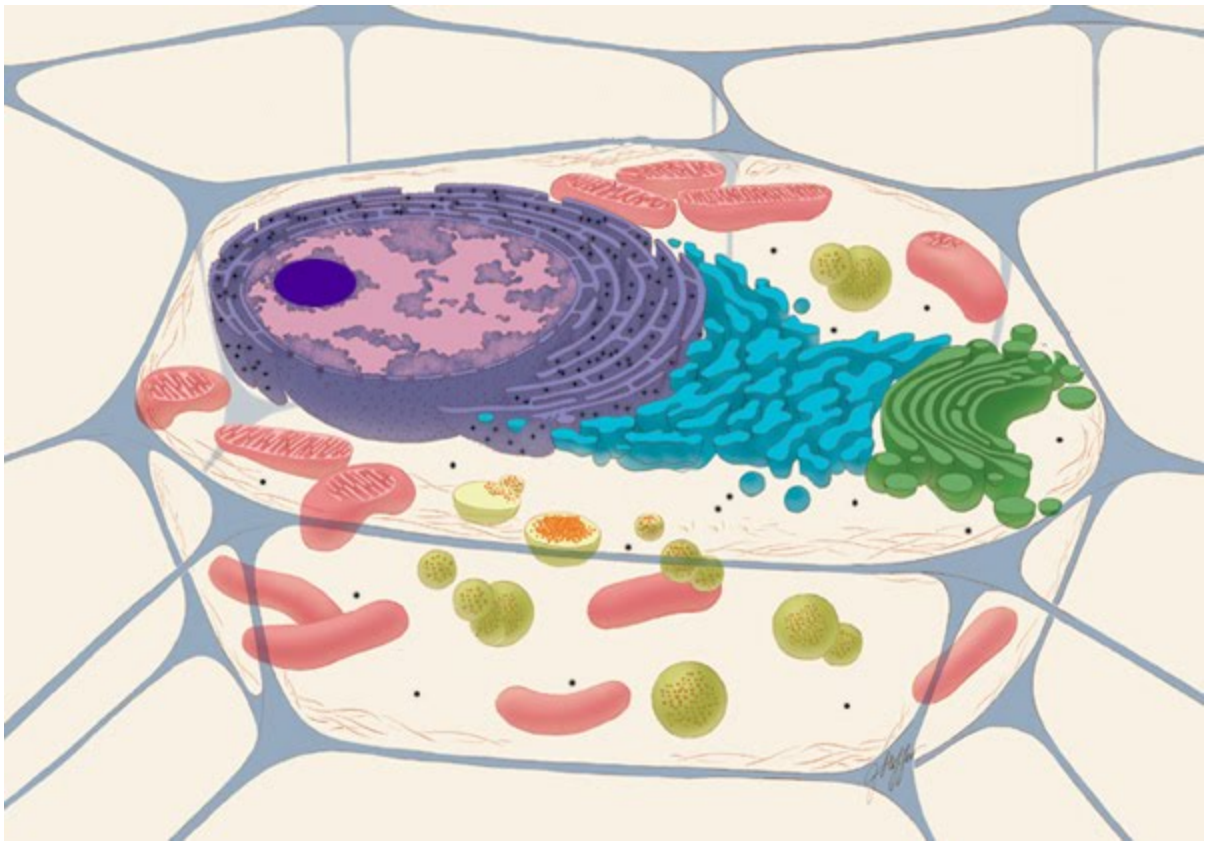


Image accessed from http://publications.nigms.nih.gov/insidethecell/chapter1.html (Links to an external site.)

Licensing & Attributions

CC licensed content, Shared previously

- Study Guide: The Cell . **Authored by**: Wendy Riggs. **Provided by**: College of the Redwoods. **Project**: Kaleidoscope. **License**: *CC BY: Attribution*

STRUCTURE AND FUNCTION OF PLASMA MEMBRANE

COMPONENTS AND STRUCTURE

5.1 Components and Structure

Summary

A cell's plasma membrane defines the cell, outlines its borders, and determines the nature of its interaction with its environment (seeTable for a summary). Cells exclude some substances, take in others, and excrete still others, all in controlled quantities. The plasma membrane must be very flexible to allow certain cells, such as red blood cells and white blood cells, to change shape as they pass through narrow capillaries. These are the more obvious functions of a plasma membrane. In addition, the surface of the plasma membrane carries markers that allow cells to recognize one another, which is vital for tissue and organ formation during early development, and which later plays a role in the "self" versus "non-self" distinction of the immune response.

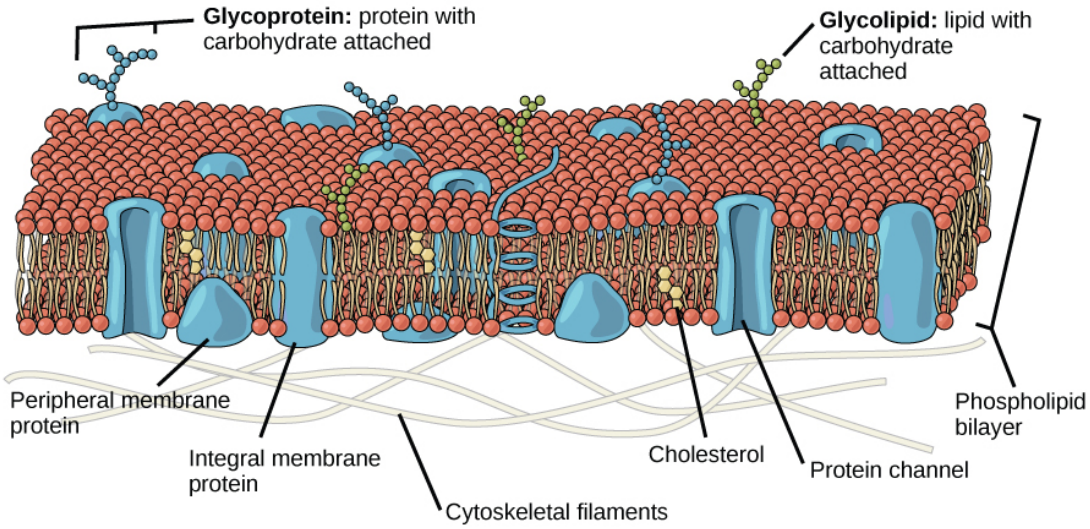
Among the most sophisticated functions of the plasma membrane is the ability to transmit signals by means of complex, integral proteins known as receptors. These proteins act both as receivers of extracellular inputs and as activators of intracellular processes. These membrane receptors provide extracellular attachment sites for effectors like hormones and growth factors, and they activate intracellular response cascades when their effectors are bound. Occasionally, receptors are hijacked by viruses (HIV, human immunodeficiency virus, is one example) that use them to gain entry into cells, and at times, the genes encoding receptors become mutated, causing the process of signal transduction to malfunction with disastrous consequences.

Fluid Mosaic Model

The existence of the plasma membrane was identified in the 1890s, and its chemical components were identified in 1915. The principal components identified at that time were lipids and proteins. The first widely accepted model of the plasma membrane's structure was proposed in 1935 by Hugh Davson and James Danielli; it was based on the "railroad track" appearance of the plasma membrane in early electron micrographs. They theorized that the structure of the plasma membrane resembles a sandwich, with protein being analogous to the bread, and lipids being analogous to the filling. In the 1950s, advances in microscopy, notably transmission electron microscopy (TEM), allowed researchers to see that the core of the plasma membrane consisted of a double, rather than a single, layer. A new model that better explains both the microscopic observations and the function of that plasma membrane was proposed by S.J. Singer and Garth L. Nicolson in 1972.

The explanation proposed by Singer and Nicolson is called the fluid mosaic model. The model has evolved somewhat over time, but it still best accounts for the structure and functions of the plasma membrane as we now understand them. The fluid mosaic model describes the structure of the plasma membrane as a mosaic of components—including phospholipids, cholesterol, proteins, and carbohydrates—that gives the membrane a fluid character. Plasma membranes range from 5 to 10 nm in thickness. For comparison, human red blood cells,

visible via light microscopy, are approximately 8 μm wide, or approximately 1,000 times wider than a plasma membrane. The membrane does look a bit like a sandwich (Figure).



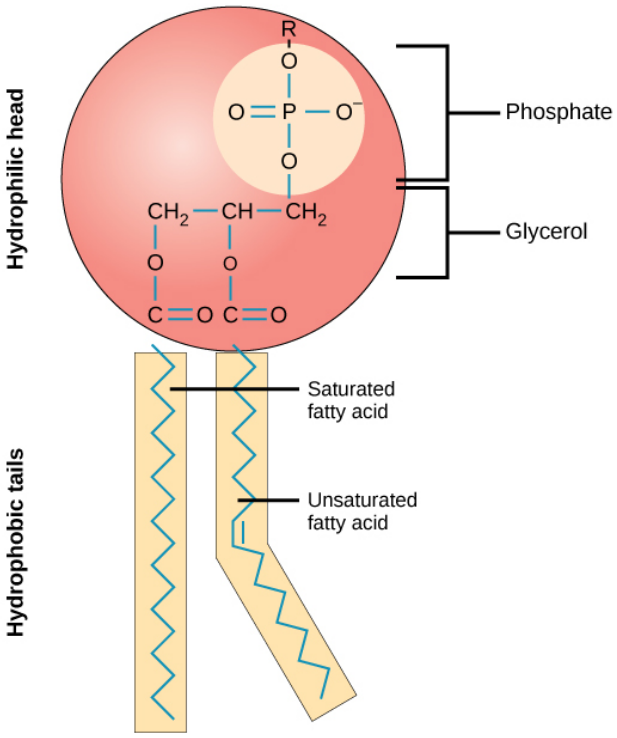
The fluid mosaic model of the plasma membrane describes the plasma membrane as a fluid combination of phospholipids, cholesterol, and proteins. Carbohydrates attached to lipids (glycolipids) and to proteins (glycoproteins) extend from the outward-facing surface of the membrane.

The principal components of a plasma membrane are lipids (phospholipids and cholesterol), proteins, and carbohydrates attached to some of the lipids and some of the proteins. A phospholipid is a molecule consisting of glycerol, two fatty acids, and a phosphate-linked head group. Cholesterol, another lipid composed of four fused carbon rings, is found alongside the phospholipids in the core of the membrane. The proportions of proteins, lipids, and carbohydrates in the plasma membrane vary with cell type, but for a typical human cell, protein accounts for about 50 percent of the composition by mass, lipids (of all types) account for about 40 percent of the composition by mass, with the remaining 10 percent of the composition by mass being carbohydrates. However, the concentration of proteins and lipids varies with different cell membranes. For example, myelin, an outgrowth of the membrane of specialized cells that insulates the axons of the peripheral nerves, contains only 18 percent protein and 76 percent lipid. The mitochondrial inner membrane contains 76 percent protein and only 24 percent lipid. The plasma membrane of human red blood cells is 30 percent lipid. Carbohydrates are present only on the exterior surface of the plasma membrane and are attached to proteins, forming glycoproteins, or attached to lipids, forming glycolipids.

Phospholipids

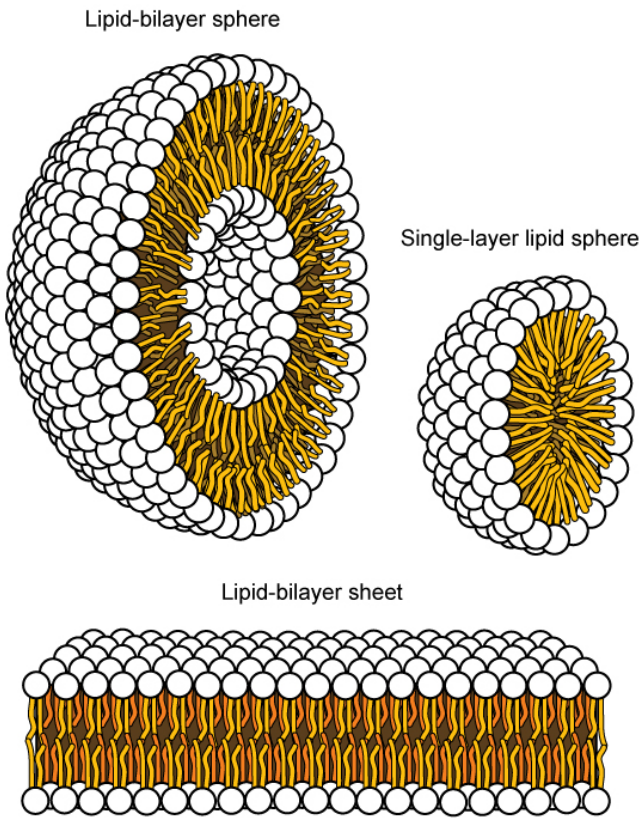
The main fabric of the membrane is composed of amphiphilic, phospholipid molecules. The hydrophilic or “water-loving” areas of these molecules (which look like a collection of balls in an artist’s rendition of the model) (Figure) are in contact with the aqueous fluid both inside and outside the cell. Hydrophobic, or water-hating molecules, tend to be non-polar. They interact with other non-polar molecules in chemical reactions, but generally do not interact with polar molecules. When placed in water, hydrophobic molecules tend to form a ball or cluster. The hydrophilic regions of the phospholipids tend to form hydrogen bonds with water and other polar molecules on both the exterior and interior of the cell. Thus, the membrane surfaces that face the interior and exterior of the cell are hydrophilic. In contrast, the interior of the cell membrane is hydrophobic and will not interact with water. Therefore, phospholipids form an excellent two-layer cell membrane that separates fluid within the cell from the fluid outside of the cell.

A phospholipid molecule (Figure) consists of a three-carbon glycerol backbone with two fatty acid molecules attached to carbons 1 and 2, and a phosphate-containing group attached to the third carbon. This arrangement gives the overall molecule an area described as its head (the phosphate-containing group), which has a polar character or negative charge, and an area called the tail (the fatty acids), which has no charge. The head can form hydrogen bonds, but the tail cannot. A molecule with this arrangement of a positively or negatively charged area and an uncharged, or non-polar, area is referred to as amphiphilic or “dual-loving.”



This phospholipid molecule is composed of a hydrophilic head and two hydrophobic tails. The hydrophilic head group consists of a phosphate-containing group attached to a glycerol molecule. The hydrophobic tails, each containing either a saturated or an unsaturated fatty acid, are long hydrocarbon chains.

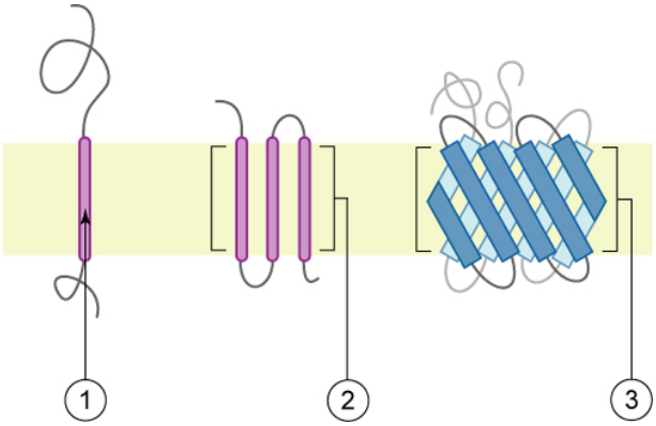
This characteristic is vital to the structure of a plasma membrane because, in water, phospholipids tend to become arranged with their hydrophobic tails facing each other and their hydrophilic heads facing out. In this way, they form a lipid bilayer—a barrier composed of a double layer of phospholipids that separates the water and other materials on one side of the barrier from the water and other materials on the other side. In fact, phospholipids heated in an aqueous solution tend to spontaneously form small spheres or droplets (called micelles or liposomes), with their hydrophilic heads forming the exterior and their hydrophobic tails on the inside (Figure).



In an aqueous solution, phospholipids tend to arrange themselves with their polar heads facing outward and their hydrophobic tails facing inward. (credit: modification of work by Mariana Ruiz Villareal)

Proteins

Proteins make up the second major component of plasma membranes. Integral proteins (some specialized types are called integrins) are, as their name suggests, integrated completely into the membrane structure, and their hydrophobic membrane-spanning regions interact with the hydrophobic region of the the phospholipid bilayer (Figure). Single-pass integral membrane proteins usually have a hydrophobic transmembrane segment that consists of 20–25 amino acids. Some span only part of the membrane—associating with a single layer—while others stretch from one side of the membrane to the other, and are exposed on either side. Some complex proteins are composed of up to 12 segments of a single protein, which are extensively folded and embedded in the membrane (Figure). This type of protein has a hydrophilic region or regions, and one or several mildly hydrophobic regions. This arrangement of regions of the protein tends to orient the protein alongside the phospholipids, with the hydrophobic region of the protein adjacent to the tails of the phospholipids and the hydrophilic region or regions of the protein protruding from the membrane and in contact with the cytosol or extracellular fluid.



Integral membranes proteins may have one or more alpha-helices that span the membrane (examples 1 and 2), or they may have beta-sheets that span the membrane (example 3). (credit: “Foobar”/Wikimedia Commons)

Peripheral proteins are found on the exterior and interior surfaces of membranes, attached either to integral proteins or to phospholipids. Peripheral proteins, along with integral proteins, may serve as enzymes, as structural attachments for the fibers of the cytoskeleton, or as part of the cell’s recognition sites. These are sometimes referred to as “cell-specific” proteins. The body recognizes its own proteins and attacks foreign proteins associated with invasive pathogens.

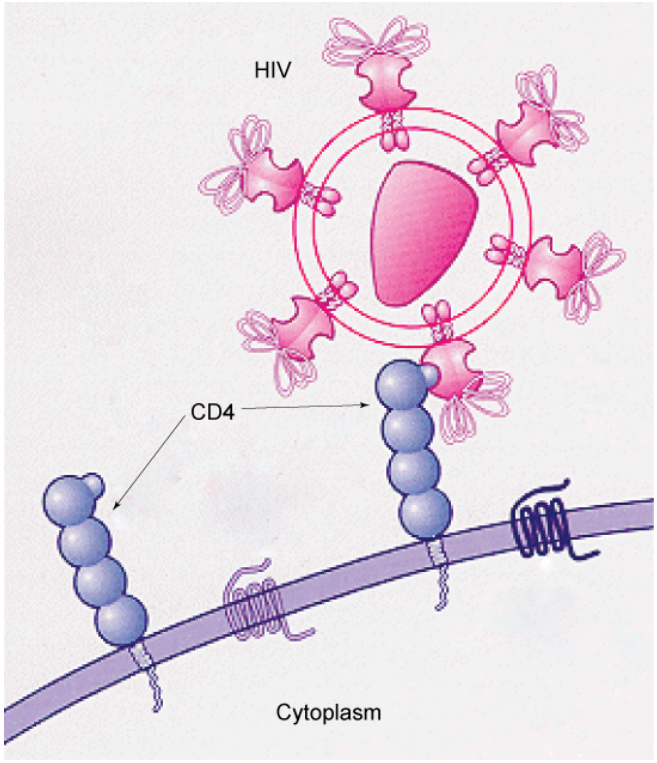
Carbohydrates

Carbohydrates are the third major component of plasma membranes. They are always found on the exterior surface of cells and are bound either to proteins (forming glycoproteins) or to lipids (forming glycolipids) (Figure). These carbohydrate chains may consist of 2–60 monosaccharide units and can be either straight or branched. Along with peripheral proteins, carbohydrates form specialized sites on the cell surface that allow cells to recognize each other. These sites have unique patterns that allow the cell to be recognized, much the way that the facial features unique to each person allow him or her to be recognized. This recognition function is very important to cells, as it allows the immune system to differentiate between body cells (called “self”) and foreign cells or tissues (called “non-self”). Similar types of glycoproteins and glycolipids are found on the surfaces of viruses and may change frequently, preventing immune cells from recognizing and attacking them.

These carbohydrates on the exterior surface of the cell—the carbohydrate components of both glycoproteins and glycolipids—are collectively referred to as the glycocalyx (meaning “sugar coating”). The glycocalyx is highly hydrophilic and attracts large amounts of water to the surface of the cell. This aids in the interaction of the cell with its watery environment and in the cell’s ability to obtain substances dissolved in the water. As discussed above, the glycocalyx is also important for cell identification, self/non-self determination, and embryonic development, and is used in cell-cell attachments to form tissues.
EVOLUTION CONNECTION

How Viruses Infect Specific OrgansGlycoprotein and glycolipid patterns on the surfaces of cells give many viruses an opportunity for infection. HIV and hepatitis viruses infect only specific organs or cells in the human body. HIV is able to penetrate the plasma membranes of a subtype of lymphocytes called T-helper cells, as well as some monocytes and central nervous system cells. The hepatitis virus attacks liver cells.

These viruses are able to invade these cells, because the cells have binding sites on their surfaces that are specific to and compatible with certain viruses (Figure). Other recognition sites on the virus’s surface interact with the human immune system, prompting the body to produce antibodies. Antibodies are made in response to the antigens or proteins associated with invasive pathogens, or in response to foreign cells, such as might occur with an organ transplant. These same sites serve as places for antibodies to attach and either destroy or inhibit the activity of the virus. Unfortunately, these recognition sites on HIV change at a rapid rate because of mutations, making the production of an effective vaccine against the virus very difficult, as the virus evolves and adapts. A person infected with HIV will quickly develop different populations, or variants, of the virus that are distinguished by differences in these recognition sites. This rapid change of surface markers decreases the effectiveness of the person’s immune system in attacking the virus, because the antibodies will not recognize the new variations of the surface patterns. In the case of HIV, the problem is compounded by the fact that the virus specifically infects and destroys cells involved in the immune response, further incapacitating the host.



HIV binds to the CD4 receptor, a glycoprotein on the surfaces of T cells. (credit: modification of work by NIH, NIAID)

Membrane Fluidity

The mosaic characteristic of the membrane, described in the fluid mosaic model, helps to illustrate its nature. The integral proteins and lipids exist in the membrane as separate but loosely attached molecules. These resemble the separate, multicolored tiles of a mosaic picture, and they float, moving somewhat with respect to one another. The membrane is not like a balloon, however, that can expand and contract; rather, it is fairly rigid and can burst if penetrated or if a cell takes in too much water. However, because of its mosaic nature, a very fine needle can easily penetrate a plasma membrane without causing it to burst, and the membrane will flow and self-seal when the needle is extracted.

The mosaic characteristics of the membrane explain some but not all of its fluidity. There are two other factors that help maintain this fluid characteristic. One factor is the nature of the phospholipids themselves. In their saturated form, the fatty acids in phospholipid tails are saturated with bound hydrogen atoms. There are no double bonds between adjacent carbon atoms. This results in tails that are relatively straight. In contrast, unsaturated fatty acids do not contain a maximal number of hydrogen atoms, but they do contain some double bonds between adjacent carbon atoms; a double bond results in a bend in the string of carbons of approximately 30 degrees (Figure).

Thus, if saturated fatty acids, with their straight tails, are compressed by decreasing temperatures, they press in on each other, making a dense and fairly rigid membrane. If unsaturated fatty acids are compressed, the “kinks” in their tails elbow adjacent phospholipid molecules away, maintaining some space between the phospholipid molecules. This “elbow room” helps to maintain fluidity in the membrane at temperatures at which membranes with saturated fatty acid tails in their phospholipids would “freeze” or solidify. The relative fluidity of the membrane is particularly important in a cold environment. A cold environment tends to compress membranes composed largely of saturated fatty acids, making them less fluid and more susceptible to rupturing. Many organisms (fish are one example) are capable of adapting to cold environments by changing the proportion of unsaturated fatty acids in their membranes in response to the lowering of the temperature.

Link to Learning

Visit this site to see animations of the fluidity and mosaic quality of membranes.

Animals have an additional membrane constituent that assists in maintaining fluidity. Cholesterol, which lies alongside the phospholipids in the membrane, tends to dampen the effects of temperature on the membrane. Thus, this lipid functions as a buffer, preventing lower temperatures from inhibiting fluidity and preventing increased temperatures from increasing fluidity too much. Thus, cholesterol extends, in both directions, the range of temperature in which the membrane is appropriately fluid and consequently functional. Cholesterol also serves other functions, such as organizing clusters of transmembrane proteins into lipid rafts.

The Components and Functions of the Plasma Membrane	
Component	Location
Phospholipid	Main fabric of the membrane
Cholesterol	Attached between phospholipids and between the two phospholipid layers
Integral proteins (for example, integrins)	Embedded within the phospholipid layer(s). May or may not penetrate through both layers
Peripheral proteins	On the inner or outer surface of the phospholipid bilayer; not embedded within the phospholipids
Carbohydrates (components of glycoproteins and glycolipids)	Generally attached to proteins on the outside membrane layer

CAREER CONNECTION

ImmunologistThe variations in peripheral proteins and carbohydrates that affect a cell's recognition sites are of prime interest in immunology. These changes are taken into consideration in vaccine development. Many infectious diseases, such as smallpox, polio, diphtheria, and tetanus, were conquered by the use of vaccines.

Immunologists are the physicians and scientists who research and develop vaccines, as well as treat and study allergies or other immune problems. Some immunologists study and treat autoimmune problems (diseases in which a person's immune system attacks his or her own cells or tissues, such as lupus) and immunodeficiencies, whether acquired (such as acquired immunodeficiency syndrome, or AIDS) or hereditary (such as severe combined immunodeficiency, or SCID). Immunologists are called in to help treat organ transplantation patients, who must have their immune systems suppressed so that their bodies will not reject a transplanted organ. Some immunologists work to understand natural immunity and the effects of a person's environment on it. Others work on questions about how the immune system affects diseases such as cancer. In the past, the importance of having a healthy immune system in preventing cancer was not at all understood.

To work as an immunologist, a PhD or MD is required. In addition, immunologists undertake at least 2–3 years of training in an accredited program and must pass an examination given by the American Board of Allergy and Immunology. Immunologists must possess knowledge of the functions of the human body as they relate to issues beyond immunization, and knowledge of pharmacology and medical technology, such as medications, therapies, test materials, and surgical procedures.

Section Summary

The modern understanding of the plasma membrane is referred to as the fluid mosaic model. The plasma membrane is composed of a bilayer of phospholipids, with their hydrophobic, fatty acid tails in contact with each other. The landscape of the membrane is studded with proteins, some of which span the membrane. Some of these proteins serve to transport materials into or out of the cell. Carbohydrates are attached to some of the

proteins and lipids on the outward-facing surface of the membrane, forming complexes that function to identify the cell to other cells. The fluid nature of the membrane is due to temperature, the configuration of the fatty acid tails (some kinked by double bonds), the presence of cholesterol embedded in the membrane, and the mosaic nature of the proteins and protein-carbohydrate combinations, which are not firmly fixed in place. Plasma membranes enclose and define the borders of cells, but rather than being a static bag, they are dynamic and constantly in flux.

Review Questions

Which plasma membrane component can be either found on its surface or embedded in the membrane structure?

1. protein
2. cholesterol
3. carbohydrate
4. phospholipid

Which characteristic of a phospholipid contributes to the fluidity of the membrane?

1. its head
2. cholesterol
3. a saturated fatty acid tail
4. double bonds in the fatty acid tail

What is the primary function of carbohydrates attached to the exterior of cell membranes?

1. identification of the cell
2. flexibility of the membrane
3. strengthening the membrane
4. channels through membrane

Free Response

Why is it advantageous for the cell membrane to be fluid in nature?

Why do phospholipids tend to spontaneously orient themselves into something resembling a membrane?

Glossary

amphiphilic molecule possessing a polar or charged area and a nonpolar or uncharged area capable of interacting with both hydrophilic and hydrophobic environments

fluid mosaic model describes the structure of the plasma membrane as a mosaic of components including phospholipids, cholesterol, proteins, glycoproteins, and glycolipids (sugar chains attached to proteins or lipids, respectively), resulting in a fluid character (fluidity)

glycolipid combination of carbohydrates and lipids

glycoprotein combination of carbohydrates and proteins

hydrophilic molecule with the ability to bond with water; “water-loving”

hydrophobic molecule that does not have the ability to bond with water; “water-hating”

integral protein protein integrated into the membrane structure that interacts extensively with the hydrocarbon chains of membrane lipids and often spans the membrane; these proteins can be removed only by the disruption of the membrane by detergents

peripheral protein protein found at the surface of a plasma membrane either on its exterior or interior side; these proteins can be removed (washed off of the membrane) by a high-salt wash

Licensing & Attributions

CC licensed content, Shared previously

- components and structure. **Authored by**: open stax. **Provided by**: open stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:QOGUelqL@9/Components-and-Structure. **License**: *CC BY: Attribution*

PASSIVE TRANSPORT

5.2 Passive Transport

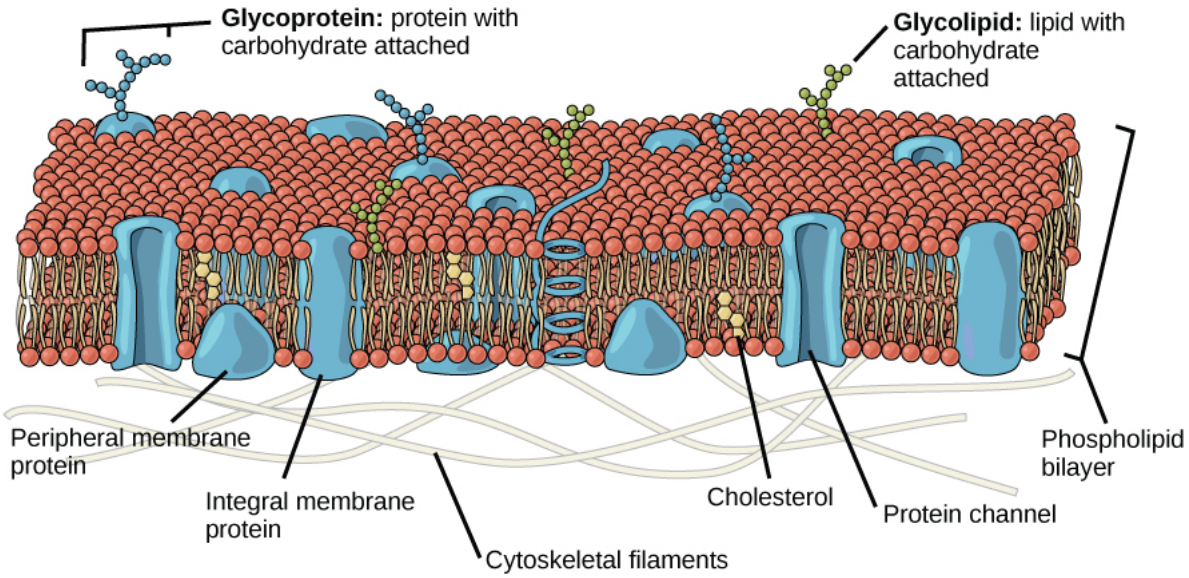
Summary

Plasma membranes must allow certain substances to enter and leave a cell, and prevent some harmful materials from entering and some essential materials from leaving. In other words, plasma membranes are selectively permeable—they allow some substances to pass through, but not others. If they were to lose this selectivity, the cell would no longer be able to sustain itself, and it would be destroyed. Some cells require larger amounts of specific substances than do other cells; they must have a way of obtaining these materials from extracellular fluids. This may happen passively, as certain materials move back and forth, or the cell may have special mechanisms that facilitate transport. Some materials are so important to a cell that it spends some of its energy, hydrolyzing adenosine triphosphate (ATP), to obtain these materials. Red blood cells use some of their energy doing just that. All cells spend the majority of their energy to maintain an imbalance of sodium and potassium ions between the interior and exterior of the cell.

The most direct forms of membrane transport are passive. Passive transport is a naturally occurring phenomenon and does not require the cell to exert any of its energy to accomplish the movement. In passive transport, substances move from an area of higher concentration to an area of lower concentration. A physical space in which there is a range of concentrations of a single substance is said to have a concentration gradient.

Selective Permeability

Plasma membranes are asymmetric: the interior of the membrane is not identical to the exterior of the membrane. In fact, there is a considerable difference between the array of phospholipids and proteins between the two leaflets that form a membrane. On the interior of the membrane, some proteins serve to anchor the membrane to fibers of the cytoskeleton. There are peripheral proteins on the exterior of the membrane that bind elements of the extracellular matrix. Carbohydrates, attached to lipids or proteins, are also found on the exterior surface of the plasma membrane. These carbohydrate complexes help the cell bind substances that the cell needs in the extracellular fluid. This adds considerably to the selective nature of plasma membranes (Figure).



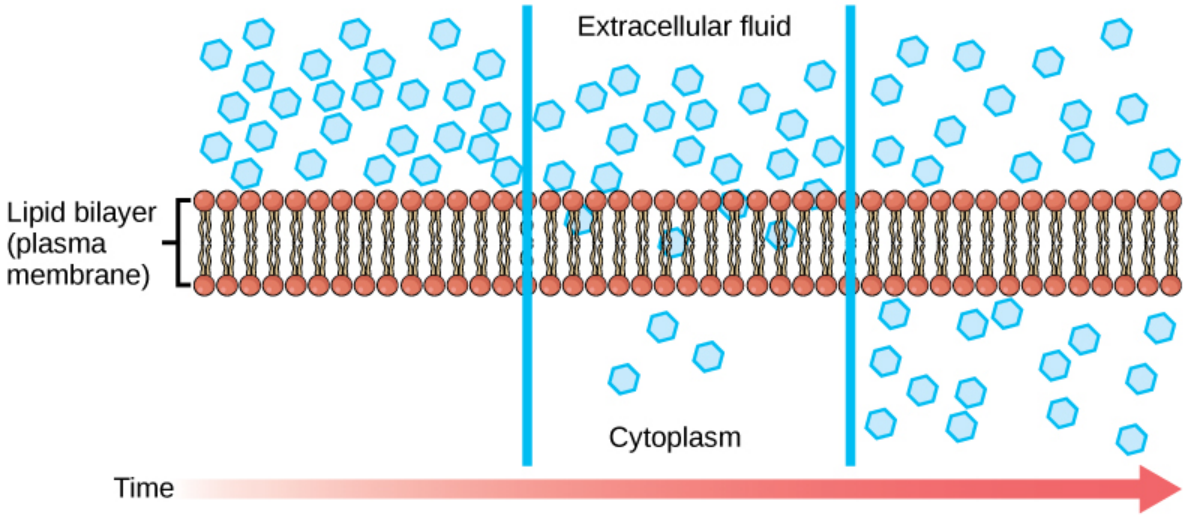
The exterior surface of the plasma membrane is not identical to the interior surface of the same membrane.

Recall that plasma membranes are amphiphilic: They have hydrophilic and hydrophobic regions. This characteristic helps the movement of some materials through the membrane and hinders the movement of others. Lipid-soluble material with a low molecular weight can easily slip through the hydrophobic lipid core of the membrane. Substances such as the fat-soluble vitamins A, D, E, and K readily pass through the plasma membranes in the digestive tract and other tissues. Fat-soluble drugs and hormones also gain easy entry into cells and are readily transported into the body's tissues and organs. Molecules of oxygen and carbon dioxide have no charge and so pass through membranes by simple diffusion.

Polar substances present problems for the membrane. While some polar molecules connect easily with the outside of a cell, they cannot readily pass through the lipid core of the plasma membrane. Additionally, while small ions could easily slip through the spaces in the mosaic of the membrane, their charge prevents them from doing so. Ions such as sodium, potassium, calcium, and chloride must have special means of penetrating plasma membranes. Simple sugars and amino acids also need help with transport across plasma membranes, achieved by various transmembrane proteins (channels).

Diffusion

Diffusion is a passive process of transport. A single substance tends to move from an area of high concentration to an area of low concentration until the concentration is equal across a space. You are familiar with diffusion of substances through the air. For example, think about someone opening a bottle of ammonia in a room filled with people. The ammonia gas is at its highest concentration in the bottle; its lowest concentration is at the edges of the room. The ammonia vapor will diffuse, or spread away, from the bottle, and gradually, more and more people will smell the ammonia as it spreads. Materials move within the cell's cytosol by diffusion, and certain materials move through the plasma membrane by diffusion (Figure). Diffusion expends no energy. On the contrary, concentration gradients are a form of potential energy, dissipated as the gradient is eliminated.



Diffusion through a permeable membrane moves a substance from an area of high concentration (extracellular fluid, in this case) down its concentration gradient (into the cytoplasm). (credit: modification of work by Mariana Ruiz Villareal)

Each separate substance in a medium, such as the extracellular fluid, has its own concentration gradient, independent of the concentration gradients of other materials. In addition, each substance will diffuse according to that gradient. Within a system, there will be different rates of diffusion of the different substances in the medium.

Factors That Affect Diffusion

Molecules move constantly in a random manner, at a rate that depends on their mass, their environment, and the amount of thermal energy they possess, which in turn is a function of temperature. This movement accounts for the diffusion of molecules through whatever medium in which they are localized. A substance will tend to move into any space available to it until it is evenly distributed throughout it. After a substance has diffused completely through a space, removing its concentration gradient, molecules will still move around in the space, but there will be no *net* movement of the number of molecules from one area to another. This lack of a concentration gradient in which there is no net movement of a substance is known as dynamic equilibrium. While diffusion will go forward in the presence of a concentration gradient of a substance, several factors affect the rate of diffusion.

- Extent of the concentration gradient: The greater the difference in concentration, the more rapid the diffusion. The closer the distribution of the material gets to equilibrium, the slower the rate of diffusion becomes.
- Mass of the molecules diffusing: Heavier molecules move more slowly; therefore, they diffuse more slowly. The reverse is true for lighter molecules.
- Temperature: Higher temperatures increase the energy and therefore the movement of the molecules, increasing the rate of diffusion. Lower temperatures decrease the energy of the molecules, thus decreasing the rate of diffusion.
- Solvent density: As the density of a solvent increases, the rate of diffusion decreases. The molecules slow down because they have a more difficult time getting through the denser medium. If the medium is less dense, diffusion increases. Because cells primarily use diffusion to move materials within the cytoplasm, any increase in the cytoplasm's density will inhibit the movement of the materials. An example of this is a person experiencing dehydration. As the body's cells lose water, the rate of diffusion decreases in the cytoplasm, and the cells' functions deteriorate. Neurons tend to be very sensitive to this effect. Dehydration frequently leads to unconsciousness and possibly coma because of the decrease in diffusion rate within the cells.
- Solubility: As discussed earlier, nonpolar or lipid-soluble materials pass through plasma membranes more easily than polar materials, allowing a faster rate of diffusion.
- Surface area and thickness of the plasma membrane: Increased surface area increases the rate of diffusion, whereas a thicker membrane reduces it.
- Distance travelled: The greater the distance that a substance must travel, the slower the rate of diffusion. This places an upper limitation on cell size. A large, spherical cell will die because nutrients or waste

cannot reach or leave the center of the cell, respectively. Therefore, cells must either be small in size, as in the case of many prokaryotes, or be flattened, as with many single-celled eukaryotes.

A variation of diffusion is the process of filtration. In filtration, material moves according to its concentration gradient through a membrane; sometimes the rate of diffusion is enhanced by pressure, causing the substances to filter more rapidly. This occurs in the kidney, where blood pressure forces large amounts of water and accompanying dissolved substances, or solutes, out of the blood and into the renal tubules. The rate of diffusion in this instance is almost totally dependent on pressure. One of the effects of high blood pressure is the appearance of protein in the urine, which is “squeezed through” by the abnormally high pressure.

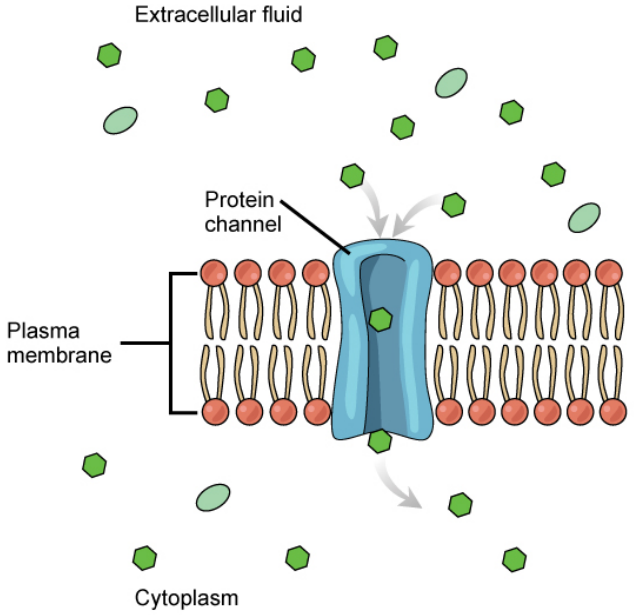
Facilitated transport

In facilitated transport, also called facilitated diffusion, materials diffuse across the plasma membrane with the help of membrane proteins. A concentration gradient exists that would allow these materials to diffuse into the cell without expending cellular energy. However, these materials are ions are polar molecules that are repelled by the hydrophobic parts of the cell membrane. Facilitated transport proteins shield these materials from the repulsive force of the membrane, allowing them to diffuse into the cell.

The material being transported is first attached to protein or glycoprotein receptors on the exterior surface of the plasma membrane. This allows the material that is needed by the cell to be removed from the extracellular fluid. The substances are then passed to specific integral proteins that facilitate their passage. Some of these integral proteins are collections of beta pleated sheets that form a pore or channel through the phospholipid bilayer. Others are carrier proteins which bind with the substance and aid its diffusion through the membrane.

Channels

The integral proteins involved in facilitated transport are collectively referred to as transport proteins, and they function as either channels for the material or carriers. In both cases, they are transmembrane proteins. Channels are specific for the substance that is being transported. Channel proteins have hydrophilic domains exposed to the intracellular and extracellular fluids; they additionally have a hydrophilic channel through their core that provides a hydrated opening through the membrane layers (Figure). Passage through the channel allows polar compounds to avoid the nonpolar central layer of the plasma membrane that would otherwise slow or prevent their entry into the cell. Aquaporins are channel proteins that allow water to pass through the membrane at a very high rate.

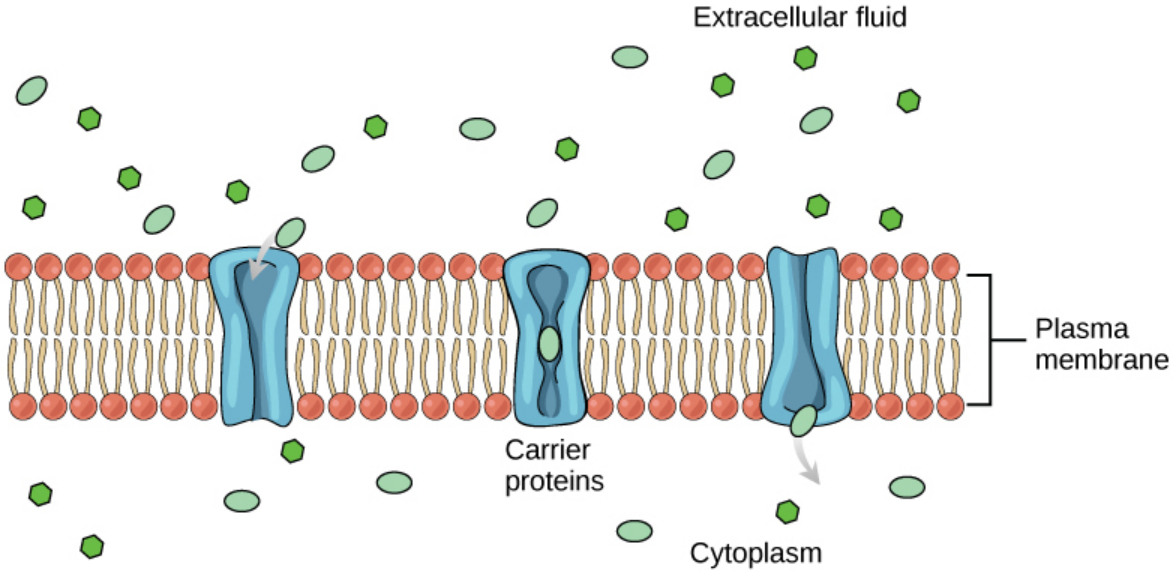


Facilitated transport moves substances down their concentration gradients. They may cross the plasma membrane with the aid of channel proteins. (credit: modification of work by Mariana Ruiz Villareal)

Channel proteins are either open at all times or they are “gated,” which controls the opening of the channel. The attachment of a particular ion to the channel protein may control the opening, or other mechanisms or substances may be involved. In some tissues, sodium and chloride ions pass freely through open channels, whereas in other tissues a gate must be opened to allow passage. An example of this occurs in the kidney, where both forms of channels are found in different parts of the renal tubules. Cells involved in the transmission of electrical impulses, such as nerve and muscle cells, have gated channels for sodium, potassium, and calcium in their membranes. Opening and closing of these channels changes the relative concentrations on opposing sides of the membrane of these ions, resulting in the facilitation of electrical transmission along membranes (in the case of nerve cells) or in muscle contraction (in the case of muscle cells).

Carrier Proteins

Another type of protein embedded in the plasma membrane is a carrier protein. This aptly named protein binds a substance and, in doing so, triggers a change of its own shape, moving the bound molecule from the outside of the cell to its interior (Figure); depending on the gradient, the material may move in the opposite direction. Carrier proteins are typically specific for a single substance. This selectivity adds to the overall selectivity of the plasma membrane. The exact mechanism for the change of shape is poorly understood. Proteins can change shape when their hydrogen bonds are affected, but this may not fully explain this mechanism. Each carrier protein is specific to one substance, and there are a finite number of these proteins in any membrane. This can cause problems in transporting enough of the material for the cell to function properly. When all of the proteins are bound to their ligands, they are saturated and the rate of transport is at its maximum. Increasing the concentration gradient at this point will not result in an increased rate of transport.



Some substances are able to move down their concentration gradient across the plasma membrane with the aid of carrier proteins. Carrier proteins change shape as they move molecules across the membrane. (credit: modification of work by Mariana Ruiz Villareal)

An example of this process occurs in the kidney. Glucose, water, salts, ions, and amino acids needed by the body are filtered in one part of the kidney. This filtrate, which includes glucose, is then reabsorbed in another part of the kidney. Because there are only a finite number of carrier proteins for glucose, if more glucose is present than the proteins can handle, the excess is not transported and it is excreted from the body in the urine. In a diabetic individual, this is described as “spilling glucose into the urine.” A different group of carrier proteins called glucose transport proteins, or GLUTs, are involved in transporting glucose and other hexose sugars through plasma membranes within the body.

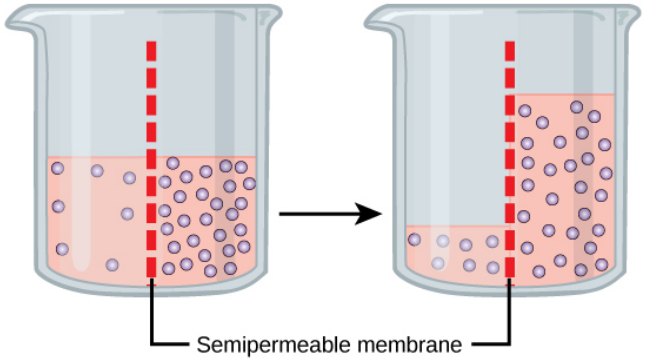
Channel and carrier proteins transport material at different rates. Channel proteins transport much more quickly than do carrier proteins. Channel proteins facilitate diffusion at a rate of tens of millions of molecules per second, whereas carrier proteins work at a rate of a thousand to a million molecules per second.

Osmosis

Osmosis is the movement of water through a semipermeable membrane according to the concentration gradient of water across the membrane, which is inversely proportional to the concentration of solutes. While diffusion transports material across membranes and within cells, osmosis transports *only water* across a membrane and the membrane limits the diffusion of solutes in the water. Not surprisingly, the aquaporins that facilitate water movement play a large role in osmosis, most prominently in red blood cells and the membranes of kidney tubules.

Mechanism

Osmosis is a special case of diffusion. Water, like other substances, moves from an area of high concentration to one of low concentration. An obvious question is what makes water move at all? Imagine a beaker with a semipermeable membrane separating the two sides or halves (Figure). On both sides of the membrane the water level is the same, but there are different concentrations of a dissolved substance, or solute, that cannot cross the membrane (otherwise the concentrations on each side would be balanced by the solute crossing the membrane). If the volume of the solution on both sides of the membrane is the same, but the concentrations of solute are different, then there are different amounts of water, the solvent, on either side of the membrane.



In osmosis, water always moves from an area of higher water concentration to one of lower concentration. In the diagram shown, the solute cannot pass through the selectively permeable membrane, but the water can.

To illustrate this, imagine two full glasses of water. One has a single teaspoon of sugar in it, whereas the second one contains one-quarter cup of sugar. If the total volume of the solutions in both cups is the same, which cup contains more water? Because the large amount of sugar in the second cup takes up much more space than the teaspoon of sugar in the first cup, the first cup has more water in it.

Returning to the beaker example, recall that it has a mixture of solutes on either side of the membrane. A principle of diffusion is that the molecules move around and will spread evenly throughout the medium if they can. However, only the material capable of getting through the membrane will diffuse through it. In this example, the solute cannot diffuse through the membrane, but the water can. Water has a concentration gradient in this system. Thus, water will diffuse down its concentration gradient, crossing the membrane to the side where it is less concentrated. This diffusion of water through the membrane—osmosis—will continue until the concentration gradient of water goes to zero or until the hydrostatic pressure of the water balances the osmotic pressure. Osmosis proceeds constantly in living systems.

Tonicity

Tonicity describes how an extracellular solution can change the volume of a cell by affecting osmosis. A solution's tonicity often directly correlates with the osmolarity of the solution. Osmolarity describes the total solute concentration of the solution. A solution with low osmolarity has a greater number of water molecules relative to the number of solute particles; a solution with high osmolarity has fewer water molecules with respect to solute particles. In a situation in which solutions of two different osmolarities are separated by a membrane permeable to water, though not to the solute, water will move from the side of the membrane with lower osmolarity (and more water) to the side with higher osmolarity (and less water). This effect makes sense if you remember that the solute cannot move across the membrane, and thus the only component in the system that can move—the

water—moves along its own concentration gradient. An important distinction that concerns living systems is that osmolarity measures the number of particles (which may be molecules) in a solution. Therefore, a solution that is cloudy with cells may have a lower osmolarity than a solution that is clear, if the second solution contains more dissolved molecules than there are cells.

Hypotonic Solutions

Three terms—hypotonic, isotonic, and hypertonic—are used to relate the osmolarity of a cell to the osmolarity of the extracellular fluid that contains the cells. In a hypotonic situation, the extracellular fluid has lower osmolarity than the fluid inside the cell, and water enters the cell. (In living systems, the point of reference is always the cytoplasm, so the prefix *hypo–* means that the extracellular fluid has a lower concentration of solutes, or a lower osmolarity, than the cell cytoplasm.) It also means that the extracellular fluid has a higher concentration of water in the solution than does the cell. In this situation, water will follow its concentration gradient and enter the cell.

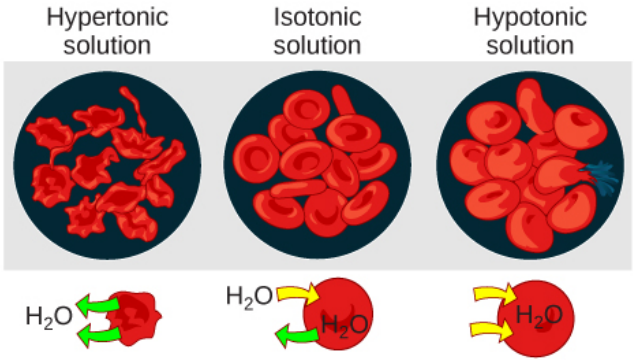
Hypertonic Solutions

As for a hypertonic solution, the prefix *hyper–* refers to the extracellular fluid having a higher osmolarity than the cell's cytoplasm; therefore, the fluid contains less water than the cell does. Because the cell has a relatively higher concentration of water, water will leave the cell.

Isotonic Solutions

In an isotonic solution, the extracellular fluid has the same osmolarity as the cell. If the osmolarity of the cell matches that of the extracellular fluid, there will be no net movement of water into or out of the cell, although water will still move in and out. Blood cells and plant cells in hypertonic, isotonic, and hypotonic solutions take on characteristic appearances (Figure).

ART CONNECTION



Osmotic pressure changes the shape of red blood cells in hypertonic, isotonic, and hypotonic solutions. (credit: Mariana Ruiz Villareal)

A doctor injects a patient with what the doctor thinks is an isotonic saline solution. The patient dies, and an autopsy reveals that many red blood cells have been destroyed. Do you think the solution the doctor injected was really isotonic?

Link to Learning

For a video illustrating the process of diffusion in solutions, visit this site.

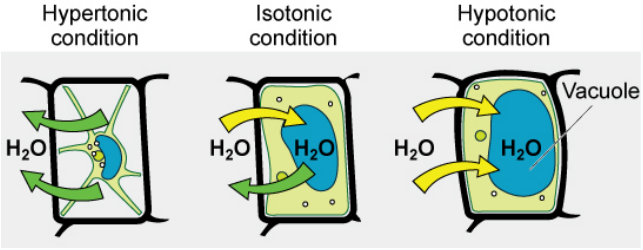
Tonicity in Living Systems

In a hypotonic environment, water enters a cell, and the cell swells. In an isotonic condition, the relative concentrations of solute and solvent are equal on both sides of the membrane. There is no net water movement; therefore, there is no change in the size of the cell. In a hypertonic solution, water leaves a cell and the cell shrinks. If either the hypo- or hyper- condition goes to excess, the cell's functions become compromised, and the cell may be destroyed.

A red blood cell will burst, or lyse, when it swells beyond the plasma membrane's capability to expand. Remember, the membrane resembles a mosaic, with discrete spaces between the molecules composing it. If the cell swells, and the spaces between the lipids and proteins become too large, the cell will break apart.

In contrast, when excessive amounts of water leave a red blood cell, the cell shrinks, or crenates. This has the effect of concentrating the solutes left in the cell, making the cytosol denser and interfering with diffusion within the cell. The cell's ability to function will be compromised and may also result in the death of the cell.

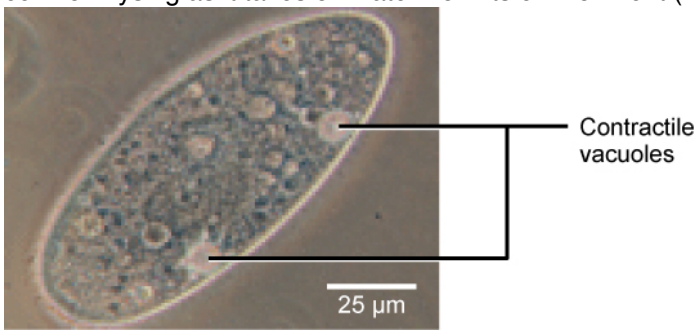
Various living things have ways of controlling the effects of osmosis—a mechanism called osmoregulation. Some organisms, such as plants, fungi, bacteria, and some protists, have cell walls that surround the plasma membrane and prevent cell lysis in a hypotonic solution. The plasma membrane can only expand to the limit of the cell wall, so the cell will not lyse. In fact, the cytoplasm in plants is always slightly hypertonic to the cellular environment, and water will always enter a cell if water is available. This inflow of water produces turgor pressure, which stiffens the cell walls of the plant (Figure). In nonwoody plants, turgor pressure supports the plant. Conversly, if the plant is not watered, the extracellular fluid will become hypertonic, causing water to leave the cell. In this condition, the cell does not shrink because the cell wall is not flexible. However, the cell membrane detaches from the wall and constricts the cytoplasm. This is called plasmolysis. Plants lose turgor pressure in this condition and wilt (Figure).



The turgor pressure within a plant cell depends on the tonicity of the solution that it is bathed in. (credit: modification of work by Mariana Ruiz Villareal)

Without adequate water, the plant on the left has lost turgor pressure, visible in its wilting; the turgor pressure is restored by watering it (right). (credit: Victor M. Vicente Selvas)

Tonicity is a concern for all living things. For example, paramecia and amoebas, which are protists that lack cell walls, have contractile vacuoles. This vesicle collects excess water from the cell and pumps it out, keeping the cell from lysing as it takes on water from its environment (Figure).



A paramecium's contractile vacuole, here visualized using bright field light microscopy at 480x magnification, continuously pumps water out of the organism's body to keep it from bursting in a hypotonic medium. (credit: modification of work by NIH; scale-bar data from Matt Russell)

Many marine invertebrates have internal salt levels matched to their environments, making them isotonic with the water in which they live. Fish, however, must spend approximately five percent of their metabolic energy maintaining osmotic homeostasis. Freshwater fish live in an environment that is hypotonic to their cells. These fish actively take in salt through their gills and excrete diluted urine to rid themselves of excess water. Saltwater fish live in the reverse environment, which is hypertonic to their cells, and they secrete salt through their gills and excrete highly concentrated urine.

In vertebrates, the kidneys regulate the amount of water in the body. Osmoreceptors are specialized cells in the brain that monitor the concentration of solutes in the blood. If the levels of solutes increase beyond a certain range, a hormone is released that retards water loss through the kidney and dilutes the blood to safer levels. Animals also have high concentrations of albumin, which is produced by the liver, in their blood. This protein is too large to pass easily through plasma membranes and is a major factor in controlling the osmotic pressures applied to tissues.

Section Summary

The passive forms of transport, diffusion and osmosis, move materials of small molecular weight across membranes. Substances diffuse from areas of high concentration to areas of lower concentration, and this process continues until the substance is evenly distributed in a system. In solutions containing more than one substance, each type of molecule diffuses according to its own concentration gradient, independent of the diffusion of other substances. Many factors can affect the rate of diffusion, including concentration gradient, size of the particles that are diffusing, temperature of the system, and so on.

In living systems, diffusion of substances into and out of cells is mediated by the plasma membrane. Some materials diffuse readily through the membrane, but others are hindered, and their passage is made possible by specialized proteins, such as channels and transporters. The chemistry of living things occurs in aqueous solutions, and balancing the concentrations of those solutions is an ongoing problem. In living systems, diffusion of some substances would be slow or difficult without membrane proteins that facilitate transport.

Art Connections

Figure A doctor injects a patient with what the doctor thinks is an isotonic saline solution. The patient dies, and an autopsy reveals that many red blood cells have been destroyed. Do you think the solution the doctor injected was really isotonic?

Review Questions

Water moves via osmosis _____.

1. throughout the cytoplasm
2. from an area with a high concentration of other solutes to a lower one
3. from an area with a high concentration of water to one of lower concentration
4. from an area with a low concentration of water to one of higher concentration

The principal force driving movement in diffusion is the _____.

1. temperature
2. particle size
3. concentration gradient
4. membrane surface area

What problem is faced by organisms that live in fresh water?

1. Their bodies tend to take in too much water.
2. They have no way of controlling their tonicity.
3. Only salt water poses problems for animals that live in it.
4. Their bodies tend to lose too much water to their environment.

Free Response

Discuss why the following affect the rate of diffusion: molecular size, temperature, solution density, and the distance that must be traveled.

Why does water move through a membrane?

Both of the regular intravenous solutions administered in medicine, normal saline and lactated Ringer's solution, are isotonic. Why is this important?

Glossary

aquaporin channel protein that allows water through the membrane at a very high rate

carrier protein membrane protein that moves a substance across the plasma membrane by changing its own shape

channel protein membrane protein that allows a substance to pass through its hollow core across the plasma membrane

concentration gradient area of high concentration adjacent to an area of low concentration

diffusion passive process of transport of low-molecular weight material according to its concentration gradient

facilitated transport process by which material moves down a concentration gradient (from high to low concentration) using integral membrane proteins

hypertonic situation in which extracellular fluid has a higher osmolarity than the fluid inside the cell, resulting in water moving out of the cell

hypotonic situation in which extracellular fluid has a lower osmolarity than the fluid inside the cell, resulting in water moving into the cell

isotonic situation in which the extracellular fluid has the same osmolarity as the fluid inside the cell, resulting in no net movement of water into or out of the cell

osmolarity total amount of substances dissolved in a specific amount of solution

osmosis transport of water through a semipermeable membrane according to the concentration gradient of water across the membrane that results from the presence of solute that cannot pass through the membrane

passive transport method of transporting material through a membrane that does not require energy

plasmolysis detaching of the cell membrane from the cell wall and constriction of the cell membrane when a plant cell is in a hypertonic solution

selectively permeable characteristic of a membrane that allows some substances through but not others

solute substance dissolved in a liquid to form a solution

tonicity amount of solute in a solution

transport protein membrane protein that facilitates passage of a substance across a membrane by binding it

Licensing & Attributions

CC licensed content, Shared previously

- passive transport. **Authored by**: open stax. **Provided by**: open stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:6X9o-0n6@12/Passive-Transport. **License**: *CC BY: Attribution*

ACTIVE TRANSPORT

5.3 Active Transport

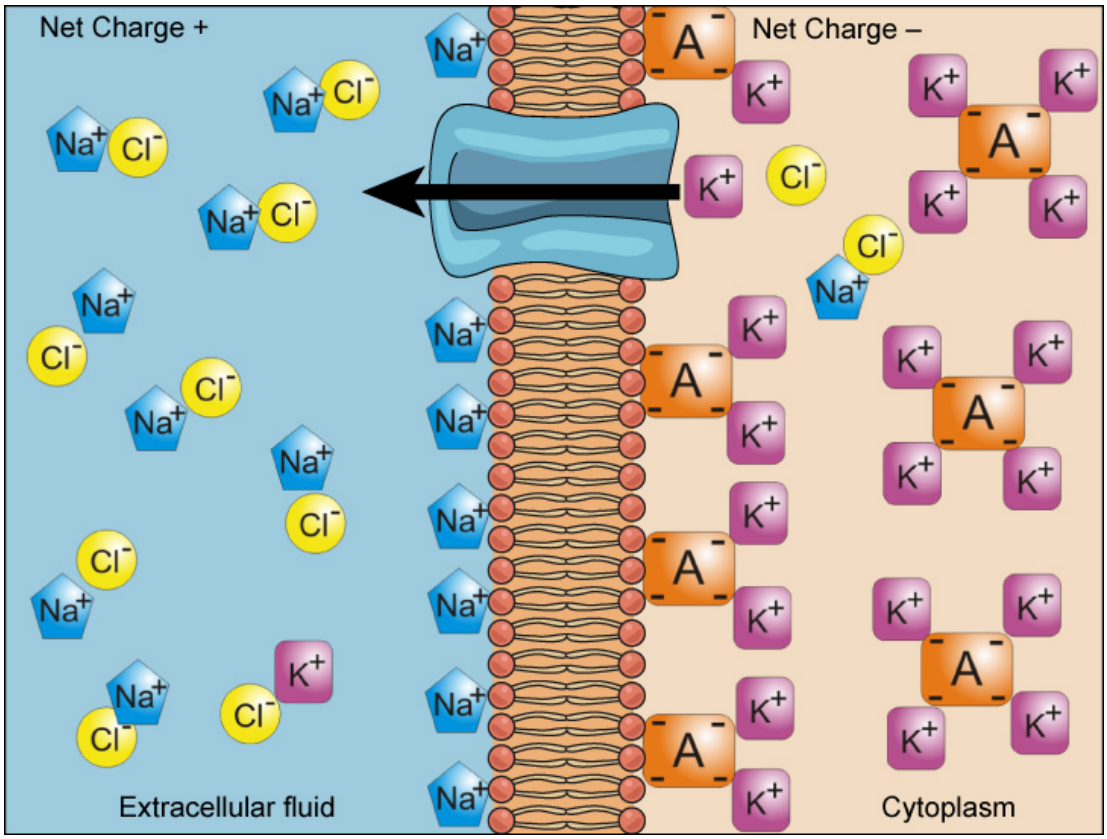
Summary

Active transport mechanisms require the use of the cell's energy, usually in the form of adenosine triphosphate (ATP). If a substance must move into the cell against its concentration gradient—that is, if the concentration of the substance inside the cell is greater than its concentration in the extracellular fluid (and vice versa)—the cell must use energy to move the substance. Some active transport mechanisms move small-molecular weight materials, such as ions, through the membrane. Other mechanisms transport much larger molecules.

Electrochemical Gradient

We have discussed simple concentration gradients—differential concentrations of a substance across a space or a membrane—but in living systems, gradients are more complex. Because ions move into and out of cells and because cells contain proteins that do not move across the membrane and are mostly negatively charged, there is also an electrical gradient, a difference of charge, across the plasma membrane. The interior of living cells is electrically negative with respect to the extracellular fluid in which they are bathed, and at the same time, cells have higher concentrations of potassium (K^+) and lower concentrations of sodium (Na^+) than does the extracellular fluid. So in a living cell, the concentration gradient of Na^+ tends to drive it into the cell, and the electrical gradient of Na^+ (a positive ion) also tends to drive it inward to the negatively charged interior. The situation is more complex, however, for other elements such as potassium. The electrical gradient of K^+, a positive ion, also tends to drive it into the cell, but the concentration gradient of K^+ tends to drive K^+ *out* of the cell (Figure). The combined gradient of concentration and electrical charge that affects an ion is called its electrochemical gradient.

ART CONNECTION



Electrochemical gradients arise from the combined effects of concentration gradients and electrical gradients. (credit: “Synaptitude”/Wikimedia Commons)

Injection of a potassium solution into a person’s blood is lethal; this is used in capital punishment and euthanasia. Why do you think a potassium solution injection is lethal?

Moving Against a Gradient

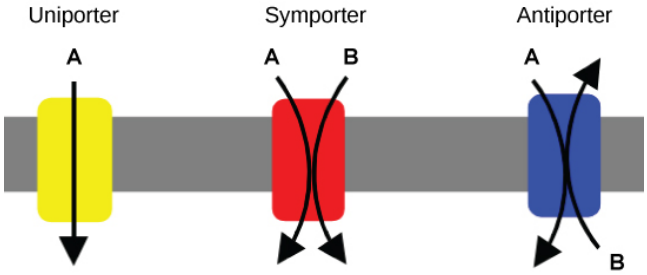
To move substances against a concentration or electrochemical gradient, the cell must use energy. This energy is harvested from ATP generated through the cell’s metabolism. Active transport mechanisms, collectively called pumps, work against electrochemical gradients. Small substances constantly pass through plasma membranes. Active transport maintains concentrations of ions and other substances needed by living cells in the face of these passive movements. Much of a cell’s supply of metabolic energy may be spent maintaining these processes. (Most of a red blood cell’s metabolic energy is used to maintain the imbalance between exterior and interior sodium and potassium levels required by the cell.) Because active transport mechanisms depend on a cell’s metabolism for energy, they are sensitive to many metabolic poisons that interfere with the supply of ATP.

Two mechanisms exist for the transport of small-molecular weight material and small molecules. Primary active transport moves ions across a membrane and creates a difference in charge across that membrane, which is directly dependent on ATP. Secondary active transport describes the movement of material that is due to the electrochemical gradient established by primary active transport that does not directly require ATP.

Carrier Proteins for Active Transport

An important membrane adaption for active transport is the presence of specific carrier proteins or pumps to facilitate movement: there are three types of these proteins or transporters (Figure). A uniporter carries one specific ion or molecule. A symportercarries two different ions or molecules, both in the same direction. An antiporter also carries two different ions or molecules, but in different directions. All of these transporters can also transport small, uncharged organic molecules like glucose. These three types of carrier proteins are also found in facilitated diffusion, but they do not require ATP to work in that process. Some examples of pumps for active

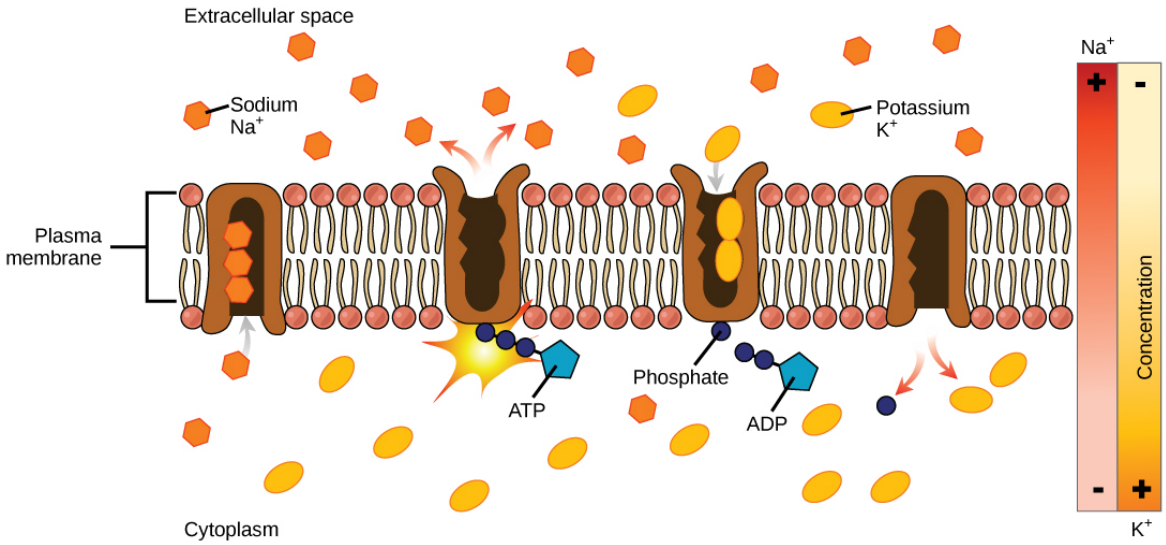
transport are Na^+-K^+ ATPase, which carries sodium and potassium ions, and H^+-K^+ ATPase, which carries hydrogen and potassium ions. Both of these are antiporter carrier proteins. Two other carrier proteins are Ca^{2+} ATPase and H^+ ATPase, which carry only calcium and only hydrogen ions, respectively. Both are pumps.



A uniporter carries one molecule or ion. A symporter carries two different molecules or ions, both in the same direction. An antiporter also carries two different molecules or ions, but in different directions. (credit: modification of work by “Lupask”/Wikimedia Commons)

Primary Active Transport

The primary active transport that functions with the active transport of sodium and potassium allows secondary active transport to occur. The second transport method is still considered active because it depends on the use of energy as does primary transport (Figure).



Primary active transport moves ions across a membrane, creating an electrochemical gradient (electrogenic transport). (credit: modification of work by Mariana Ruiz Villareal)

One of the most important pumps in animals cells is the sodium-potassium pump (Na^+-K^+ ATPase), which maintains the electrochemical gradient (and the correct concentrations of Na^+ and K^+) in living cells. The sodium-potassium pump moves K^+ into the cell while moving Na^+ out at the same time, at a ratio of three Na^+ for every two K^+ ions moved in. The Na^+-K^+ ATPase exists in two forms, depending on its orientation to the interior or exterior of the cell and its affinity for either sodium or potassium ions. The process consists of the following six steps.

1. With the enzyme oriented towards the interior of the cell, the carrier has a high affinity for sodium ions. Three ions bind to the protein.
2. ATP is hydrolyzed by the protein carrier and a low-energy phosphate group attaches to it.
3. As a result, the carrier changes shape and re-orients itself towards the exterior of the membrane. The protein’s affinity for sodium decreases and the three sodium ions leave the carrier.
4. The shape change increases the carrier’s affinity for potassium ions, and two such ions attach to the protein. Subsequently, the low-energy phosphate group detaches from the carrier.

5. With the phosphate group removed and potassium ions attached, the carrier protein repositions itself towards the interior of the cell.
6. The carrier protein, in its new configuration, has a decreased affinity for potassium, and the two ions are released into the cytoplasm. The protein now has a higher affinity for sodium ions, and the process starts again.

Several things have happened as a result of this process. At this point, there are more sodium ions outside of the cell than inside and more potassium ions inside than out. For every three ions of sodium that move out, two ions of potassium move in. This results in the interior being slightly more negative relative to the exterior. This difference in charge is important in creating the conditions necessary for the secondary process. The sodium-potassium pump is, therefore, an electrogenic pump (a pump that creates a charge imbalance), creating an electrical imbalance across the membrane and contributing to the membrane potential.

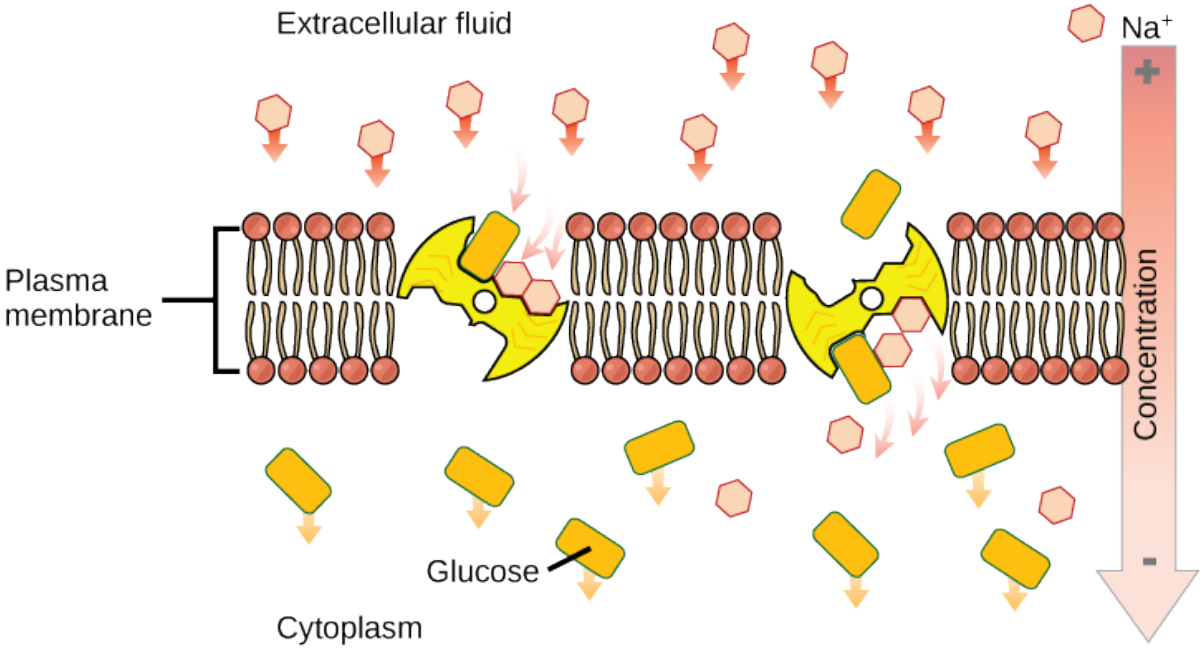
Link to Learning

Watch the video to see a simulation of active transport in a sodium-potassium ATPase.

Visit this page in your course online to use this activity.

Secondary Active Transport (Co-transport)

Secondary active transport brings sodium ions, and possibly other compounds, into the cell. As sodium ion concentrations build outside of the plasma membrane because of the action of the primary active transport process, an electrochemical gradient is created. If a channel protein exists and is open, the sodium ions will be pulled through the membrane. This movement is used to transport other substances that can attach themselves to the transport protein through the membrane (Figure). Many amino acids, as well as glucose, enter a cell this way. This secondary process is also used to store high-energy hydrogen ions in the mitochondria of plant and animal cells for the production of ATP. The potential energy that accumulates in the stored hydrogen ions is translated into kinetic energy as the ions surge through the channel protein ATP synthase, and that energy is used to convert ADP into ATP.
ART CONNECTION



An electrochemical gradient, created by primary active transport, can move other substances against their concentration gradients, a process called co-transport or secondary active transport. (credit: modification of work by Mariana Ruiz Villareal)

If the pH outside the cell decreases, would you expect the amount of amino acids transported into the cell to increase or decrease?

Section Summary

The combined gradient that affects an ion includes its concentration gradient and its electrical gradient. A positive ion, for example, might tend to diffuse into a new area, down its concentration gradient, but if it is diffusing into an area of net positive charge, its diffusion will be hampered by its electrical gradient. When dealing with ions in aqueous solutions, a combination of the electrochemical and concentration gradients, rather than just the concentration gradient alone, must be considered. Living cells need certain substances that exist inside the cell in concentrations greater than they exist in the extracellular space. Moving substances up their electrochemical gradients requires energy from the cell. Active transport uses energy stored in ATP to fuel this transport. Active transport of small molecular-sized materials uses integral proteins in the cell membrane to move the materials: These proteins are analogous to pumps. Some pumps, which carry out primary active transport, couple directly with ATP to drive their action. In co-transport (or secondary active transport), energy from primary transport can be used to move another substance into the cell and up its concentration gradient.

Art Connections

Figure Injection of a potassium solution into a person's blood is lethal; this is used in capital punishment and euthanasia. Why do you think a potassium solution injection is lethal?

Figure If the pH outside the cell decreases, would you expect the amount of amino acids transported into the cell to increase or decrease?

Review Questions

Active transport must function continuously because _____.

1. plasma membranes wear out
2. not all membranes are amphiphilic
3. facilitated transport opposes active transport
4. diffusion is constantly moving solutes in opposite directions

How does the sodium-potassium pump make the interior of the cell negatively charged?

1. by expelling anions
2. by pulling in anions
3. by expelling more cations than are taken in
4. by taking in and expelling an equal number of cations

What is the combination of an electrical gradient and a concentration gradient called?

1. potential gradient
2. electrical potential
3. concentration potential
4. electrochemical gradient

Free Response

Where does the cell get energy for active transport processes?

How does the sodium-potassium pump contribute to the net negative charge of the interior of the cell?

Glossary

active transport method of transporting material that requires energy

antiporter transporter that carries two ions or small molecules in different directions

electrochemical gradient gradient produced by the combined forces of an electrical gradient and a chemical gradient

electrogenic pump pump that creates a charge imbalance

primary active transport active transport that moves ions or small molecules across a membrane and may create a difference in charge across that membrane

pump active transport mechanism that works against electrochemical gradients

secondary active transport movement of material that is due to the electrochemical gradient established by primary active transport

symporter transporter that carries two different ions or small molecules, both in the same direction

transporter specific carrier proteins or pumps that facilitate movement

uniporter transporter that carries one specific ion or molecule

Licensing & Attributions

CC licensed content, Shared previously

- active transport. **Authored by**: open stax. **Provided by**: open stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:CmTJkys8@6/Active-Transport. **License**: *CC BY: Attribution*

BULK TRANSPORT

5.4 Bulk Transport

Summary

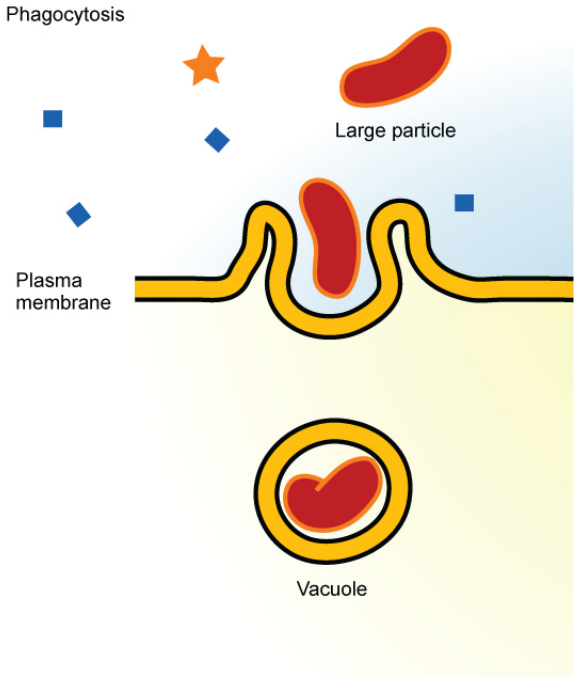
In addition to moving small ions and molecules through the membrane, cells also need to remove and take in larger molecules and particles (see Table for examples). Some cells are even capable of engulfing entire unicellular microorganisms. You might have correctly hypothesized that the uptake and release of large particles by the cell requires energy. A large particle, however, cannot pass through the membrane, even with energy supplied by the cell.

Endocytosis

Endocytosis is a type of active transport that moves particles, such as large molecules, parts of cells, and even whole cells, into a cell. There are different variations of endocytosis, but all share a common characteristic: The plasma membrane of the cell invaginates, forming a pocket around the target particle. The pocket pinches off, resulting in the particle being contained in a newly created intracellular vesicle formed from the plasma membrane.

Phagocytosis

Phagocytosis (the condition of “cell eating”) is the process by which large particles, such as cells or relatively large particles, are taken in by a cell. For example, when microorganisms invade the human body, a type of white blood cell called a neutrophil will remove the invaders through this process, surrounding and engulfing the microorganism, which is then destroyed by the neutrophil (Figure).

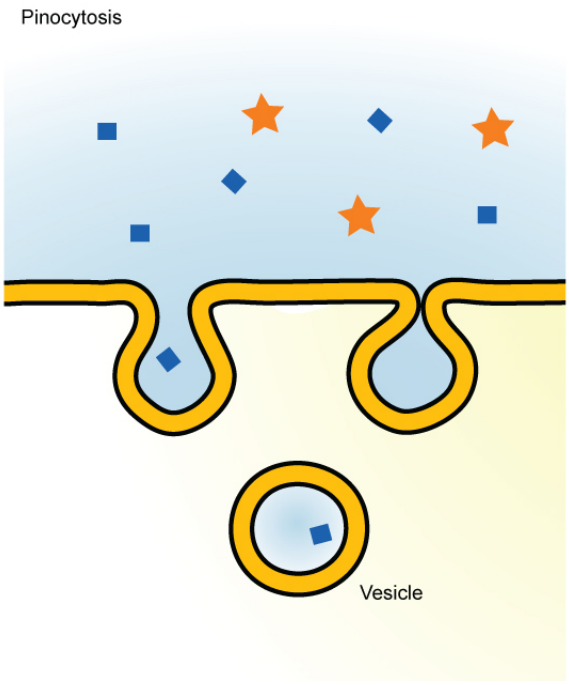


In phagocytosis, the cell membrane surrounds the particle and engulfs it. (credit: Mariana Ruiz Villareal)

In preparation for phagocytosis, a portion of the inward-facing surface of the plasma membrane becomes coated with a protein called clathrin, which stabilizes this section of the membrane. The coated portion of the membrane then extends from the body of the cell and surrounds the particle, eventually enclosing it. Once the vesicle containing the particle is enclosed within the cell, the clathrin disengages from the membrane and the vesicle merges with a lysosome for the breakdown of the material in the newly formed compartment (endosome). When accessible nutrients from the degradation of the vesicular contents have been extracted, the newly formed endosome merges with the plasma membrane and releases its contents into the extracellular fluid. The endosomal membrane again becomes part of the plasma membrane.

Pinocytosis

A variation of endocytosis is called pinocytosis. This literally means “cell drinking” and was named at a time when the assumption was that the cell was purposefully taking in extracellular fluid. In reality, this is a process that takes in molecules, including water, which the cell needs from the extracellular fluid. Pinocytosis results in a much smaller vesicle than does phagocytosis, and the vesicle does not need to merge with a lysosome (Figure).

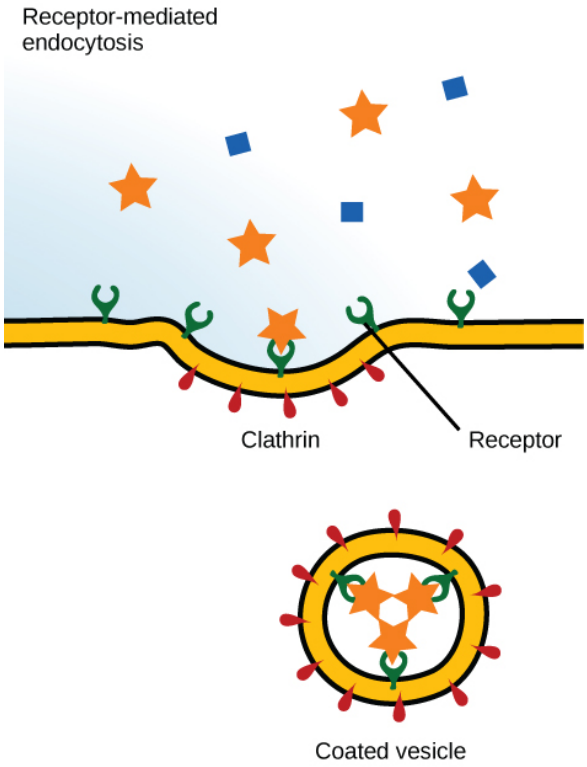


In pinocytosis, the cell membrane invaginates, surrounds a small volume of fluid, and pinches off. (credit: Mariana Ruiz Villareal)

A variation of pinocytosis is called potocytosis. This process uses a coating protein, called caveolin, on the cytoplasmic side of the plasma membrane, which performs a similar function to clathrin. The cavities in the plasma membrane that form the vacuoles have membrane receptors and lipid rafts in addition to caveolin. The vacuoles or vesicles formed in caveolae (singular caveola) are smaller than those in pinocytosis. Potocytosis is used to bring small molecules into the cell and to transport these molecules through the cell for their release on the other side of the cell, a process called transcytosis.

Receptor-mediated Endocytosis

A targeted variation of endocytosis employs receptor proteins in the plasma membrane that have a specific binding affinity for certain substances (Figure).



In receptor-mediated endocytosis, uptake of substances by the cell is targeted to a single type of substance that binds to the receptor on the external surface of the cell membrane. (credit: modification of work by Mariana Ruiz Villareal)

In receptor-mediated endocytosis, as in phagocytosis, clathrin is attached to the cytoplasmic side of the plasma membrane. If uptake of a compound is dependent on receptor-mediated endocytosis and the process is ineffective, the material will not be removed from the tissue fluids or blood. Instead, it will stay in those fluids and increase in concentration. Some human diseases are caused by the failure of receptor-mediated endocytosis. For example, the form of cholesterol termed low-density lipoprotein or LDL (also referred to as “bad” cholesterol) is removed from the blood by receptor-mediated endocytosis. In the human genetic disease familial hypercholesterolemia, the LDL receptors are defective or missing entirely. People with this condition have life-threatening levels of cholesterol in their blood, because their cells cannot clear LDL particles from their blood.

Although receptor-mediated endocytosis is designed to bring specific substances that are normally found in the extracellular fluid into the cell, other substances may gain entry into the cell at the same site. Flu viruses, diphtheria, and cholera toxin all have sites that cross-react with normal receptor-binding sites and gain entry into cells.

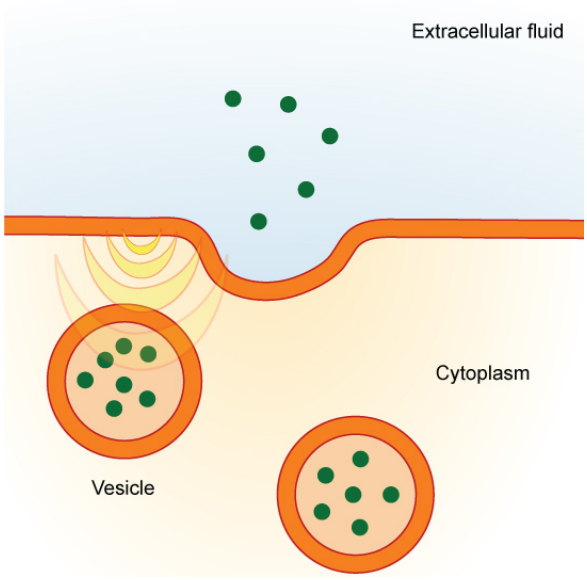
Link to Learning

See receptor-mediated endocytosis in action, and click on different parts for a focused animation.

Exocytosis

The reverse process of moving material into a cell is the process of exocytosis. Exocytosis is the opposite of the processes discussed above in that its purpose is to expel material from the cell into the extracellular fluid. Waste material is enveloped in a membrane and fuses with the interior of the plasma membrane. This fusion opens the membranous envelope on the exterior of the cell, and the waste material is expelled into the extracellular space (Figure). Other examples of cells releasing molecules via exocytosis include the secretion of proteins of the extracellular matrix and secretion of neurotransmitters into the synaptic cleft by synaptic vesicles.

Exocytosis



In exocytosis, vesicles containing substances fuse with the plasma membrane. The contents are then released to the exterior of the cell. (credit: modification of work by Mariana Ruiz Villareal)

Methods of Transport, Energy Requirements, and Types of Material Transported		
Transport Method	Active/Passive	Material Transported
Diffusion	Passive	Small-molecular weight material
Osmosis	Passive	Water
Facilitated transport/diffusion	Passive	Sodium, potassium, calcium, glucose
Primary active transport	Active	Sodium, potassium, calcium
Secondary active transport	Active	Amino acids, lactose
Phagocytosis	Active	Large macromolecules, whole cells, or cellular structures
Pinocytosis and potocytosis	Active	Small molecules (liquids/water)
Receptor-mediated endocytosis	Active	Large quantities of macromolecules

Section Summary

Active transport methods require the direct use of ATP to fuel the transport. Large particles, such as macromolecules, parts of cells, or whole cells, can be engulfed by other cells in a process called phagocytosis. In phagocytosis, a portion of the membrane invaginates and flows around the particle, eventually pinching off and leaving the particle entirely enclosed by an envelope of plasma membrane. Vesicle contents are broken down by the cell, with the particles either used as food or dispatched. Pinocytosis is a similar process on a smaller scale. The plasma membrane invaginates and pinches off, producing a small envelope of fluid from outside the cell. Pinocytosis imports substances that the cell needs from the extracellular fluid. The cell expels waste in a similar but reverse manner: it pushes a membranous vacuole to the plasma membrane, allowing the vacuole to fuse with the membrane and incorporate itself into the membrane structure, releasing its contents to the exterior.

Review Questions

What happens to the membrane of a vesicle after exocytosis?

1. It leaves the cell.
2. It is disassembled by the cell.
3. It fuses with and becomes part of the plasma membrane.
4. It is used again in another exocytosis event.

Which transport mechanism can bring whole cells into a cell?

1. pinocytosis
2. phagocytosis
3. facilitated transport
4. primary active transport

In what important way does receptor-mediated endocytosis differ from phagocytosis?

1. It transports only small amounts of fluid.
2. It does not involve the pinching off of membrane.
3. It brings in only a specifically targeted substance.
4. It brings substances into the cell, while phagocytosis removes substances.

Free Response

Why is it important that there are different types of proteins in plasma membranes for the transport of materials into and out of a cell?

Why do ions have a difficult time getting through plasma membranes despite their small size?

Glossary

caveolin protein that coats the cytoplasmic side of the plasma membrane and participates in the process of liquid update by potocytosis

clathrin protein that coats the inward-facing surface of the plasma membrane and assists in the formation of specialized structures, like coated pits, for phagocytosis

endocytosis type of active transport that moves substances, including fluids and particles, into a cell

exocytosis process of passing bulk material out of a cell

pinocytosis a variation of endocytosis that imports macromolecules that the cell needs from the extracellular fluid

potocytosis variation of pinocytosis that uses a different coating protein (caveolin) on the cytoplasmic side of the plasma membrane

receptor-mediated endocytosis variation of endocytosis that involves the use of specific binding proteins in the plasma membrane for specific molecules or particles, and clathrin-coated pits that become clathrin-coated vesicles

STUDY GUIDE: CELL MEMBRANE

Study Questions

Objective: Relate the structure of the cell membrane to is function as a semi-permeable barrier between intracellular fluid and extracellular fluid.

Use this page to check your understanding of the's content.

Vocabulary

1. Phospholipid
2. Phospholipid bilayer
3. Semipermeable
4. Exocytosis
5. Endocytosis
6. Tonicity
7. Osmosis
8. Diffusion
9. Extracellular fluid
10. Intracellular fluid

Study Guide Questions

1. What is the purpose of the cell membrane?
2. Describe the structure of the cell membrane.
3. How does the structure of the cell membrane affect its function? What IS its function?Use the words "hydrophobic" and "hydrophyllic" to describe the cell membrane.
4. Why is the cell membrane said to be a "fluid mosaic"?
5. What role(s) do proteins play in the cell membrane?
6. What role(s) does cholesterol play in the cell membrane?
7. What is unique about the cell membrane of an archaean prokaryote?
8. How do substances get into and out of the cell?
9. Compare and contrast primary and secondary active transport.
10. Compare and contrast active and passive transport.
11. Compare and contrast simple diffusion and facilitated diffusion
12. Compare and contrast endocytosis and exocytosis.
13. Define osmosis.
14. Given any scenario, predict the movement of water molecules across a semipermeable membrane.
15. Describe the concept of "tonicity".
16. Given any solution, be able to label the solution "hyper, hypo or isotonic".

Licensing & Attributions

CC licensed content, Shared previously

- Study Guide: Cell Membrane. **Authored by**: Wendy Riggs. **Provided by**: College of the Redwoods. **Project**: Kaleidoscope. **License**: *CC BY: Attribution*

ENERGY AND METABOLISM

INTRO

Introduction

A hummingbird needs energy to maintain prolonged periods of flight. The bird obtains its energy from taking in food and transforming the nutrients into energy through a series of biochemical reactions. The flight muscles in birds are extremely efficient in energy production. (credit: modification of work by Cory Zanker)

Virtually every task performed by living organisms requires energy. Energy is needed to perform heavy labor and exercise, but humans also use a great deal of energy while thinking, and even during sleep. In fact, the living cells of every organism constantly use energy. Nutrients and other molecules are imported, metabolized (broken down) and possibly synthesized into new molecules, modified if needed, transported around the cell, and may be distributed to the entire organism. For example, the large proteins that make up muscles are actively built from smaller molecules. Complex carbohydrates are broken down into simple sugars that the cell uses for energy. Just as energy is required to both build and demolish a building, energy is required for both the synthesis and breakdown of molecules. Additionally, signaling molecules such as hormones and neurotransmitters are transported between cells. Pathogenic bacteria and viruses are ingested and broken down by cells. Cells must also export waste and toxins to stay healthy, and many cells must swim or move surrounding materials via the beating motion of cellular appendages like cilia and flagella.

The cellular processes listed above require a steady supply of energy. From where, and in what form, does this energy come? How do living cells obtain energy, and how do they use it? This chapter will discuss different forms of energy and the physical laws that govern energy transfer. This chapter will also describe how cells use energy and replenish it, and how chemical reactions in the cell are performed with great efficiency.

Licensing & Attributions

Public domain content

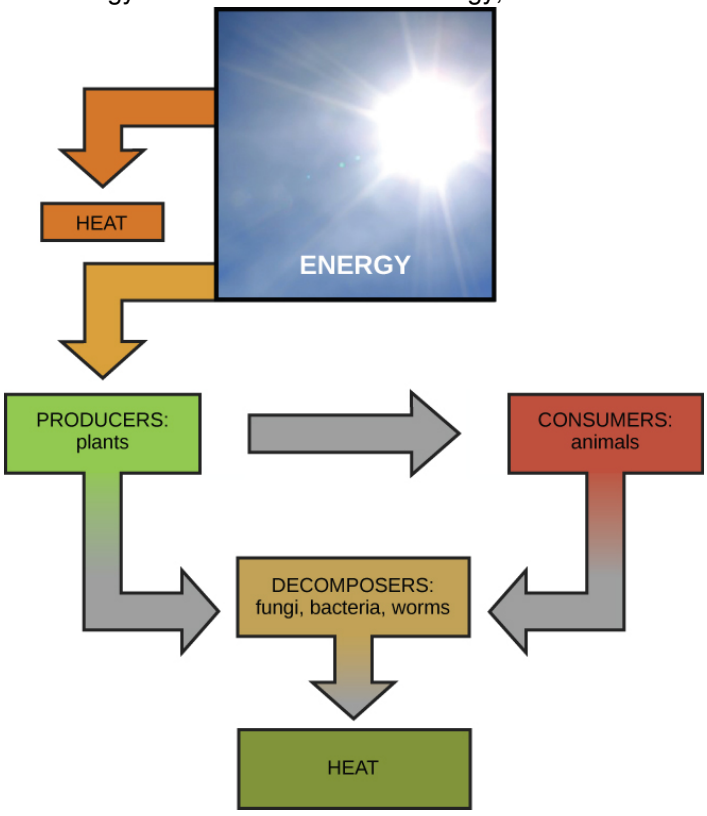
- Biology book chapter 6. **Authored by**: open stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.58:d9FA8FWC@2/Introduction. **License**: *CC BY: Attribution*

ENERGY AND METABOLISM

6.1 Energy and Metabolism

Summary

Scientists use the term bioenergetics to discuss the concept of energy flow (Figure) through living systems, such as cells. Cellular processes such as the building and breaking down of complex molecules occur through stepwise chemical reactions. Some of these chemical reactions are spontaneous and release energy, whereas others require energy to proceed. Just as living things must continually consume food to replenish what has been used, cells must continually produce more energy to replenish that used by the many energy-requiring chemical reactions that constantly take place. All of the chemical reactions that take place inside cells, including those that use energy and those that release energy, are the cell's metabolism.



Most life forms on earth get their energy from the sun. Plants use photosynthesis to capture sunlight, and herbivores eat those plants to obtain energy. Carnivores eat the herbivores, and decomposers digest plant and animal matter.

Metabolism of Carbohydrates

The metabolism of sugar (a simple carbohydrate) is a classic example of the many cellular processes that use and produce energy. Living things consume sugar as a major energy source, because sugar molecules have a

great deal of energy stored within their bonds. The breakdown of glucose, a simple sugar, is described by the equation:
C6H12O6+6O2→6CO2+6H2O+energysize 12{C rSub { size 8{6} } H rSub { size 8{12} } O rSub { size 8{2} } } {}

Carbohydrates that are consumed have their origins in photosynthesizing organisms like plants (Figure). During photosynthesis, plants use the energy of sunlight to convert carbon dioxide gas (CO_2) into sugar molecules, like glucose ($C_6H_{12}O_6$). Because this process involves synthesizing a larger, energy-storing molecule, it requires an input of energy to proceed. The synthesis of glucose is described by this equation (notice that it is the reverse of the previous equation):
6CO2+6H2O+energy→C6H12O6+6O2size 12{C rSub { size 8{6} } H rSub { size 8{12} } O rSub { size 8{2} } } {}

During the chemical reactions of photosynthesis, energy is provided in the form of a very high-energy molecule called ATP, or adenosine triphosphate, which is the primary energy currency of all cells. Just as the dollar is used as currency to buy goods, cells use molecules of ATP as energy currency to perform immediate work. The sugar (glucose) is stored as starch or glycogen. Energy-storing polymers like these are broken down into glucose to supply molecules of ATP.

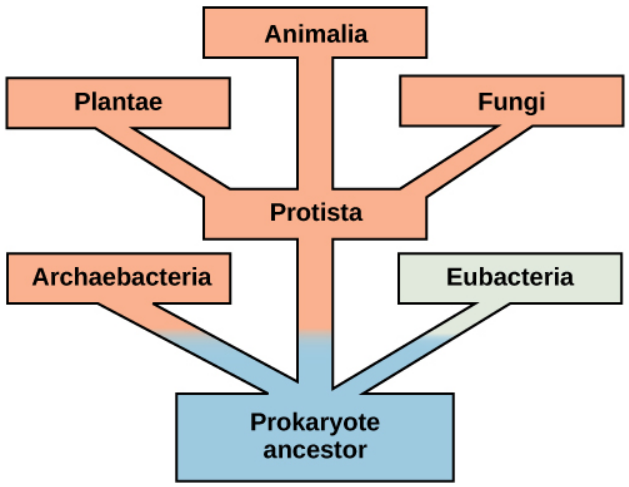
Solar energy is required to synthesize a molecule of glucose during the reactions of photosynthesis. In photosynthesis, light energy from the sun is initially transformed into chemical energy that is temporally stored in the energy carrier molecules ATP and NADPH (nicotinamide adenine dinucleotide phosphate). The stored energy in ATP and NADPH is then used later in photosynthesis to build one molecule of glucose from six molecules of CO_2. This process is analogous to eating breakfast in the morning to acquire energy for your body that can be used later in the day. Under ideal conditions, energy from 18 molecules of ATP is required to synthesize one molecule of glucose during the reactions of photosynthesis. Glucose molecules can also be combined with and converted into other types of sugars. When sugars are consumed, molecules of glucose eventually make their way into each living cell of the organism. Inside the cell, each sugar molecule is broken down through a complex series of chemical reactions. The goal of these reactions is to harvest the energy stored inside the sugar molecules. The harvested energy is used to make high-energy ATP molecules, which can be used to perform work, powering many chemical reactions in the cell. The amount of energy needed to make one molecule of glucose from six molecules of carbon dioxide is 18 molecules of ATP and 12 molecules of NADPH (each one of which is energetically equivalent to three molecules of ATP), or a total of 54 molecule equivalents required for the synthesis of one molecule of glucose. This process is a fundamental and efficient way for cells to generate the molecular energy that they require.

Plants, like this oak tree and acorn, use energy from sunlight to make sugar and other organic molecules. Both plants and animals (like this squirrel) use cellular respiration to derive energy from the organic molecules originally produced by plants. (credit “acorn”: modification of work by Noel Reynolds; credit “squirrel”: modification of work by Dawn Huczek)

Metabolic Pathways

The processes of making and breaking down sugar molecules illustrate two types of metabolic pathways. A metabolic pathway is a series of interconnected biochemical reactions that convert a substrate molecule or molecules, step-by-step, through a series of metabolic intermediates, eventually yielding a final product or products. In the case of sugar metabolism, the first metabolic pathway synthesized sugar from smaller molecules, and the other pathway broke sugar down into smaller molecules. These two opposite processes—the first requiring energy and the second producing energy—are referred to as anabolic (building) and catabolic (breaking down) pathways, respectively. Consequently, metabolism is composed of building (anabolism) and degradation (catabolism).

EVOLUTION CONNECTION



This tree shows the evolution of the various branches of life. The vertical dimension is time. Early life forms, in blue, used anaerobic metabolism to obtain energy from their surroundings.

Evolution of Metabolic PathwaysThere is more to the complexity of metabolism than understanding the metabolic pathways alone. Metabolic complexity varies from organism to organism. Photosynthesis is the primary pathway in which photosynthetic organisms like plants (the majority of global synthesis is done by planktonic algae) harvest the sun's energy and convert it into carbohydrates. The by-product of photosynthesis is oxygen, required by some cells to carry out cellular respiration. During cellular respiration, oxygen aids in the catabolic breakdown of carbon compounds, like carbohydrates. Among the products of this catabolism are CO_2 and ATP. In addition, some eukaryotes perform catabolic processes without oxygen (fermentation); that is, they perform or use anaerobic metabolism.

Organisms probably evolved anaerobic metabolism to survive (living organisms came into existence about 3.8 billion years ago, when the atmosphere lacked oxygen). Despite the differences between organisms and the complexity of metabolism, researchers have found that all branches of life share some of the same metabolic pathways, suggesting that all organisms evolved from the same ancient common ancestor (Figure). Evidence indicates that over time, the pathways diverged, adding specialized enzymes to allow organisms to better adapt to their environment, thus increasing their chance to survive. However, the underlying principle remains that all organisms must harvest energy from their environment and convert it to ATP to carry out cellular functions.

Anabolic and Catabolic Pathways

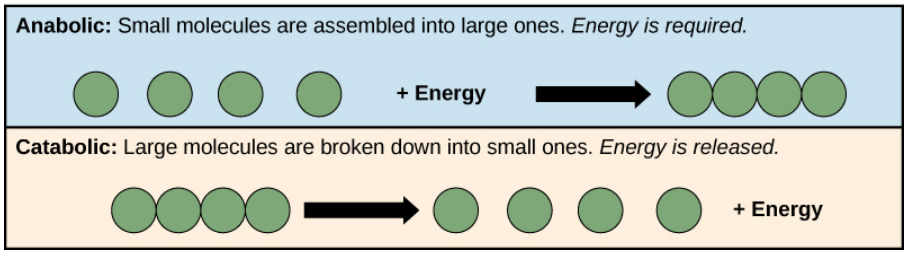
Anabolic pathways require an input of energy to synthesize complex molecules from simpler ones. Synthesizing sugar from CO_2 is one example. Other examples are the synthesis of large proteins from amino acid building blocks, and the synthesis of new DNA strands from nucleic acid building blocks. These biosynthetic processes are critical to the life of the cell, take place constantly, and demand energy provided by ATP and other high-energy molecules like NADH (nicotinamide adenine dinucleotide) and NADPH (Figure).

ATP is an important molecule for cells to have in sufficient supply at all times. The breakdown of sugars illustrates how a single molecule of glucose can store enough energy to make a great deal of ATP, 36 to 38 molecules. This

is a catabolic pathway. Catabolic pathways involve the degradation (or breakdown) of complex molecules into simpler ones. Molecular energy stored in the bonds of complex molecules is released in catabolic pathways and harvested in such a way that it can be used to produce ATP. Other energy-storing molecules, such as fats, are also broken down through similar catabolic reactions to release energy and make ATP (Figure).

It is important to know that the chemical reactions of metabolic pathways don't take place spontaneously. Each reaction step is facilitated, or catalyzed, by a protein called an enzyme. Enzymes are important for catalyzing all types of biological reactions—those that require energy as well as those that release energy.

Metabolic pathways



Anabolic pathways are those that require energy to synthesize larger molecules. Catabolic pathways are those that generate energy by breaking down larger molecules. Both types of pathways are required for maintaining the cell's energy balance.

Section Summary

Cells perform the functions of life through various chemical reactions. A cell's metabolism refers to the chemical reactions that take place within it. There are metabolic reactions that involve the breaking down of complex chemicals into simpler ones, such as the breakdown of large macromolecules. This process is referred to as catabolism, and such reactions are associated with a release of energy. On the other end of the spectrum, anabolism refers to metabolic processes that build complex molecules out of simpler ones, such as the synthesis of macromolecules. Anabolic processes require energy. Glucose synthesis and glucose breakdown are examples of anabolic and catabolic pathways, respectively.

Multiple Choice

Energy is stored long-term in the bonds of _____ and used short-term to perform work from a(n) _____ molecule.

1. ATP : glucose
2. an anabolic molecule : catabolic molecule
3. glucose : ATP
4. a catabolic molecule : anabolic molecule

DNA replication involves unwinding two strands of parent DNA, copying each strand to synthesize complementary strands, and releasing the parent and daughter DNA. Which of the following accurately describes this process?

1. This is an anabolic process
2. This is a catabolic process
3. This is both anabolic and catabolic
4. This is a metabolic process but is neither anabolic nor catabolic

Free Response

Does physical exercise involve anabolic and/or catabolic processes? Give evidence for your answer.

Name two different cellular functions that require energy that parallel human energy-requiring functions.

Glossary

anabolic (also, anabolism) pathways that require an input of energy to synthesize complex molecules from simpler ones

bioenergetics study of energy flowing through living systems

catabolic (also, catabolism) pathways in which complex molecules are broken down into simpler ones

metabolism all the chemical reactions that take place inside cells, including anabolism and catabolism

Licensing & Attributions

CC licensed content, Shared previously

- energy and metabolism. **Authored by**: open stax. **Provided by**: open stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:7V33N3f_@8/Energy-and-Metabolism. **License**: *CC BY: Attribution*

POTENTIAL, KINETIC, FREE, AND ACTIVATION ENERGY

Potential, Kinetic, Free, and Activation Energy

Summary

Energy is defined as the ability to do work. As you've learned, energy exists in different forms. For example, electrical energy, light energy, and heat energy are all different types of energy. While these are all familiar types of energy that one can see or feel, there is another type of energy that is much less tangible. This energy is associated with something as simple as an object held above the ground. In order to appreciate the way energy flows into and out of biological systems, it is important to understand more about the different types of energy that exist in the physical world.

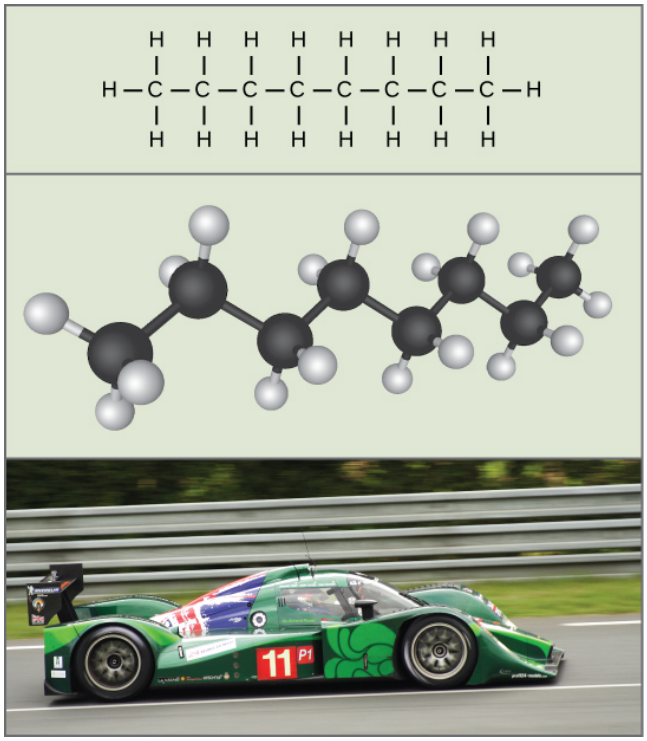
Types of Energy

When an object is in motion, there is energy associated with that object. In the example of an airplane in flight, there is a great deal of energy associated with the motion of the airplane. This is because moving objects are capable of enacting a change, or doing work. Think of a wrecking ball. Even a slow-moving wrecking ball can do a great deal of damage to other objects. However, a wrecking ball that is not in motion is incapable of performing work. Energy associated with objects in motion is called kinetic energy. A speeding bullet, a walking person, the rapid movement of molecules in the air (which produces heat), and electromagnetic radiation like light all have kinetic energy.

Now what if that same motionless wrecking ball is lifted two stories above a car with a crane? If the suspended wrecking ball is unmoving, is there energy associated with it? The answer is yes. The suspended wrecking ball has energy associated with it that is fundamentally different from the kinetic energy of objects in motion. This form of energy results from the fact that there is the*potential* for the wrecking ball to do work. If it is released, indeed it would do work. Because this type of energy refers to the potential to do work, it is called potential energy. Objects transfer their energy between kinetic and potential in the following way: As the wrecking ball hangs motionless, it has 0 kinetic and 100 percent potential energy. Once it is released, its kinetic energy begins to increase because it builds speed due to gravity. At the same time, as it nears the ground, it loses potential energy. Somewhere mid-fall it has 50 percent kinetic and 50 percent potential energy. Just before it hits the ground, the ball has nearly lost its potential energy and has near-maximal kinetic energy. Other examples of potential energy include the energy of water held behind a dam (Figure), or a person about to skydive out of an airplane.

Water behind a dam has potential energy. Moving water, such as in a waterfall or a rapidly flowing river, has kinetic energy. (credit “dam”: modification of work by “Pascal”/Flickr; credit “waterfall”: modification of work by Frank Gualtieri)

Potential energy is not only associated with the location of matter (such as a child sitting on a tree branch), but also with the structure of matter. A spring on the ground has potential energy if it is compressed; so does a rubber band that is pulled taut. The very existence of living cells relies heavily on structural potential energy. On a chemical level, the bonds that hold the atoms of molecules together have potential energy. Remember that anabolic cellular pathways require energy to synthesize complex molecules from simpler ones, and catabolic pathways release energy when complex molecules are broken down. The fact that energy can be released by the breakdown of certain chemical bonds implies that those bonds have potential energy. In fact, there is potential energy stored within the bonds of all the food molecules we eat, which is eventually harnessed for use. This is because these bonds can release energy when broken. The type of potential energy that exists within chemical bonds, and is released when those bonds are broken, is called chemical energy (Figure). Chemical energy is responsible for providing living cells with energy from food. The release of energy is brought about by breaking the molecular bonds within fuel molecules.



The molecules in gasoline (octane, the chemical formula shown) contain chemical energy within the chemical bonds. This energy is transformed into kinetic energy that allows a car to race on a racetrack. (credit “car”: modification of work by Russell Trow)

Link to Learning

Visit this site and select “A simple pendulum” on the menu (under “Harmonic Motion”) to see the shifting kinetic (K) and potential energy (U) of a pendulum in motion.

Free Energy

After learning that chemical reactions release energy when energy-storing bonds are broken, an important next question is how is the energy associated with chemical reactions quantified and expressed? How can the energy released from one reaction be compared to that of another reaction? A measurement of free energy is used to quantitate these energy transfers. Free energy is called Gibbs free energy (abbreviated with the letter G) after Josiah Willard Gibbs, the scientist who developed the measurement. Recall that according to the second law of thermodynamics, all energy transfers involve the loss of some amount of energy in an unusable form such as heat, resulting in entropy. Gibbs free energy specifically refers to the energy associated with a chemical reaction that is available after entropy is accounted for. In other words, Gibbs free energy is usable energy, or energy that is available to do work.

Every chemical reaction involves a change in free energy, called delta G (ΔG). The change in free energy can be calculated for any system that undergoes such a change, such as a chemical reaction. To calculate ΔG, subtract the amount of energy lost to entropy (denoted as ΔS) from the total energy change of the system. This total energy change in the system is called enthalpyand is denoted as ΔH . The formula for calculating ΔG is as follows, where the symbol T refers to absolute temperature in Kelvin (degrees Celsius + 273):
$\Delta G=\Delta H-T\Delta S\Delta G=\Delta H-T\Delta S$

The standard free energy change of a chemical reaction is expressed as an amount of energy per mole of the reaction product (either in kilojoules or kilocalories, kJ/mol or kcal/mol; 1 kJ = 0.239 kcal) under standard pH, temperature, and pressure conditions. Standard pH, temperature, and pressure conditions are generally calculated at pH 7.0 in biological systems, 25 degrees Celsius, and 100 kilopascals (1 atm pressure), respectively. It is important to note that cellular conditions vary considerably from these standard conditions, and so standard calculated ΔG values for biological reactions will be different inside the cell.

Endergonic Reactions and Exergonic Reactions

If energy is released during a chemical reaction, then the resulting value from the above equation will be a negative number. In other words, reactions that release energy have a $\Delta G < 0$. A negative ΔG also means that the products of the reaction have less free energy than the reactants, because they gave off some free energy during the reaction. Reactions that have a negative ΔG and consequently release free energy are called exergonic reactions. Think: *ex*ergonic means energy is *ex*iting the system. These reactions are also referred to as spontaneous reactions, because they can occur without the addition of energy into the system. Understanding which chemical reactions are spontaneous and release free energy is extremely useful for biologists, because these reactions can be harnessed to perform work inside the cell. An important distinction must be drawn between the term spontaneous and the idea of a chemical reaction that occurs immediately. Contrary to the everyday use of the term, a spontaneous reaction is not one that suddenly or quickly occurs. The rusting of iron is an example of a spontaneous reaction that occurs slowly, little by little, over time.

If a chemical reaction requires an input of energy rather than releasing energy, then the ΔG for that reaction will be a positive value. In this case, the products have more free energy than the reactants. Thus, the products of these reactions can be thought of as energy-storing molecules. These chemical reactions are called endergonic reactions, and they are non-spontaneous. An endergonic reaction will not take place on its own without the addition of free energy.

Let’s revisit the example of the synthesis and breakdown of the food molecule, glucose. Remember that the building of complex molecules, such as sugars, from simpler ones is an anabolic process and requires energy. Therefore, the chemical reactions involved in anabolic processes are endergonic reactions. On the other hand, the catabolic process of breaking sugar down into simpler molecules releases energy in a series of exergonic reactions. Like the example of rust above, the breakdown of sugar involves spontaneous reactions, but these reactions don’t occur instantaneously. Figure shows some other examples of endergonic and exergonic reactions.

Later sections will provide more information about what else is required to make even spontaneous reactions happen more efficiently.

ART CONNECTION

(a)

(b)

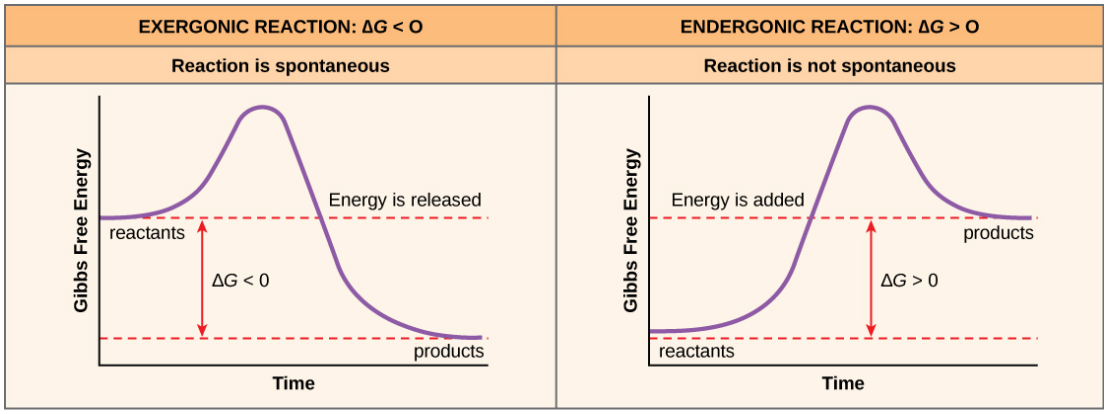
(c)

(d)

Shown are some examples of endergonic processes (ones that require energy) and exergonic processes (ones that release energy). These include (a) a compost pile decomposing, (b) a chick hatching from a fertilized egg, (c) sand art being destroyed, and (d) a ball rolling down a hill. (credit a: modification of work by Natalie Maynor; credit b: modification of work by USDA; credit c: modification of work by “Athlex”/Flickr; credit d: modification of work by Harry Malsch)

Look at each of the processes shown, and decide if it is endergonic or exergonic. In each case, does enthalpy increase or decrease, and does entropy increase or decrease?

An important concept in the study of metabolism and energy is that of chemical equilibrium. Most chemical reactions are reversible. They can proceed in both directions, releasing energy into their environment in one direction, and absorbing it from the environment in the other direction (Figure). The same is true for the chemical reactions involved in cell metabolism, such as the breaking down and building up of proteins into and from individual amino acids, respectively. Reactants within a closed system will undergo chemical reactions in both directions until a state of equilibrium is reached. This state of equilibrium is one of the lowest possible free energy and a state of maximal entropy. Energy must be put into the system to push the reactants and products away from a state of equilibrium. Either reactants or products must be added, removed, or changed. If a cell were a closed system, its chemical reactions would reach equilibrium, and it would die because there would be insufficient free energy left to perform the work needed to maintain life. In a living cell, chemical reactions are constantly moving towards equilibrium, but never reach it. This is because a living cell is an open system. Materials pass in and out, the cell recycles the products of certain chemical reactions into other reactions, and chemical equilibrium is never reached. In this way, living organisms are in a constant energy-requiring, uphill battle against equilibrium and entropy. This constant supply of energy ultimately comes from sunlight, which is used to produce nutrients in the process of photosynthesis.



Exergonic and endergonic reactions result in changes in Gibbs free energy. Exergonic reactions release energy; endergonic reactions require energy to proceed.

Activation Energy

There is another important concept that must be considered regarding endergonic and exergonic reactions. Even exergonic reactions require a small amount of energy input to get going before they can proceed with their energy-releasing steps. These reactions have a net release of energy, but still require some energy in the beginning. This small amount of energy input necessary for all chemical reactions to occur is called the activation energy (or free energy of activation) and is abbreviated E_A(Figure).

Why would an energy-releasing, negative ΔG reaction actually require some energy to proceed? The reason lies in the steps that take place during a chemical reaction. During chemical reactions, certain chemical bonds are broken and new ones are formed. For example, when a glucose molecule is broken down, bonds between the carbon atoms of the molecule are broken. Since these are energy-storing bonds, they release energy when broken. However, to get them into a state that allows the bonds to break, the molecule must be somewhat contorted. A small energy input is required to achieve this contorted state. This contorted state is called the transition state, and it is a high-energy, unstable state. For this reason, reactant molecules don't last long in their transition state, but very quickly proceed to the next steps of the chemical reaction. Free energy diagrams illustrate the energy profiles for a given reaction. Whether the reaction is exergonic or endergonic determines whether the products in the diagram will exist at a lower or higher energy state than both the reactants and the products. However, regardless of this measure, the transition state of the reaction exists at a higher energy state than the reactants, and thus, E_A is always positive.

Link to Learning

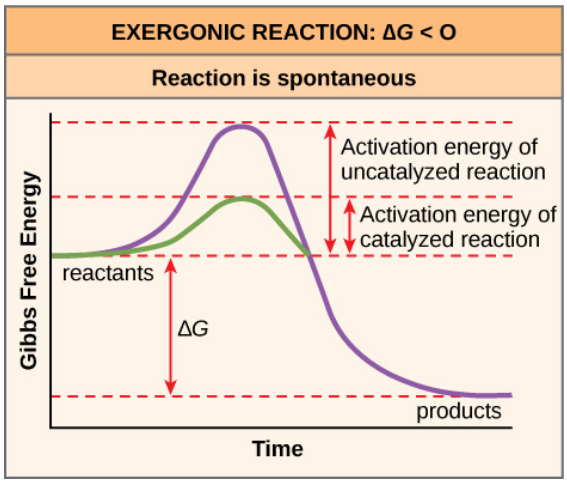
Watch an animation of the move from free energy to transition state at this site.

Where does the activation energy required by chemical reactants come from? The source of the activation energy needed to push reactions forward is typically heat energy from the surroundings. Heat energy (the total bond energy of reactants or products in a chemical reaction) speeds up the motion of molecules, increasing the frequency and force with which they collide; it also moves atoms and bonds within the molecule slightly, helping them reach their transition state. For this reason, heating up a system will cause chemical reactants within that system to react more frequently. Increasing the pressure on a system has the same effect. Once reactants have absorbed enough heat energy from their surroundings to reach the transition state, the reaction will proceed.

The activation energy of a particular reaction determines the rate at which it will proceed. The higher the activation energy, the slower the chemical reaction will be. The example of iron rusting illustrates an inherently slow reaction. This reaction occurs slowly over time because of its high E_A. Additionally, the burning of many fuels, which is strongly exergonic, will take place at a negligible rate unless their activation energy is overcome by sufficient heat from a spark. Once they begin to burn, however, the chemical reactions release enough heat to continue the burning process, supplying the activation energy for surrounding fuel molecules. Like these reactions outside of cells, the activation energy for most cellular reactions is too high for heat energy to overcome at efficient rates. In other words, in order for important cellular reactions to occur at appreciable rates (number of

reactions per unit time), their activation energies must be lowered (Figure); this is referred to as catalysis. This is a very good thing as far as living cells are concerned. Important macromolecules, such as proteins, DNA, and RNA, store considerable energy, and their breakdown is exergonic. If cellular temperatures alone provided enough heat energy for these exergonic reactions to overcome their activation barriers, the essential components of a cell would disintegrate.

ART CONNECTION



Activation energy is the energy required for a reaction to proceed, and it is lower if the reaction is catalyzed. The horizontal axis of this diagram describes the sequence of events in time.

If no activation energy were required to break down sucrose (table sugar), would you be able to store it in a sugar bowl?

Section Summary

Energy comes in many different forms. Objects in motion do physical work, and kinetic energy is the energy of objects in motion. Objects that are not in motion may have the potential to do work, and thus, have potential energy. Molecules also have potential energy because the breaking of molecular bonds has the potential to release energy. Living cells depend on the harvesting of potential energy from molecular bonds to perform work. Free energy is a measure of energy that is available to do work. The free energy of a system changes during energy transfers such as chemical reactions, and this change is referred to as ΔG.

The ΔG of a reaction can be negative or positive, meaning that the reaction releases energy or consumes energy, respectively. A reaction with a negative ΔG that gives off energy is called an exergonic reaction. One with a positive ΔG that requires energy input is called an endergonic reaction. Exergonic reactions are said to be spontaneous, because their products have less energy than their reactants. The products of endergonic reactions have a higher energy state than the reactants, and so these are nonspontaneous reactions. However, all reactions (including spontaneous -ΔG reactions) require an initial input of energy in order to reach the transition state, at which they'll proceed. This initial input of energy is called the activation energy.

Licensing & Attributions

CC licensed content, Shared previously

- potential kinetic free and activation energy. **Authored by**: open stax. **Provided by**: open stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:FrKEod5j@7/Potential-Kinetic-Free-and-Act. **License**: *CC BY: Attribution*

THE LAWS OF THERMODYNAMICS

The Laws of Thermodynamics

Summary

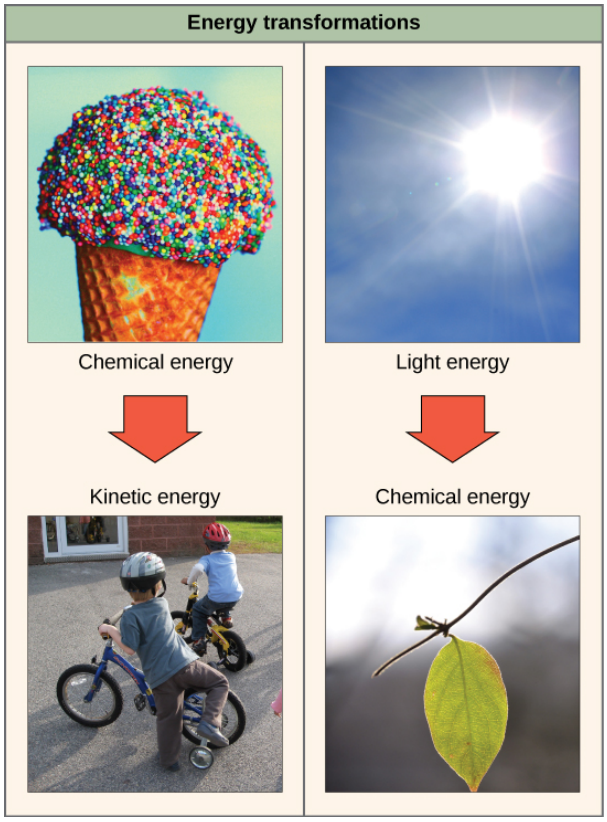
Thermodynamics refers to the study of energy and energy transfer involving physical matter. The matter and its environment relevant to a particular case of energy transfer are classified as a system, and everything outside of that system is called the surroundings. For instance, when heating a pot of water on the stove, the system includes the stove, the pot, and the water. Energy is transferred within the system (between the stove, pot, and water). There are two types of systems: open and closed. An open system is one in which energy can be transferred between the system and its surroundings. The stovetop system is open because heat can be lost into the air. A closed system is one that cannot transfer energy to its surroundings.

Biological organisms are open systems. Energy is exchanged between them and their surroundings, as they consume energy-storing molecules and release energy to the environment by doing work. Like all things in the physical world, energy is subject to the laws of physics. The laws of thermodynamics govern the transfer of energy in and among all systems in the universe.

The First Law of Thermodynamics

The first law of thermodynamics deals with the total amount of energy in the universe. It states that this total amount of energy is constant. In other words, there has always been, and always will be, exactly the same amount of energy in the universe. Energy exists in many different forms. According to the first law of thermodynamics, energy may be transferred from place to place or transformed into different forms, but it cannot be created or destroyed. The transfers and transformations of energy take place around us all the time. Light bulbs transform electrical energy into light energy. Gas stoves transform chemical energy from natural gas into heat energy. Plants perform one of the most biologically useful energy transformations on earth: that of converting the energy of sunlight into the chemical energy stored within organic molecules ([link]). Some examples of energy transformations are shown in Figure.

The challenge for all living organisms is to obtain energy from their surroundings in forms that they can transfer or transform into usable energy to do work. Living cells have evolved to meet this challenge very well. Chemical energy stored within organic molecules such as sugars and fats is transformed through a series of cellular chemical reactions into energy within molecules of ATP. Energy in ATP molecules is easily accessible to do work. Examples of the types of work that cells need to do include building complex molecules, transporting materials, powering the beating motion of cilia or flagella, contracting muscle fibers to create movement, and reproduction.



Shown are two examples of energy being transferred from one system to another and transformed from one form to another. Humans can convert the chemical energy in food, like this ice cream cone, into kinetic energy (the energy of movement to ride a bicycle). Plants can convert electromagnetic radiation (light energy) from the sun into chemical energy. (credit “ice cream”: modification of work by D. Sharon Pruitt; credit “kids on bikes”: modification of work by Michelle Riggen-Ransom; credit “leaf”: modification of work by Cory Zanker)

The Second Law of Thermodynamics

A living cell’s primary tasks of obtaining, transforming, and using energy to do work may seem simple. However, the second law of thermodynamics explains why these tasks are harder than they appear. None of the energy transfers we’ve discussed, along with all energy transfers and transformations in the universe, is completely efficient. In every energy transfer, some amount of energy is lost in a form that is unusable. In most cases, this form is heat energy. Thermodynamically, heat energy is defined as the energy transferred from one system to another that is not doing work. For example, when an airplane flies through the air, some of the energy of the flying plane is lost as heat energy due to friction with the surrounding air. This friction actually heats the air by temporarily increasing the speed of air molecules. Likewise, some energy is lost as heat energy during cellular metabolic reactions. This is good for warm-blooded creatures like us, because heat energy helps to maintain our body temperature. Strictly speaking, no energy transfer is completely efficient, because some energy is lost in an unusable form.

An important concept in physical systems is that of order and disorder (also known as randomness). The more energy that is lost by a system to its surroundings, the less ordered and more random the system is. Scientists refer to the measure of randomness or disorder within a system as entropy. High entropy means high disorder and low energy (Figure). To better understand entropy, think of a student’s bedroom. If no energy or work were put into it, the room would quickly become messy. It would exist in a very disordered state, one of high entropy. Energy must be put into the system, in the form of the student doing work and putting everything away, in order to bring the room back to a state of cleanliness and order. This state is one of low entropy. Similarly, a car or house must be constantly maintained with work in order to keep it in an ordered state. Left alone, the entropy of the house or car gradually increases through rust and degradation. Molecules and chemical reactions have varying amounts of entropy as well. For example, as chemical reactions reach a state of equilibrium, entropy increases, and as molecules at a high concentration in one place diffuse and spread out, entropy also increases.

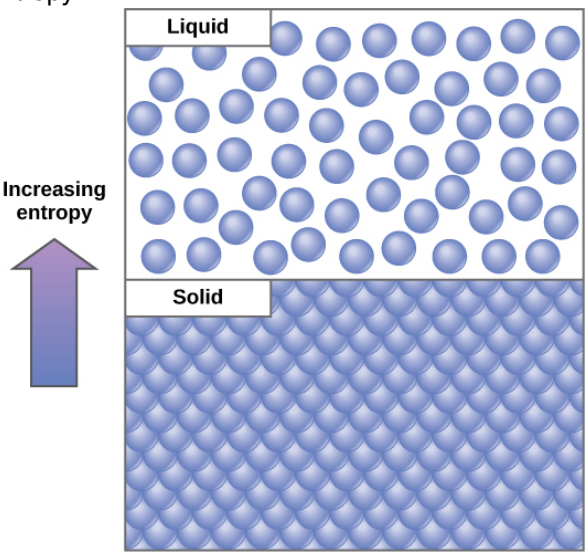
SCIENTIFIC CONNECTION

Transfer of Energy and the Resulting Entropy

Set up a simple experiment to understand how energy is transferred and how a change in entropy results.

1. Take a block of ice. This is water in solid form, so it has a high structural order. This means that the molecules cannot move very much and are in a fixed position. The temperature of the ice is 0°C. As a result, the entropy of the system is low.
2. Allow the ice to melt at room temperature. What is the state of molecules in the liquid water now? How did the energy transfer take place? Is the entropy of the system higher or lower? Why?
3. Heat the water to its boiling point. What happens to the entropy of the system when the water is heated?

All physical systems can be thought of in this way: Living things are highly ordered, requiring constant energy input to be maintained in a state of low entropy. As living systems take in energy-storing molecules and transform them through chemical reactions, they lose some amount of usable energy in the process, because no reaction is completely efficient. They also produce waste and by-products that aren't useful energy sources. This process increases the entropy of the system's surroundings. Since all energy transfers result in the loss of some usable energy, the second law of thermodynamics states that every energy transfer or transformation increases the entropy of the universe. Even though living things are highly ordered and maintain a state of low entropy, the entropy of the universe in total is constantly increasing due to the loss of usable energy with each energy transfer that occurs. Essentially, living things are in a continuous uphill battle against this constant increase in universal entropy.



Entropy is a measure of randomness or disorder in a system. Gases have higher entropy than liquids, and liquids have higher entropy than solids.

Section Summary

In studying energy, scientists use the term "system" to refer to the matter and its environment involved in energy transfers. Everything outside of the system is called the surroundings. Single cells are biological systems. Systems can be thought of as having a certain amount of order. It takes energy to make a system more ordered. The more ordered a system is, the lower its entropy. Entropy is a measure of the disorder of a system. As a system becomes more disordered, the lower its energy and the higher its entropy become.

A series of laws, called the laws of thermodynamics, describe the properties and processes of energy transfer. The first law states that the total amount of energy in the universe is constant. This means that energy can't be created or destroyed, only transferred or transformed. The second law of thermodynamics states that every energy transfer involves some loss of energy in an unusable form, such as heat energy, resulting in a more disordered system. In other words, no energy transfer is completely efficient and tends toward disorder.

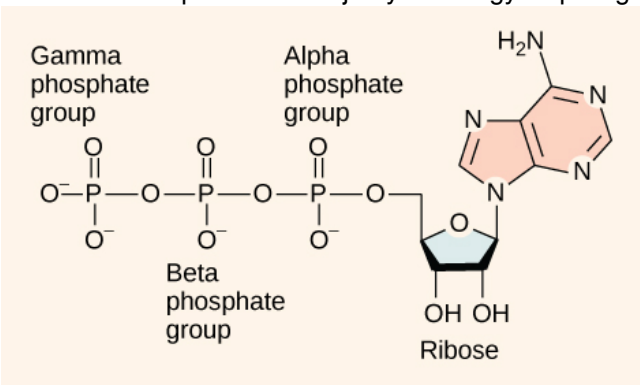
Licensing & Attributions

CC licensed content, Shared previously

- the laws of thermodynamics. **Authored by**: open stax. **Provided by**: open stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:HZOoT9Pk@7/The-Laws-of-Thermodynamics. **License**: *CC BY: Attribution*

ATP

Even exergonic, energy-releasing reactions require a small amount of activation energy in order to proceed. However, consider endergonic reactions, which require much more energy input, because their products have more free energy than their reactants. Within the cell, where does energy to power such reactions come from? The answer lies with an energy-supplying molecule called adenosine triphosphate, or ATP. ATP is a small, relatively simple molecule (Figure), but within some of its bonds, it contains the potential for a quick burst of energy that can be harnessed to perform cellular work. This molecule can be thought of as the primary energy currency of cells in much the same way that money is the currency that people exchange for things they need. ATP is used to power the majority of energy-requiring cellular reactions.



ATP is the primary energy currency of the cell. It has an adenosine backbone with three phosphate groups attached.

As its name suggests, adenosine triphosphate is comprised of adenosine bound to three phosphate groups (Figure). Adenosine is a nucleoside consisting of the nitrogenous base adenine and a five-carbon sugar, ribose. The three phosphate groups, in order of closest to furthest from the ribose sugar, are labeled alpha, beta, and gamma. Together, these chemical groups constitute an energy powerhouse. However, not all bonds within this molecule exist in a particularly high-energy state. Both bonds that link the phosphates are equally high-energy bonds (phosphoanhydride bonds) that, when broken, release sufficient energy to power a variety of cellular reactions and processes. These high-energy bonds are the bonds between the second and third (or beta and gamma) phosphate groups and between the first and second phosphate groups. The reason that these bonds are considered “high-energy” is because the products of such bond breaking—adenosine diphosphate (ADP) and one inorganic phosphate group (P_i)—have considerably lower free energy than the reactants: ATP and a water molecule. Because this reaction takes place with the use of a water molecule, it is considered a hydrolysis reaction. In other words, ATP is hydrolyzed into ADP in the following reaction:
ATP+H2O→ADP+Pi+free energysize 12{{ATP} + H rSub { size 8{2} } O ADP + P rSub { size 8{i}} + {free energy} } {}

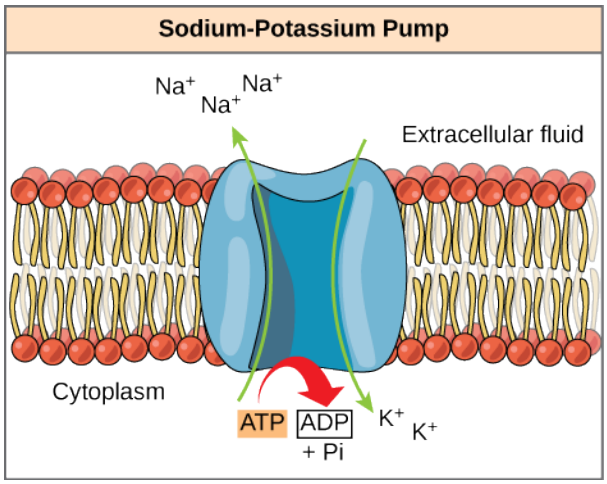
Like most chemical reactions, the hydrolysis of ATP to ADP is reversible. The reverse reaction regenerates ATP from ADP + P_i. Indeed, cells rely on the regeneration of ATP just as people rely on the regeneration of spent money through some sort of income. Since ATP hydrolysis releases energy, ATP regeneration must require an input of free energy. The formation of ATP is expressed in this equation:
ADP+Pi+free energy→ATP+H2Osize 12{{ATP} + H rSub { size 8{2} } O ADP + P rSub { size 8{i}} + {free energy} } {}

Two prominent questions remain with regard to the use of ATP as an energy source. Exactly how much free energy is released with the hydrolysis of ATP, and how is that free energy used to do cellular work? The

calculated ΔG for the hydrolysis of one mole of ATP into ADP and P_i is –7.3 kcal/mole (–30.5 kJ/mol). Since this calculation is true under standard conditions, it would be expected that a different value exists under cellular conditions. In fact, the ΔG for the hydrolysis of one mole of ATP in a living cell is almost double the value at standard conditions: 14 kcal/mol (–57 kJ/mol).

ATP is a highly unstable molecule. Unless quickly used to perform work, ATP spontaneously dissociates into ADP + P_i, and the free energy released during this process is lost as heat. The second question posed above, that is, how the energy released by ATP hydrolysis is used to perform work inside the cell, depends on a strategy called energy coupling. Cells couple the exergonic reaction of ATP hydrolysis with endergonic reactions, allowing them to proceed. One example of energy coupling using ATP involves a transmembrane ion pump that is extremely important for cellular function. This sodium-potassium pump (Na^+/K^+pump) drives sodium out of the cell and potassium into the cell (Figure). A large percentage of a cell's ATP is spent powering this pump, because cellular processes bring a great deal of sodium into the cell and potassium out of the cell. The pump works constantly to stabilize cellular concentrations of sodium and potassium. In order for the pump to turn one cycle (exporting three Na+ ions and importing two K^+ ions), one molecule of ATP must be hydrolyzed. When ATP is hydrolyzed, its gamma phosphate doesn't simply float away, but is actually transferred onto the pump protein. This process of a phosphate group binding to a molecule is called phosphorylation. As with most cases of ATP hydrolysis, a phosphate from ATP is transferred onto another molecule. In a phosphorylated state, the Na^+/K^+ pump has more free energy and is triggered to undergo a conformational change. This change allows it to release Na^+ to the outside of the cell. It then binds extracellular K^+, which, through another conformational change, causes the phosphate to detach from the pump. This release of phosphate triggers the K^+ to be released to the inside of the cell. Essentially, the energy released from the hydrolysis of ATP is coupled with the energy required to power the pump and transport Na^+ and K^+ ions. ATP performs cellular work using this basic form of energy coupling through phosphorylation.

ART CONNECTION



The sodium-potassium pump is an example of energy coupling. The energy derived from exergonic ATP hydrolysis is used to pump sodium and potassium ions across the cell membrane.

The hydrolysis of one ATP molecule releases 7.3 kcal/mol of energy (ΔG = –7.3 kcal/mol of energy). If it takes 2.1 kcal/mol of energy to move one Na^+ across the membrane (ΔG = +2.1 kcal/mol of energy), how many sodium ions could be moved by the hydrolysis of one ATP molecule?

Often during cellular metabolic reactions, such as the synthesis and breakdown of nutrients, certain molecules must be altered slightly in their conformation to become substrates for the next step in the reaction series. One example is during the very first steps of cellular respiration, when a molecule of the sugar glucose is broken down in the process of glycolysis. In the first step of this process, ATP is required for the phosphorylation of glucose, creating a high-energy but unstable intermediate. This phosphorylation reaction powers a conformational change that allows the phosphorylated glucose molecule to be converted to the phosphorylated sugar fructose. Fructose is a necessary intermediate for glycolysis to move forward. Here, the exergonic reaction of ATP hydrolysis is coupled with the endergonic reaction of converting glucose into a phosphorylated intermediate in the pathway. Once again, the energy released by breaking a phosphate bond within ATP was used for the phosphorylation of another molecule, creating an unstable intermediate and powering an important conformational change.

Link to Learning

See an interactive animation of the ATP-producing glycolysis process at this site.

Section Summary

ATP is the primary energy-supplying molecule for living cells. ATP is made up of a nucleotide, a five-carbon sugar, and three phosphate groups. The bonds that connect the phosphates (phosphoanhydride bonds) have high-energy content. The energy released from the hydrolysis of ATP into ADP + P_i is used to perform cellular work. Cells use ATP to perform work by coupling the exergonic reaction of ATP hydrolysis with endergonic reactions. ATP donates its phosphate group to another molecule via a process known as phosphorylation. The phosphorylated molecule is at a higher-energy state and is less stable than its unphosphorylated form, and this added energy from the addition of the phosphate allows the molecule to undergo its endergonic reaction.

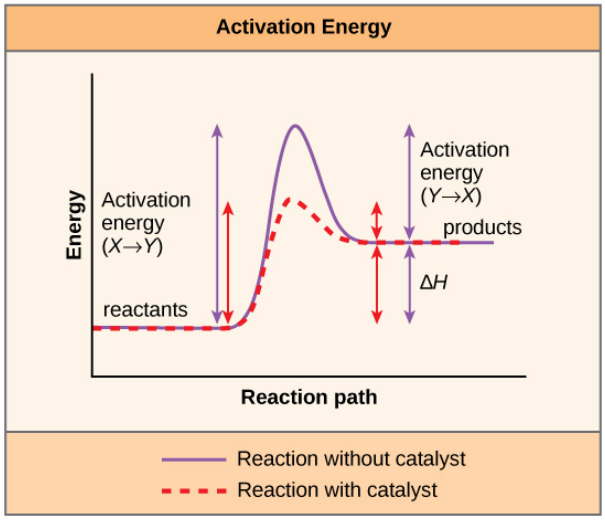
Licensing & Attributions

CC licensed content, Shared previously

- ATP. **Authored by:** open stax. **Provided by:** open stax. **Located at:** http://cnx.org/contents/GFy_h8cu@10.59:mcCadNdV@7/ATP-Adenosine-Triphosphate. **License:** *CC BY: Attribution*

ENZYMES

A substance that helps a chemical reaction to occur is a catalyst, and the special molecules that catalyze biochemical reactions are called enzymes. Almost all enzymes are proteins, made up of chains of amino acids, and they perform the critical task of lowering the activation energies of chemical reactions inside the cell. Enzymes do this by binding to the reactant molecules, and holding them in such a way as to make the chemical bond-breaking and bond-forming processes take place more readily. It is important to remember that enzymes don't change the ΔG of a reaction. In other words, they don't change whether a reaction is exergonic (spontaneous) or endergonic. This is because they don't change the free energy of the reactants or products. They only reduce the activation energy required to reach the transition state (Figure).



Enzymes lower the activation energy of the reaction but do not change the free energy of the reaction.

Enzyme Active Site and Substrate Specificity

The chemical reactants to which an enzyme binds are the enzyme's substrates. There may be one or more substrates, depending on the particular chemical reaction. In some reactions, a single-reactant substrate is

broken down into multiple products. In others, two substrates may come together to create one larger molecule. Two reactants might also enter a reaction, both become modified, and leave the reaction as two products. The location within the enzyme where the substrate binds is called the enzyme's active site. The active site is where the "action" happens, so to speak. Since enzymes are proteins, there is a unique combination of amino acid residues (also called side chains, or R groups) within the active site. Each residue is characterized by different properties. Residues can be large or small, weakly acidic or basic, hydrophilic or hydrophobic, positively or negatively charged, or neutral. The unique combination of amino acid residues, their positions, sequences, structures, and properties, creates a very specific chemical environment within the active site. This specific environment is suited to bind, albeit briefly, to a specific chemical substrate (or substrates). Due to this jigsaw puzzle-like match between an enzyme and its substrates (which adapts to find the best fit between the transition state and the active site), enzymes are known for their specificity. The "best fit" results from the shape and the amino acid functional group's attraction to the substrate. There is a specifically matched enzyme for each substrate and, thus, for each chemical reaction; however, there is flexibility as well.

The fact that active sites are so perfectly suited to provide specific environmental conditions also means that they are subject to influences by the local environment. It is true that increasing the environmental temperature generally increases reaction rates, enzyme-catalyzed or otherwise. However, increasing or decreasing the temperature outside of an optimal range can affect chemical bonds within the active site in such a way that they are less well suited to bind substrates. High temperatures will eventually cause enzymes, like other biological molecules, to denature, a process that changes the natural properties of a substance. Likewise, the pH of the local environment can also affect enzyme function. Active site amino acid residues have their own acidic or basic properties that are optimal for catalysis. These residues are sensitive to changes in pH that can impair the way substrate molecules bind. Enzymes are suited to function best within a certain pH range, and, as with temperature, extreme pH values (acidic or basic) of the environment can cause enzymes to denature.

Induced Fit and Enzyme Function

For many years, scientists thought that enzyme-substrate binding took place in a simple "lock-and-key" fashion. This model asserted that the enzyme and substrate fit together perfectly in one instantaneous step. However, current research supports a more refined view called induced fit (Figure). The induced-fit model expands upon the lock-and-key model by describing a more dynamic interaction between enzyme and substrate. As the enzyme and substrate come together, their interaction causes a mild shift in the enzyme's structure that confirms an ideal binding arrangement between the enzyme and the transition state of the substrate. This ideal binding maximizes the enzyme's ability to catalyze its reaction.

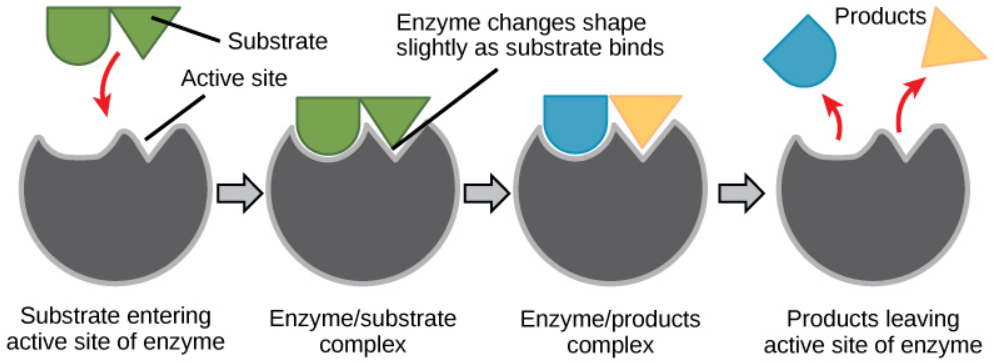
Link to Learning

View an animation of induced fit at this website.

When an enzyme binds its substrate, an enzyme-substrate complex is formed. This complex lowers the activation energy of the reaction and promotes its rapid progression in one of many ways. On a basic level, enzymes promote chemical reactions that involve more than one substrate by bringing the substrates together in an optimal orientation. The appropriate region (atoms and bonds) of one molecule is juxtaposed to the appropriate region of the other molecule with which it must react. Another way in which enzymes promote the reaction of their substrates is by creating an optimal environment within the active site for the reaction to occur. Certain chemical reactions might proceed best in a slightly acidic or non-polar environment. The chemical properties that emerge from the particular arrangement of amino acid residues within an active site create the perfect environment for an enzyme's specific substrates to react.

You've learned that the activation energy required for many reactions includes the energy involved in manipulating or slightly contorting chemical bonds so that they can easily break and allow others to reform. Enzymatic action can aid this process. The enzyme-substrate complex can lower the activation energy by contorting substrate molecules in such a way as to facilitate bond-breaking, helping to reach the transition state. Finally, enzymes can also lower activation energies by taking part in the chemical reaction itself. The amino acid residues can provide certain ions or chemical groups that actually form covalent bonds with substrate molecules as a necessary step of the reaction process. In these cases, it is important to remember that the enzyme will always return to its original state at the completion of the reaction. One of the hallmark properties of enzymes is

that they remain ultimately unchanged by the reactions they catalyze. After an enzyme is done catalyzing a reaction, it releases its product(s).



According to the induced-fit model, both enzyme and substrate undergo dynamic conformational changes upon binding. The enzyme contorts the substrate into its transition state, thereby increasing the rate of the reaction.

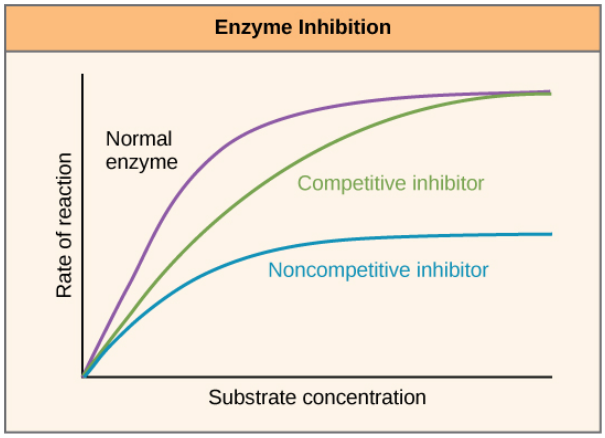
Control of Metabolism Through Enzyme Regulation

It would seem ideal to have a scenario in which all of the enzymes encoded in an organism's genome existed in abundant supply and functioned optimally under all cellular conditions, in all cells, at all times. In reality, this is far from the case. A variety of mechanisms ensure that this does not happen. Cellular needs and conditions vary from cell to cell, and change within individual cells over time. The required enzymes and energetic demands of stomach cells are different from those of fat storage cells, skin cells, blood cells, and nerve cells. Furthermore, a digestive cell works much harder to process and break down nutrients during the time that closely follows a meal compared with many hours after a meal. As these cellular demands and conditions vary, so do the amounts and functionality of different enzymes.

Since the rates of biochemical reactions are controlled by activation energy, and enzymes lower and determine activation energies for chemical reactions, the relative amounts and functioning of the variety of enzymes within a cell ultimately determine which reactions will proceed and at which rates. This determination is tightly controlled. In certain cellular environments, enzyme activity is partly controlled by environmental factors, like pH and temperature. There are other mechanisms through which cells control the activity of enzymes and determine the rates at which various biochemical reactions will occur.

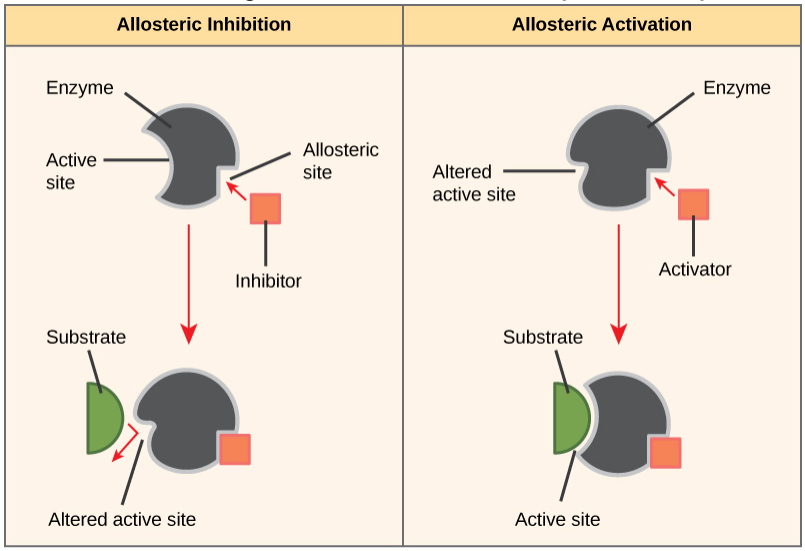
Regulation of Enzymes by Molecules

Enzymes can be regulated in ways that either promote or reduce their activity. There are many different kinds of molecules that inhibit or promote enzyme function, and various mechanisms exist for doing so. In some cases of enzyme inhibition, for example, an inhibitor molecule is similar enough to a substrate that it can bind to the active site and simply block the substrate from binding. When this happens, the enzyme is inhibited through competitive inhibition, because an inhibitor molecule competes with the substrate for active site binding (Figure). On the other hand, in noncompetitive inhibition, an inhibitor molecule binds to the enzyme in a location other than an allosteric site and still manages to block substrate binding to the active site.



Competitive and noncompetitive inhibition affect the rate of reaction differently. Competitive inhibitors affect the initial rate but do not affect the maximal rate, whereas noncompetitive inhibitors affect the maximal rate.

Some inhibitor molecules bind to enzymes in a location where their binding induces a conformational change that reduces the affinity of the enzyme for its substrate. This type of inhibition is called allosteric inhibition (Figure). Most allosterically regulated enzymes are made up of more than one polypeptide, meaning that they have more than one protein subunit. When an allosteric inhibitor binds to an enzyme, all active sites on the protein subunits are changed slightly such that they bind their substrates with less efficiency. There are allosteric activators as well as inhibitors. Allosteric activators bind to locations on an enzyme away from the active site, inducing a conformational change that increases the affinity of the enzyme's active site(s) for its substrate(s).



Allosteric inhibitors modify the active site of the enzyme so that substrate binding is reduced or prevented. In contrast, allosteric activators modify the active site of the enzyme so that the affinity for the substrate increases.
EVERYDAY CONNECTION

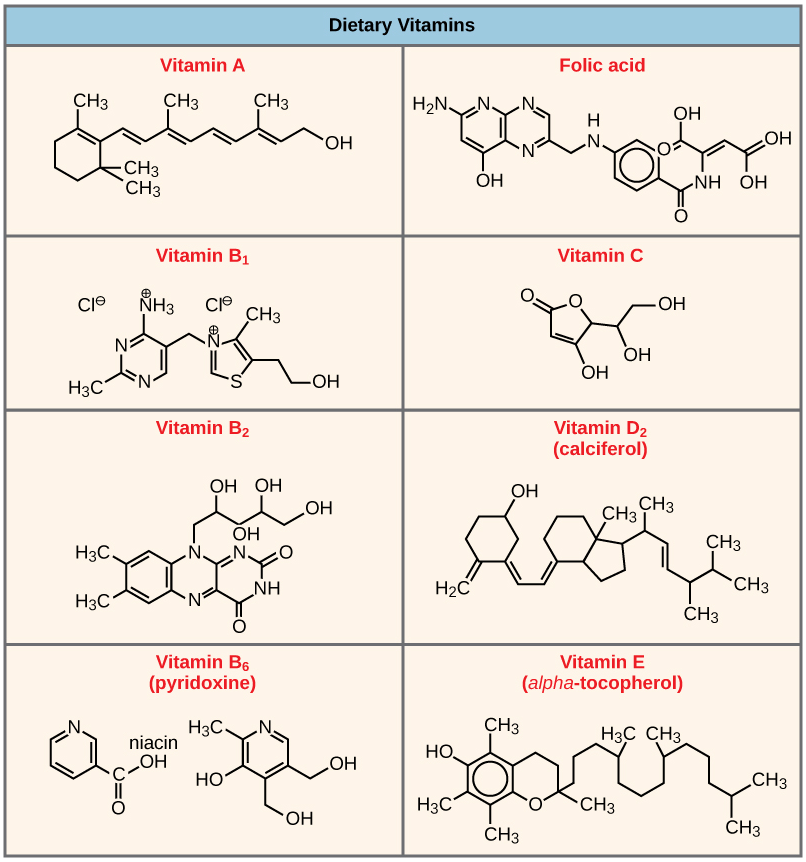
Have you ever wondered how pharmaceutical drugs are developed? (credit: Deborah Austin)

Drug Discovery by Looking for Inhibitors of Key Enzymes in Specific PathwaysEnzymes are key components of metabolic pathways. Understanding how enzymes work and how they can be regulated is a key principle behind the development of many of the pharmaceutical drugs (Figure) on the market today. Biologists working in this field collaborate with other scientists, usually chemists, to design drugs.

Consider statins for example—which is the name given to the class of drugs that reduces cholesterol levels. These compounds are essentially inhibitors of the enzyme HMG-CoA reductase. HMG-CoA reductase is the enzyme that synthesizes cholesterol from lipids in the body. By inhibiting this enzyme, the levels of cholesterol synthesized in the body can be reduced. Similarly, acetaminophen, popularly marketed under the brand name Tylenol, is an inhibitor of the enzyme cyclooxygenase. While it is effective in providing relief from fever and inflammation (pain), its mechanism of action is still not completely understood.

How are drugs developed? One of the first challenges in drug development is identifying the specific molecule that the drug is intended to target. In the case of statins, HMG-CoA reductase is the drug target. Drug targets are identified through painstaking research in the laboratory. Identifying the target alone is not sufficient; scientists also need to know how the target acts inside the cell and which reactions go awry in the case of disease. Once the target and the pathway are identified, then the actual process of drug design begins. During this stage, chemists and biologists work together to design and synthesize molecules that can either block or activate a particular reaction. However, this is only the beginning: both if and when a drug prototype is successful in performing its function, then it must undergo many tests from in vitro experiments to clinical trials before it can get FDA approval to be on the market.

Many enzymes don’t work optimally, or even at all, unless bound to other specific non-protein helper molecules, either temporarily through ionic or hydrogen bonds or permanently through stronger covalent bonds. Two types of helper molecules arecofactors and coenzymes. Binding to these molecules promotes optimal conformation and function for their respective enzymes. Cofactors are inorganic ions such as iron (Fe++) and magnesium (Mg++). One example of an enzyme that requires a metal ion as a cofactor is the enzyme that builds DNA molecules, DNA polymerase, which requires bound zinc ion (Zn++) to function. Coenzymes are organic helper molecules, with a basic atomic structure made up of carbon and hydrogen, which are required for enzyme action. The most common sources of coenzymes are dietary vitamins (Figure). Some vitamins are precursors to coenzymes and others act directly as coenzymes. Vitamin C is a coenzyme for multiple enzymes that take part in building the important connective tissue component, collagen. An important step in the breakdown of glucose to yield energy is catalysis by a multi-enzyme complex called pyruvate dehydrogenase. Pyruvate dehydrogenase is a complex of several enzymes that actually requires one cofactor (a magnesium ion) and five different organic coenzymes to catalyze its specific chemical reaction. Therefore, enzyme function is, in part, regulated by an abundance of various cofactors and coenzymes, which are supplied primarily by the diets of most organisms.



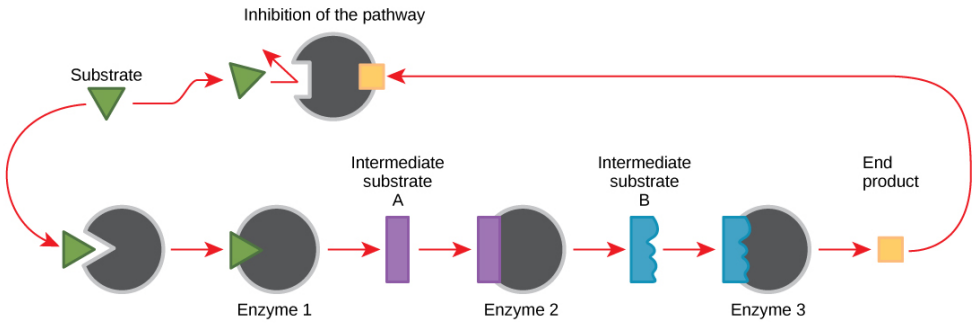
Vitamins are important coenzymes or precursors of coenzymes, and are required for enzymes to function properly. Multivitamin capsules usually contain mixtures of all the vitamins at different percentages.

Enzyme Compartmentalization

In eukaryotic cells, molecules such as enzymes are usually compartmentalized into different organelles. This allows for yet another level of regulation of enzyme activity. Enzymes required only for certain cellular processes can be housed separately along with their substrates, allowing for more efficient chemical reactions. Examples of this sort of enzyme regulation based on location and proximity include the enzymes involved in the latter stages of cellular respiration, which take place exclusively in the mitochondria, and the enzymes involved in the digestion of cellular debris and foreign materials, located within lysosomes.

Feedback Inhibition in Metabolic Pathways

Molecules can regulate enzyme function in many ways. A major question remains, however: What are these molecules and where do they come from? Some are cofactors and coenzymes, ions, and organic molecules, as you've learned. What other molecules in the cell provide enzymatic regulation, such as allosteric modulation, and competitive and noncompetitive inhibition? The answer is that a wide variety of molecules can perform these roles. Some of these molecules include pharmaceutical and non-pharmaceutical drugs, toxins, and poisons from the environment. Perhaps the most relevant sources of enzyme regulatory molecules, with respect to cellular metabolism, are the products of the cellular metabolic reactions themselves. In a most efficient and elegant way, cells have evolved to use the products of their own reactions for feedback inhibition of enzyme activity.Feedback inhibition involves the use of a reaction product to regulate its own further production (Figure). The cell responds to the abundance of specific products by slowing down production during anabolic or catabolic reactions. Such reaction products may inhibit the enzymes that catalyzed their production through the mechanisms described above.



Metabolic pathways are a series of reactions catalyzed by multiple enzymes. Feedback inhibition, where the end product of the pathway inhibits an upstream step, is an important regulatory mechanism in cells.

The production of both amino acids and nucleotides is controlled through feedback inhibition. Additionally, ATP is an allosteric regulator of some of the enzymes involved in the catabolic breakdown of sugar, the process that produces ATP. In this way, when ATP is abundant, the cell can prevent its further production. Remember that ATP is an unstable molecule that can spontaneously dissociate into ADP. If too much ATP were present in a cell, much of it would go to waste. On the other hand, ADP serves as a positive allosteric regulator (an allosteric activator) for some of the same enzymes that are inhibited by ATP. Thus, when relative levels of ADP are high compared to ATP, the cell is triggered to produce more ATP through the catabolism of sugar.

Section Summary

Enzymes are chemical catalysts that accelerate chemical reactions at physiological temperatures by lowering their activation energy. Enzymes are usually proteins consisting of one or more polypeptide chains. Enzymes have an active site that provides a unique chemical environment, made up of certain amino acid R groups (residues). This unique environment is perfectly suited to convert particular chemical reactants for that enzyme, called substrates, into unstable intermediates called transition states. Enzymes and substrates are thought to bind with an induced fit, which means that enzymes undergo slight conformational adjustments upon substrate contact, leading to full, optimal binding. Enzymes bind to substrates and catalyze reactions in four different ways: bringing substrates together in an optimal orientation, compromising the bond structures of substrates so that bonds can be more easily broken, providing optimal environmental conditions for a reaction to occur, or participating directly in their chemical reaction by forming transient covalent bonds with the substrates.

Enzyme action must be regulated so that in a given cell at a given time, the desired reactions are being catalyzed and the undesired reactions are not. Enzymes are regulated by cellular conditions, such as temperature and pH. They are also regulated through their location within a cell, sometimes being compartmentalized so that they can only catalyze reactions under certain circumstances. Inhibition and activation of enzymes via other molecules are other important ways that enzymes are regulated. Inhibitors can act competitively, noncompetitively, or allosterically; noncompetitive inhibitors are usually allosteric. Activators can also enhance the function of enzymes allosterically. The most common method by which cells regulate the enzymes in metabolic pathways is through feedback inhibition. During feedback inhibition, the products of a metabolic pathway serve as inhibitors (usually allosteric) of one or more of the enzymes (usually the first committed enzyme of the pathway) involved in the pathway that produces them.

Licensing & Attributions

CC licensed content, Shared previously

- enzymes. **Authored by:** open stax. **Provided by:** open stax. **Located at:** http://cnx.org/contents/GFy_h8cu@10.59:MnC6GuJi@7/Enzymes. **License:** *CC BY: Attribution*

STUDY GUIDE: ENERGY

Study Questions

Objective: Define energy and describe applications to living systems.

Use this page to check your understanding of the content.

Vocabulary

1. Energy
2. Work
3. Calorie
4. calorie
5. Activation energy
6. Kinetic energy
7. Potential energy
8. Chemical energy
9. Heat

Study Guide Questions

1. Compare and contrast several different forms of energy.
2. Understand the energetic dynamics of chemical bonds. In other words, know whether energy is USED UP or RELEASED when chemical bonds break and form.
3. Define “activation energy” and explain how enzymes affect activation energy.
4. What is an enzyme? How are enzymes related to energy in the body?
5. What is the first law of thermodynamics? Be able to describe an example illustrating the law.
6. What is the second law of thermodynamics? Be able to describe an example illustrating the law.
7. Be able to trace the source of all energy in living systems. Your answer should include the source of energy in the sun.
8. Know the structure of ATP. Be able to explain HOW ATP can function as a source of energy in the body.

Licensing & Attributions

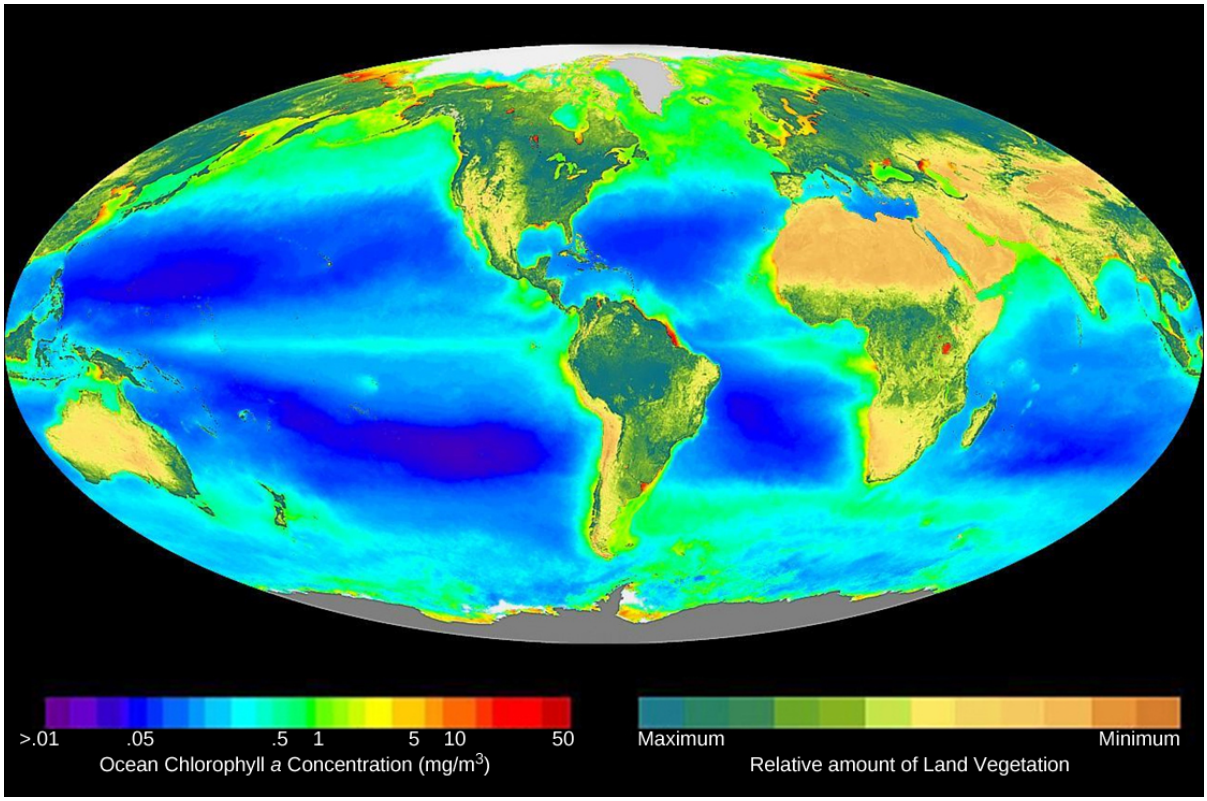
CC licensed content, Shared previously

- Study Guide: Energy. **Authored by**: Wendy Riggs. **Provided by**: College of the Redwoods. **Project**: Kaleidoscope. **License**: *CC BY: Attribution*

PHOTOSYNTHESIS

INTRODUCTION

Chapter 8. Introduction



This world map shows Earth's distribution of photosynthesis as seen via chlorophyll *a* concentrations. On land, this is evident via terrestrial plants, and in oceanic zones, via phytoplankton. (credit: modification of work by SeaWiFS Project, NASA/Goddard Space Flight Center and ORBIMAGE)

The processes in all organisms—from bacteria to humans—require energy. To get this energy, many organisms access stored energy by eating, that is, by ingesting other organisms. But where does the stored energy in food originate? All of this energy can be traced back to photosynthesis.

Licensing & Attributions

CC licensed content, Shared previously

- introduction. **Authored by**: open stax. **Provided by**: open stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:SCQoV1nR@2/Introduction. **License**: *CC BY: Attribution*

OVERVIEW OF PHOTOSYNTHESIS

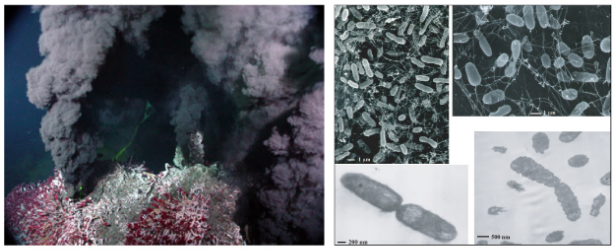
8.1 Overview of Photosynthesis

Summary

Photosynthesis is essential to all life on earth; both plants and animals depend on it. It is the only biological process that can capture energy that originates in outer space (sunlight) and convert it into chemical compounds (carbohydrates) that every organism uses to power its metabolism. In brief, the energy of sunlight is captured and used to energize electrons, which are then stored in the covalent bonds of sugar molecules. How long lasting and stable are those covalent bonds? The energy extracted today by the burning of coal and petroleum products represents sunlight energy captured and stored by photosynthesis almost 200 million years ago.

Plants, algae, and a group of bacteria called cyanobacteria are the only organisms capable of performing photosynthesis (Figure). Because they use light to manufacture their own food, they are called photoautotrophs (literally, “self-feeders using light”). Other organisms, such as animals, fungi, and most other bacteria, are termed heterotrophs (“other feeders”), because they must rely on the sugars produced by photosynthetic organisms for their energy needs. A third very interesting group of bacteria synthesize sugars, not by using sunlight’s energy, but by extracting energy from inorganic chemical compounds; hence, they are referred to as chemoautotrophs.

(a) (b) (c)



(d) (e)

Photoautotrophs including (a) plants, (b) algae, and (c) cyanobacteria synthesize their organic compounds via photosynthesis using sunlight as an energy source. Cyanobacteria and planktonic algae can grow over enormous areas in water, at times completely covering the surface. In a (d) deep sea vent, chemoautotrophs, such as these (e) thermophilic bacteria, capture energy from inorganic compounds to produce organic compounds. The ecosystem surrounding the vents has a diverse array of animals, such as tubeworms, crustaceans, and octopi that derive energy from the bacteria. (credit a: modification of work by Steve Hillebrand, U.S. Fish and Wildlife Service; credit b: modification of work by “eutrophication&hypoxia”/Flickr; credit c: modification of work by NASA; credit d: University of Washington, NOAA; credit e: modification of work by Mark Amend, West Coast and Polar Regions Undersea Research Center, UAF, NOAA)

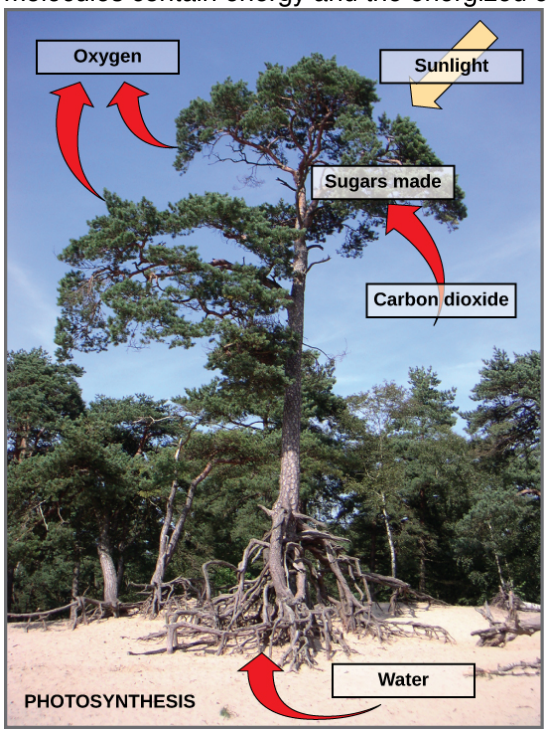
The importance of photosynthesis is not just that it can capture sunlight’s energy. A lizard sunning itself on a cold day can use the sun’s energy to warm up. Photosynthesis is vital because it evolved as a way to store the energy in solar radiation (the “photo-” part) as high-energy electrons in the carbon-carbon bonds of carbohydrate molecules (the “-synthesis” part). Those carbohydrates are the energy source that heterotrophs use to power the synthesis of ATP via respiration. Therefore, photosynthesis powers 99 percent of Earth’s ecosystems. When a top

predator, such as a wolf, preys on a deer (Figure), the wolf is at the end of an energy path that went from nuclear reactions on the surface of the sun, to light, to photosynthesis, to vegetation, to deer, and finally to wolf.

The energy stored in carbohydrate molecules from photosynthesis passes through the food chain. The predator that eats these deer receives a portion of the energy that originated in the photosynthetic vegetation that the deer consumed. (credit: modification of work by Steve VanRiper, U.S. Fish and Wildlife Service)

Main Structures and Summary of Photosynthesis

Photosynthesis is a multi-step process that requires sunlight, carbon dioxide (which is low in energy), and water as substrates (Figure). After the process is complete, it releases oxygen and produces glyceraldehyde-3-phosphate (GA3P), simple carbohydrate molecules (which are high in energy) that can subsequently be converted into glucose, sucrose, or any of dozens of other sugar molecules. These sugar molecules contain energy and the energized carbon that all living things need to survive.



Photosynthesis uses solar energy, carbon dioxide, and water to produce energy-storing carbohydrates. Oxygen is generated as a waste product of photosynthesis.

The following is the chemical equation for photosynthesis (Figure):

Photosynthesis Equation							
Carbon dioxide	+	Water	SUNLIGHT →	Sugar	+	Oxygen	
$6CO_2$		$6H_2O$		$C_6H_{12}O_6$		$6O_2$	

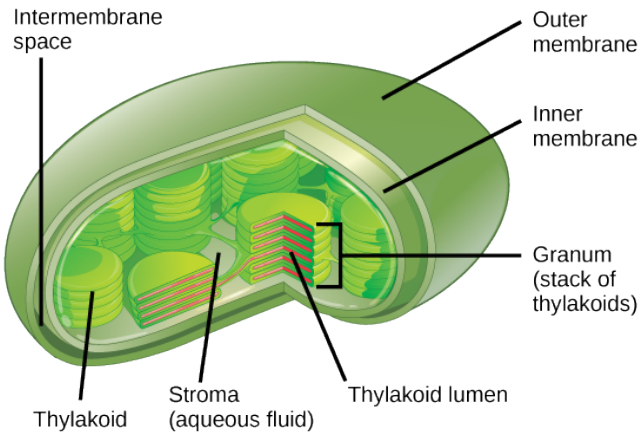
The basic equation for photosynthesis is deceptively simple. In reality, the process takes place in many steps involving intermediate reactants and products. Glucose, the primary energy source in cells, is made from two three-carbon GA3Ps.

Although the equation looks simple, the many steps that take place during photosynthesis are actually quite complex. Before learning the details of how photoautotrophs turn sunlight into food, it is important to become familiar with the structures involved.

In plants, photosynthesis generally takes place in leaves, which consist of several layers of cells. The process of photosynthesis occurs in a middle layer called the mesophyll. The gas exchange of carbon dioxide and oxygen occurs through small, regulated openings called stomata (singular: stoma), which also play roles in the regulation of gas exchange and water balance. The stomata are typically located on the underside of the leaf, which helps to minimize water loss. Each stoma is flanked by guard cells that regulate the opening and closing of the stomata by swelling or shrinking in response to osmotic changes.

In all autotrophic eukaryotes, photosynthesis takes place inside an organelle called a chloroplast. For plants, chloroplast-containing cells exist in the mesophyll. Chloroplasts have a double membrane envelope (composed of an outer membrane and an inner membrane). Within the chloroplast are stacked, disc-shaped structures called thylakoids. Embedded in the thylakoid membrane is chlorophyll, a pigment (molecule that absorbs light) responsible for the initial interaction between light and plant material, and numerous proteins that make up the electron transport chain. The thylakoid membrane encloses an internal space called the thylakoid lumen. As shown in Figure, a stack of thylakoids is called a granum, and the liquid-filled space surrounding the granum is called stroma or “bed” (not to be confused with stoma or “mouth,” an opening on the leaf epidermis).

ART CONNECTION



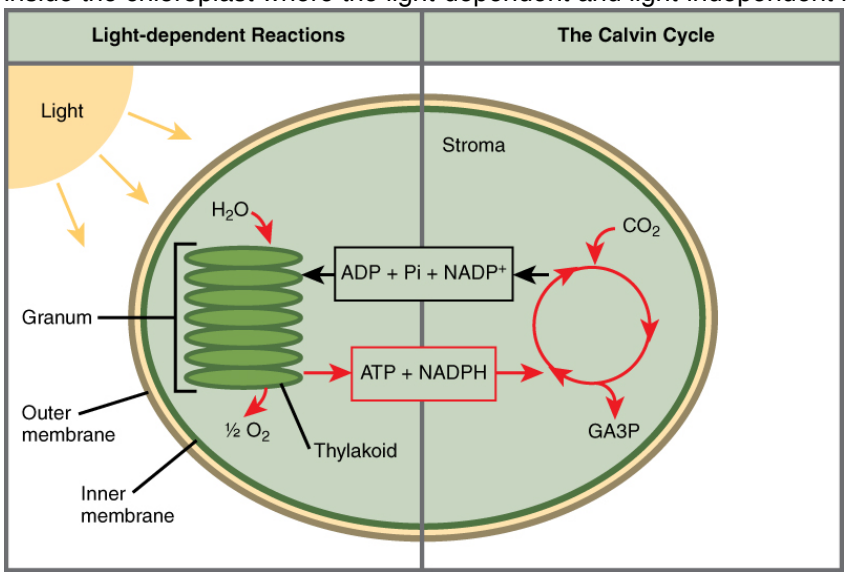
Photosynthesis takes place in chloroplasts, which have an outer membrane and an inner membrane. Stacks of thylakoids called grana form a third membrane layer.

On a hot, dry day, plants close their stomata to conserve water. What impact will this have on photosynthesis?

The Two Parts of Photosynthesis

Photosynthesis takes place in two sequential stages: the light-dependent reactions and the light independent-reactions. In thelight-dependent reactions, energy from sunlight is absorbed by chlorophyll and that energy is converted into stored chemical energy. In the light-independent reactions, the chemical energy harvested during the light-dependent reactions drive the assembly of sugar molecules from carbon dioxide. Therefore, although the light-independent reactions do not use light as a reactant, they require the products of the light-dependent reactions to function. In addition, several enzymes of the light-independent reactions are activated by light. The light-dependent reactions utilize certain molecules to temporarily store the energy: These are referred to as

energy carriers. The energy carriers that move energy from light-dependent reactions to light-independent reactions can be thought of as “full” because they are rich in energy. After the energy is released, the “empty” energy carriers return to the light-dependent reaction to obtain more energy. Figure illustrates the components inside the chloroplast where the light-dependent and light-independent reactions take place.

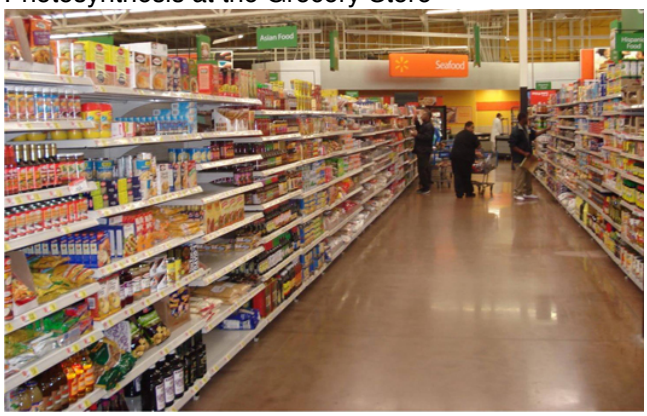


Photosynthesis takes place in two stages: light dependent reactions and the Calvin cycle. Light-dependent reactions, which take place in the thylakoid membrane, use light energy to make ATP and NADPH. The Calvin cycle, which takes place in the stroma, uses energy derived from these compounds to make GA3P from CO_2.

Link to Learning

Click the link to learn more about photosynthesis.
EVERYDAY CONNECTION

Photosynthesis at the Grocery Store



Foods that humans consume originate from photosynthesis. (credit: Associação Brasileira de Supermercados)

Major grocery stores in the United States are organized into departments, such as dairy, meats, produce, bread, cereals, and so forth. Each aisle (Figure) contains hundreds, if not thousands, of different products for customers to buy and consume.

Although there is a large variety, each item links back to photosynthesis. Meats and dairy link, because the animals were fed plant-based foods. The breads, cereals, and pastas come largely from starchy grains, which are the seeds of photosynthesis-dependent plants. What about desserts and drinks? All of these products contain sugar—sucrose is a plant product, a disaccharide, a carbohydrate molecule, which is built directly from photosynthesis. Moreover, many items are less obviously derived from plants: For instance, paper goods are generally plant products, and many plastics (abundant as products and packaging) are derived from algae.

Virtually every spice and flavoring in the spice aisle was produced by a plant as a leaf, root, bark, flower, fruit, or stem. Ultimately, photosynthesis connects to every meal and every food a person consumes.

Section Summary

The process of photosynthesis transformed life on Earth. By harnessing energy from the sun, photosynthesis evolved to allow living things access to enormous amounts of energy. Because of photosynthesis, living things gained access to sufficient energy that allowed them to build new structures and achieve the biodiversity evident today.

Only certain organisms, called photoautotrophs, can perform photosynthesis; they require the presence of chlorophyll, a specialized pigment that absorbs certain portions of the visible spectrum and can capture energy from sunlight. Photosynthesis uses carbon dioxide and water to assemble carbohydrate molecules and release oxygen as a waste product into the atmosphere. Eukaryotic autotrophs, such as plants and algae, have organelles called chloroplasts in which photosynthesis takes place, and starch accumulates. In prokaryotes, such as cyanobacteria, the process is less localized and occurs within folded membranes, extensions of the plasma membrane, and in the cytoplasm.

Art Connections

Figure On a hot, dry day, plants close their stomata to conserve water. What impact will this have on photosynthesis?

Review Questions

Which of the following components is *not* used by both plants and cyanobacteria to carry out photosynthesis?

1. chloroplasts
2. chlorophyll
3. carbon dioxide
4. water

What two main products result from photosynthesis?

1. oxygen and carbon dioxide
2. chlorophyll and oxygen
3. sugars/carbohydrates and oxygen
4. sugars/carbohydrates and carbon dioxide

In which compartment of the plant cell do the light-independent reactions of photosynthesis take place?

1. thylakoid
2. stroma
3. outer membrane
4. mesophyll

Which statement about thylakoids in eukaryotes is *not* correct?

1. Thylakoids are assembled into stacks.
2. Thylakoids exist as a maze of folded membranes.
3. The space surrounding thylakoids is called stroma.
4. Thylakoids contain chlorophyll.

Free Response

What is the overall outcome of the light reactions in photosynthesis?

Why are carnivores, such as lions, dependent on photosynthesis to survive?

Why are energy carriers thought of as either “full” or “empty”?

Glossary

chemoautotroph organism that can build organic molecules using energy derived from inorganic chemicals instead of sunlight

chloroplast organelle in which photosynthesis takes place

granum stack of thylakoids located inside a chloroplast

heterotroph organism that consumes organic substances or other organisms for food

light-dependent reaction first stage of photosynthesis where certain wavelengths of the visible light are absorbed to form two energy-carrying molecules (ATP and NADPH)

light-independent reaction second stage of photosynthesis, though which carbon dioxide is used to build carbohydrate molecules using energy from ATP and NADPH

mesophyll middle layer of chlorophyll-rich cells in a leaf

photoautotroph organism capable of producing its own organic compounds from sunlight

pigment molecule that is capable of absorbing certain wavelengths of light and reflecting others (which accounts for its color)

stoma opening that regulates gas exchange and water evaporation between leaves and the environment, typically situated on the underside of leaves

stroma fluid-filled space surrounding the grana inside a chloroplast where the light-independent reactions of photosynthesis take place

thylakoid disc-shaped, membrane-bound structure inside a chloroplast where the light-dependent reactions of photosynthesis take place; stacks of thylakoids are called grana

thylakoid lumen aqueous space bound by a thylakoid membrane where protons accumulate during light-driven electron transport

Licensing & Attributions

CC licensed content, Shared previously

- overview of photosynthesis. **Authored by**: open stax. **Provided by**: open stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:W7ctJeSl@8/Overview-of-Photosynthesis. **License**: *CC BY: Attribution*

USING LIGHT TO MAKE ORGANIC MOLECULES

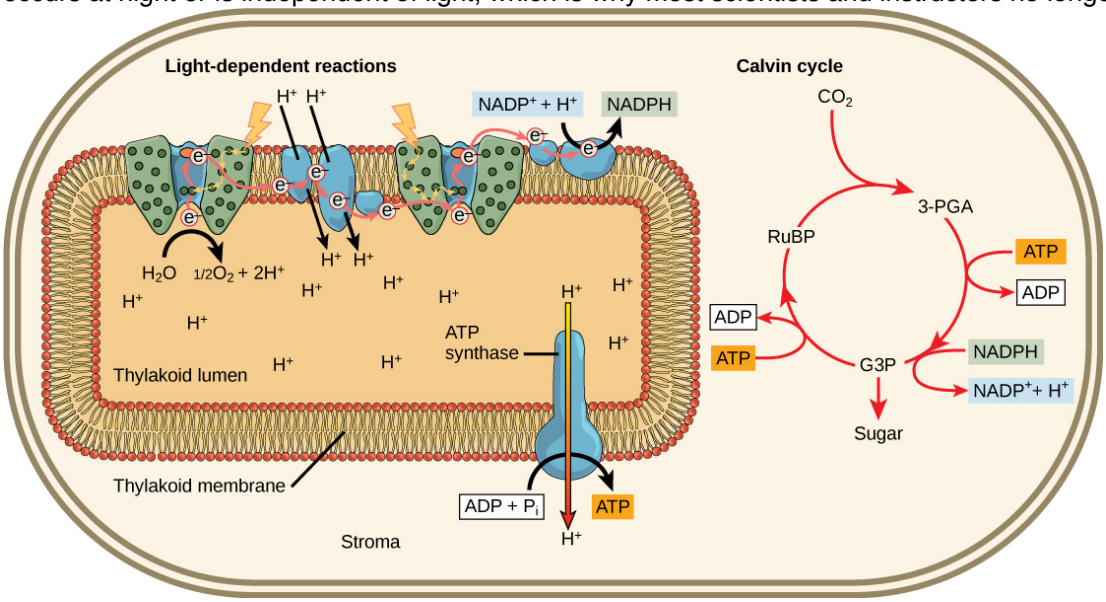
8.3 Using Light Energy to Make Organic Molecules

Summary

After the energy from the sun is converted into chemical energy and temporarily stored in ATP and NADPH molecules, the cell has the fuel needed to build carbohydrate molecules for long-term energy storage. The products of the light-dependent reactions, ATP and NADPH, have lifespans in the range of millionths of seconds, whereas the products of the light-independent reactions (carbohydrates and other forms of reduced carbon) can survive for hundreds of millions of years. The carbohydrate molecules made will have a backbone of carbon atoms. Where does the carbon come from? It comes from carbon dioxide, the gas that is a waste product of respiration in microbes, fungi, plants, and animals.

The Calvin Cycle

In plants, carbon dioxide (CO_2) enters the leaves through stomata, where it diffuses over short distances through intercellular spaces until it reaches the mesophyll cells. Once in the mesophyll cells, CO_2 diffuses into the stroma of the chloroplast—the site of light-independent reactions of photosynthesis. These reactions actually have several names associated with them. Another term, the Calvin cycle, is named for the man who discovered it, and because these reactions function as a cycle. Others call it the Calvin-Benson cycle to include the name of another scientist involved in its discovery. The most outdated name is dark reactions, because light is not directly required (Figure). However, the term dark reaction can be misleading because it implies incorrectly that the reaction only occurs at night or is independent of light, which is why most scientists and instructors no longer use it.



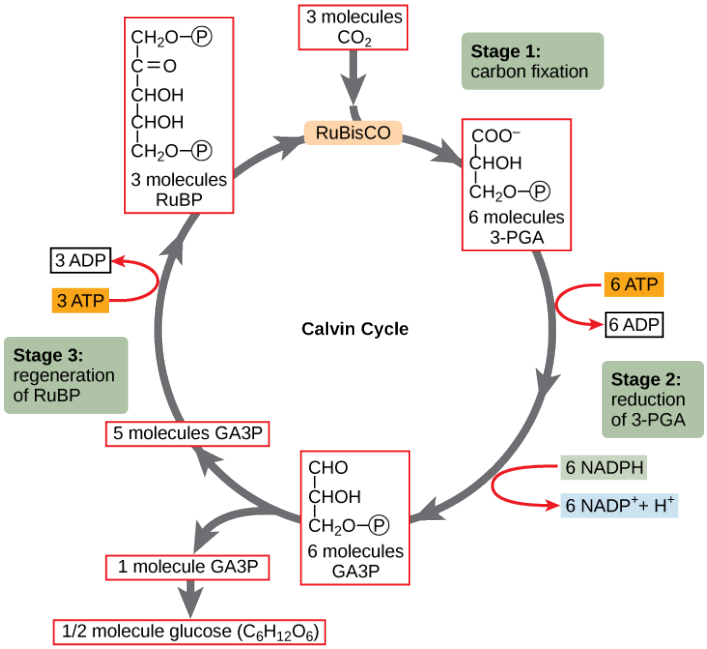
Light reactions harness energy from the sun to produce chemical bonds, ATP, and NADPH. These energy-carrying molecules are made in the stroma where carbon fixation takes place.

The light-independent reactions of the Calvin cycle can be organized into three basic stages: fixation, reduction, and regeneration.

Stage 1: Fixation

In the stroma, in addition to CO_2, two other components are present to initiate the light-independent reactions: an enzyme called ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO), and three molecules of ribulose bisphosphate (RuBP), as shown in Figure. RuBP has five atoms of carbon, flanked by two phosphates.

ART CONNECTION



The Calvin cycle has three stages. In stage 1, the enzyme RuBisCO incorporates carbon dioxide into an organic molecule, 3-PGA. In stage 2, the organic molecule is reduced using electrons supplied by NADPH. In stage 3, RuBP, the molecule that starts the cycle, is regenerated so that the cycle can continue. Only one carbon dioxide molecule is incorporated at a time, so the cycle must be completed three times to produce a single three-carbon GA3P molecule, and six times to produce a six-carbon glucose molecule.

Which of the following statements is true?

1. In photosynthesis, oxygen, carbon dioxide, ATP, and NADPH are reactants. GA3P and water are products.
2. In photosynthesis, chlorophyll, water, and carbon dioxide are reactants. GA3P and oxygen are products.
3. In photosynthesis, water, carbon dioxide, ATP, and NADPH are reactants. RuBP and oxygen are products.
4. In photosynthesis, water and carbon dioxide are reactants. GA3P and oxygen are products.

RuBisCO catalyzes a reaction between CO_2 and RuBP. For each CO_2 molecule that reacts with one RuBP, two molecules of another compound (3-PGA) form. PGA has three carbons and one phosphate. Each turn of the cycle involves only one RuBP and one carbon dioxide and forms two molecules of 3-PGA. The number of carbon atoms remains the same, as the atoms move to form new bonds during the reactions (3 atoms from $3CO_2$ + 15 atoms from 3RuBP = 18 atoms in 3 atoms of 3-PGA). This process is called carbon fixation, because CO_2 is "fixed" from an inorganic form into organic molecules.

Stage 2: Reduction

ATP and NADPH are used to convert the six molecules of 3-PGA into six molecules of a chemical called glyceraldehyde 3-phosphate (G3P). That is a reduction reaction because it involves the gain of electrons by 3-PGA. Recall that a reduction is the gain of an electron by an atom or molecule. Six molecules of both ATP and NADPH are used. For ATP, energy is released with the loss of the terminal phosphate atom, converting it into ADP; for NADPH, both energy and a hydrogen atom are lost, converting it into $NADP^+$. Both of these molecules return to the nearby light-dependent reactions to be reused and reenergized.

Stage 3: Regeneration

Interestingly, at this point, only one of the G3P molecules leaves the Calvin cycle and is sent to the cytoplasm to contribute to the formation of other compounds needed by the plant. Because the G3P exported from the chloroplast has three carbon atoms, it takes three “turns” of the Calvin cycle to fix enough net carbon to export one G3P. But each turn makes two G3Ps, thus three turns make six G3Ps. One is exported while the remaining five G3P molecules remain in the cycle and are used to regenerate RuBP, which enables the system to prepare for more CO_2 to be fixed. Three more molecules of ATP are used in these regeneration reactions.

Link to Learning

This link leads to an animation of the Calvin cycle. Click stage 1, stage 2, and then stage 3 to see G3P and ATP regenerate to form RuBP.
EVOLUTION CONNECTION

PhotosynthesisDuring the evolution of photosynthesis, a major shift occurred from the bacterial type of photosynthesis that involves only one photosystem and is typically anoxygenic (does not generate oxygen) into modern oxygenic (does generate oxygen) photosynthesis, employing two photosystems. This modern oxygenic photosynthesis is used by many organisms—from giant tropical leaves in the rainforest to tiny cyanobacterial cells—and the process and components of this photosynthesis remain largely the same. Photosystems absorb light and use electron transport chains to convert energy into the chemical energy of ATP and NADH. The subsequent light-independent reactions then assemble carbohydrate molecules with this energy.

Photosynthesis in desert plants has evolved adaptations that conserve water. In the harsh dry heat, every drop of water must be used to survive. Because stomata must open to allow for the uptake of CO_2, water escapes from the leaf during active photosynthesis. Desert plants have evolved processes to conserve water and deal with harsh conditions. A more efficient use of CO_2 allows plants to adapt to living with less water. Some plants such as cacti (Figure) can prepare materials for photosynthesis during the night by a temporary carbon fixation/storage process, because opening the stomata at this time conserves water due to cooler temperatures. In addition, cacti have evolved the ability to carry out low levels of photosynthesis without opening stomata at all, an extreme mechanism to face extremely dry periods.

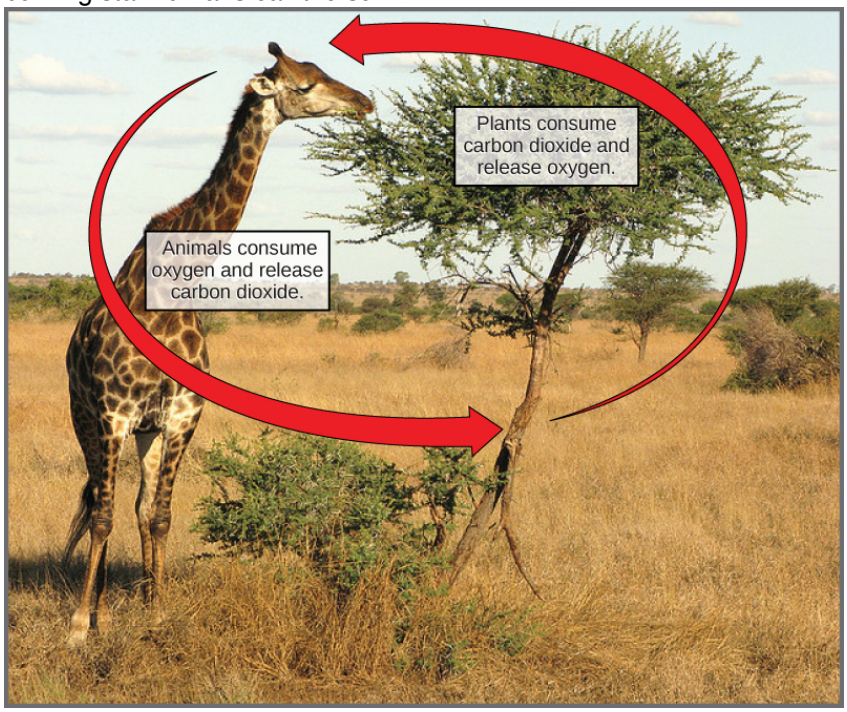
The harsh conditions of the desert have led plants like these cacti to evolve variations of the light-independent reactions of photosynthesis. These variations increase the efficiency of water usage, helping to conserve water and energy. (credit: Piotr Wojtkowski)

The Energy Cycle

Whether the organism is a bacterium, plant, or animal, all living things access energy by breaking down carbohydrate molecules. But if plants make carbohydrate molecules, why would they need to break them down, especially when it has been shown that the gas organisms release as a "waste product" (CO_2) acts as a substrate for the formation of more food in photosynthesis? Remember, living things need energy to perform life functions. In addition, an organism can either make its own food or eat another organism—either way, the food still needs to be broken down. Finally, in the process of breaking down food, called cellular respiration, heterotrophs release needed energy and produce "waste" in the form of CO_2 gas.

In nature, there is no such thing as waste. Every single atom of matter and energy is conserved, recycling over and over infinitely. Substances change form or move from one type of molecule to another, but their constituent atoms never disappear (Figure).

CO_2 is no more a form of waste than oxygen is wasteful to photosynthesis. Both are byproducts of reactions that move on to other reactions. Photosynthesis absorbs light energy to build carbohydrates in chloroplasts, and aerobic cellular respiration releases energy by using oxygen to metabolize carbohydrates in the cytoplasm and mitochondria. Both processes use electron transport chains to capture the energy necessary to drive other reactions. These two powerhouse processes, photosynthesis and cellular respiration, function in biological, cyclical harmony to allow organisms to access life-sustaining energy that originates millions of miles away in a burning star humans call the sun.



Photosynthesis consumes carbon dioxide and produces oxygen. Aerobic respiration consumes oxygen and produces carbon dioxide. These two processes play an important role in the carbon cycle. (credit: modification of work by Stuart Bassil)

Section Summary

Using the energy carriers formed in the first steps of photosynthesis, the light-independent reactions, or the Calvin cycle, take in CO_2 from the environment. An enzyme, RuBisCO, catalyzes a reaction with CO_2 and another molecule, RuBP. After three cycles, a three-carbon molecule of G3P leaves the cycle to become part of a carbohydrate molecule. The remaining G3P molecules stay in the cycle to be regenerated into RuBP, which is then ready to react with more CO_2. Photosynthesis forms an energy cycle with the process of cellular respiration.

Plants need both photosynthesis and respiration for their ability to function in both the light and dark, and to be able to interconvert essential metabolites. Therefore, plants contain both chloroplasts and mitochondria.

Art Connections

Figure Which of the following statements is true?

1. In photosynthesis, oxygen, carbon dioxide, ATP, and NADPH are reactants. G3P and water are products.
2. In photosynthesis, chlorophyll, water, and carbon dioxide are reactants. G3P and oxygen are products.
3. In photosynthesis, water, carbon dioxide, ATP, and NADPH are reactants. RuBP and oxygen are products.
4. In photosynthesis, water and carbon dioxide are reactants. G3P and oxygen are products.

Review Questions

Which molecule must enter the Calvin cycle continually for the light-independent reactions to take place?

1. RuBisCO
2. RuBP
3. 3-PGA
4. CO_2

Which order of molecular conversions is correct for the Calvin cycle?

1. RuBP + G3P → 3-PGA → sugarRuBP + G3P → 3-PGA → sugar
2. RuBisCO → CO2 → RuBP → G3PRuBisCO → CO2 → RuBP → G3P
3. RuBP + CO2 → [RuBisCO] 3-PGA → G3PRuBP + CO2 → [RuBisCO] 3-PGA → G3P
4. CO2 → 3-PGA → RuBP → G3PCO2 → 3-PGA → RuBP → G3P

Where in eukaryotic cells does the Calvin cycle take place?

1. thylakoid membrane
2. thylakoid lumen
3. chloroplast stroma
4. granum

Which statement correctly describes carbon fixation?

1. the conversion of CO_2 into an organic compound
2. the use of RuBisCO to form 3-PGA
3. the production of carbohydrate molecules from G3P
4. the formation of RuBP from G3P molecules
5. the use of ATP and NADPH to reduce CO_2

Free Response

Why is the third stage of the Calvin cycle called the regeneration stage?

Which part of the light-independent reactions would be affected if a cell could not produce the enzyme RuBisCO?

Why does it take three turns of the Calvin cycle to produce G3P, the initial product of photosynthesis?

Glossary

Calvin cycle light-independent reactions of photosynthesis that convert carbon dioxide from the atmosphere into carbohydrates using the energy and reducing power of ATP and NADPH

carbon fixation process of converting inorganic CO_2 gas into organic compounds

reduction gain of electron(s) by an atom or molecule

Licensing & Attributions

CC licensed content, Shared previously

- using light to make organic molecules. **Authored by**: open stax. **Provided by**: open stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:NrNGneVh@8/Using-Light-Energy-to-Make-Org. **License**: *CC BY: Attribution*

LIGHT DEPENDENT REACTIONS

8.2 The Light-Dependent Reactions of Photosynthesis

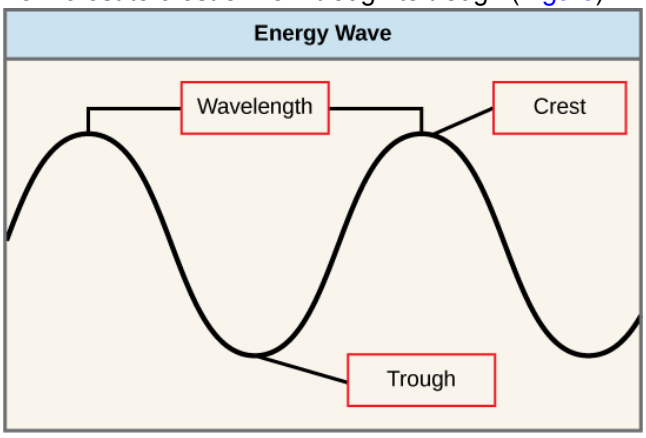
Summary

How can light be used to make food? When a person turns on a lamp, electrical energy becomes light energy. Like all other forms of kinetic energy, light can travel, change form, and be harnessed to do work. In the case of photosynthesis, light energy is converted into chemical energy, which photoautotrophs use to build carbohydrate molecules (Figure). However, autotrophs only use a few specific components of sunlight.

Photoautotrophs can capture light energy from the sun, converting it into the chemical energy used to build food molecules. (credit: Gerry Atwell)

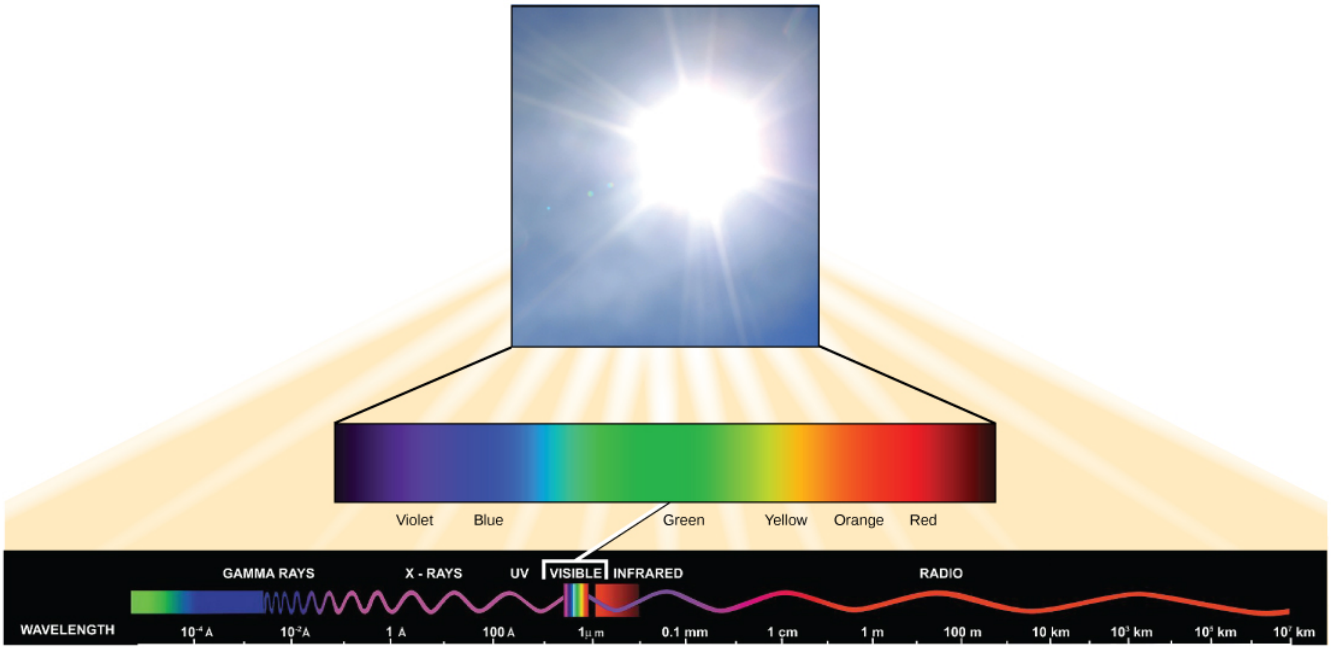
What Is Light Energy?

The sun emits an enormous amount of electromagnetic radiation (solar energy). Humans can see only a fraction of this energy, which portion is therefore referred to as “visible light.” The manner in which solar energy travels is described as waves. Scientists can determine the amount of energy of a wave by measuring its wavelength, the distance between consecutive points of a wave. A single wave is measured from two consecutive points, such as from crest to crest or from trough to trough (Figure).



The wavelength of a single wave is the distance between two consecutive points of similar position (two crests or two troughs) along the wave.

Visible light constitutes only one of many types of electromagnetic radiation emitted from the sun and other stars. Scientists differentiate the various types of radiant energy from the sun within the electromagnetic spectrum. The electromagnetic spectrum is the range of all possible frequencies of radiation (Figure). The difference between wavelengths relates to the amount of energy carried by them.



The sun emits energy in the form of electromagnetic radiation. This radiation exists at different wavelengths, each of which has its own characteristic energy. All electromagnetic radiation, including visible light, is characterized by its wavelength.

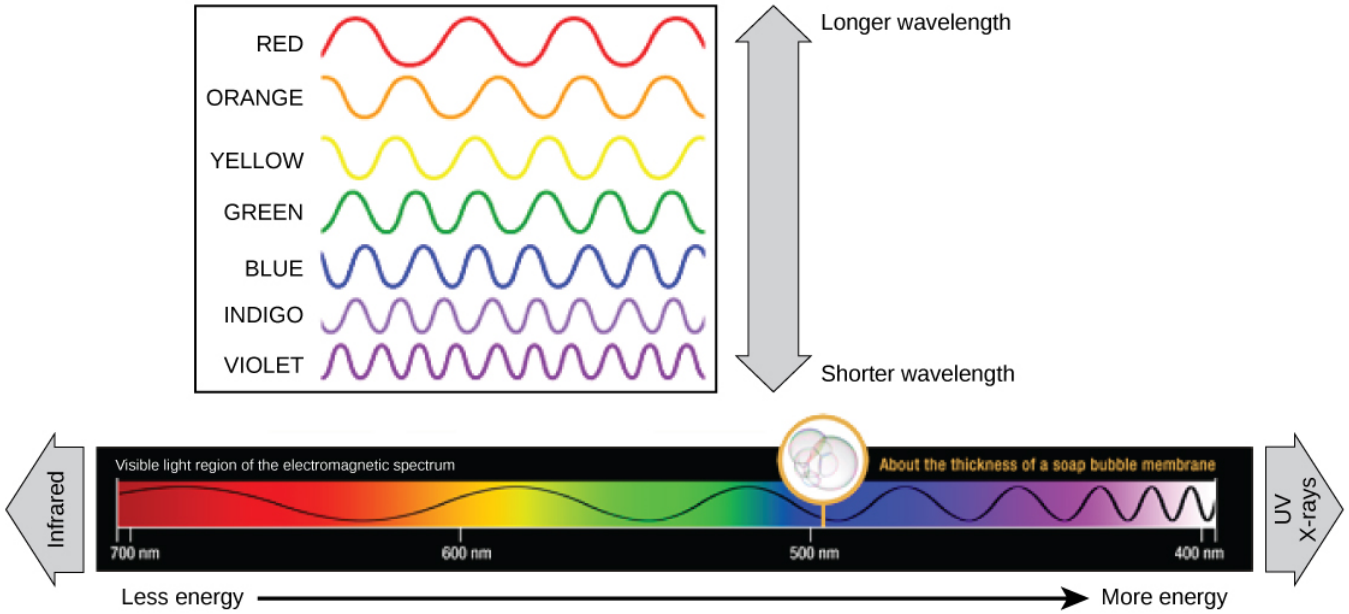
Each type of electromagnetic radiation travels at a particular wavelength. The longer the wavelength (or the more stretched out it appears in the diagram), the less energy is carried. Short, tight waves carry the most energy. This may seem illogical, but think of it in terms of a piece of moving a heavy rope. It takes little effort by a person to move a rope in long, wide waves. To make a rope move in short, tight waves, a person would need to apply significantly more energy.

The electromagnetic spectrum (Figure) shows several types of electromagnetic radiation originating from the sun, including X-rays and ultraviolet (UV) rays. The higher-energy waves can penetrate tissues and damage cells and DNA, explaining why both X-rays and UV rays can be harmful to living organisms.

Absorption of Light

Light energy initiates the process of photosynthesis when pigments absorb the light. Organic pigments, whether in the human retina or the chloroplast thylakoid, have a narrow range of energy levels that they can absorb. Energy levels lower than those represented by red light are insufficient to raise an orbital electron to a populatable, excited (quantum) state. Energy levels higher than those in blue light will physically tear the molecules apart, called bleaching. So retinal pigments can only "see" (absorb) 700 nm to 400 nm light, which is therefore called visible light. For the same reasons, plants pigment molecules absorb only light in the wavelength range of 700 nm to 400 nm; plant physiologists refer to this range for plants as photosynthetically active radiation.

The visible light seen by humans as white light actually exists in a rainbow of colors. Certain objects, such as a prism or a drop of water, disperse white light to reveal the colors to the human eye. The visible light portion of the electromagnetic spectrum shows the rainbow of colors, with violet and blue having shorter wavelengths, and therefore higher energy. At the other end of the spectrum toward red, the wavelengths are longer and have lower energy (Figure).



The colors of visible light do not carry the same amount of energy. Violet has the shortest wavelength and therefore carries the most energy, whereas red has the longest wavelength and carries the least amount of energy. (credit: modification of work by NASA)

Understanding Pigments

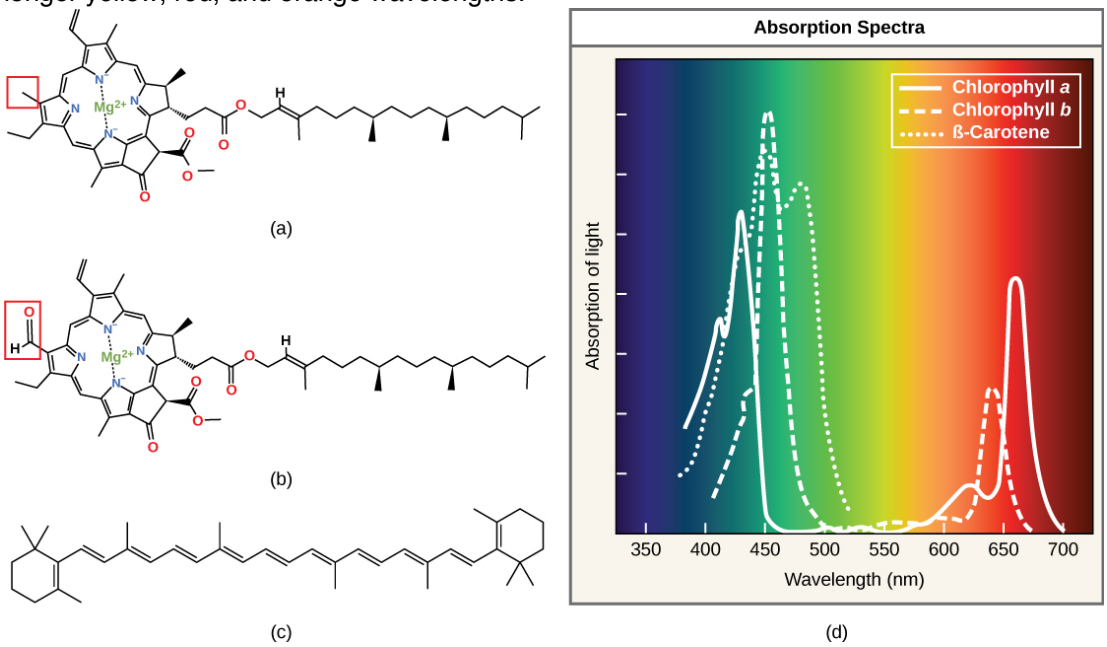
Different kinds of pigments exist, and each has evolved to absorb only certain wavelengths (colors) of visible light. Pigments reflect or transmit the wavelengths they cannot absorb, making them appear in the corresponding color.

Chlorophylls and carotenoids are the two major classes of photosynthetic pigments found in plants and algae; each class has multiple types of pigment molecules. There are five major chlorophylls: *a*, *b*, *c* and *d* and a related molecule found in prokaryotes called bacteriochlorophyll. Chlorophyll *a* and chlorophyll *b* are found in higher plant chloroplasts and will be the focus of the following discussion.

With dozens of different forms, carotenoids are a much larger group of pigments. The carotenoids found in fruit—such as the red of tomato (lycopene), the yellow of corn seeds (zeaxanthin), or the orange of an orange peel (β-carotene)—are used as advertisements to attract seed dispersers. In photosynthesis, carotenoids function as photosynthetic pigments that are very efficient molecules for the disposal of excess energy. When a leaf is

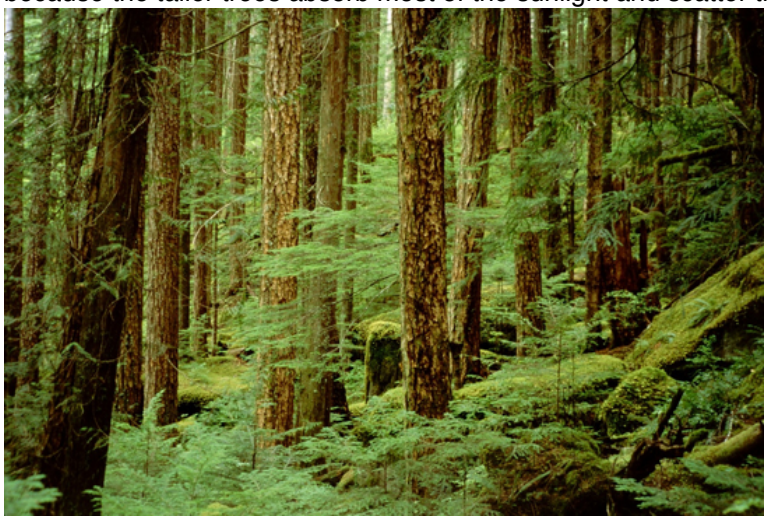
exposed to full sun, the light-dependent reactions are required to process an enormous amount of energy; if that energy is not handled properly, it can do significant damage. Therefore, many carotenoids reside in the thylakoid membrane, absorb excess energy, and safely dissipate that energy as heat.

Each type of pigment can be identified by the specific pattern of wavelengths it absorbs from visible light, which is the absorption spectrum. The graph in Figure shows the absorption spectra for chlorophyll *a*, chlorophyll *b*, and a type of carotenoid pigment called β-carotene (which absorbs blue and green light). Notice how each pigment has a distinct set of peaks and troughs, revealing a highly specific pattern of absorption. Chlorophyll *a* absorbs wavelengths from either end of the visible spectrum (blue and red), but not green. Because green is reflected or transmitted, chlorophyll appears green. Carotenoids absorb in the short-wavelength blue region, and reflect the longer yellow, red, and orange wavelengths.



(a) Chlorophyll *a*, (b) chlorophyll *b*, and (c) *β*-carotene are hydrophobic organic pigments found in the thylakoid membrane. Chlorophyll *a* and *b*, which are identical except for the part indicated in the red box, are responsible for the green color of leaves. *β*-carotene is responsible for the orange color in carrots. Each pigment has (d) a unique absorbance spectrum.

Many photosynthetic organisms have a mixture of pigments; using them, the organism can absorb energy from a wider range of wavelengths. Not all photosynthetic organisms have full access to sunlight. Some organisms grow underwater where light intensity and quality decrease and change with depth. Other organisms grow in competition for light. Plants on the rainforest floor must be able to absorb any bit of light that comes through, because the taller trees absorb most of the sunlight and scatter the remaining solar radiation (Figure).



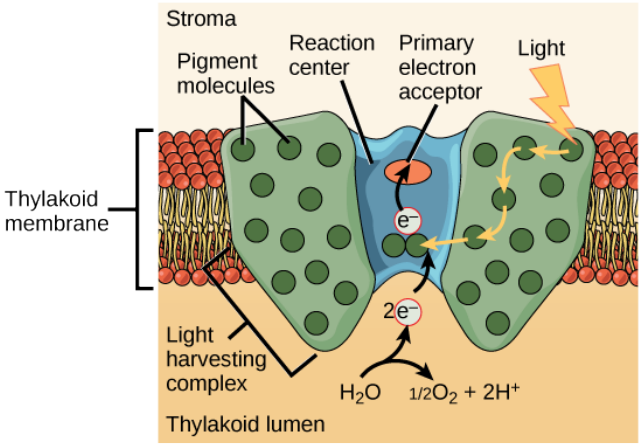
Plants that commonly grow in the shade have adapted to low levels of light by changing the relative concentrations of their chlorophyll pigments. (credit: Jason Hollinger)

When studying a photosynthetic organism, scientists can determine the types of pigments present by generating absorption spectra. An instrument called a spectrophotometer can differentiate which wavelengths of light a substance can absorb. Spectrophotometers measure transmitted light and compute from it the absorption. By extracting pigments from leaves and placing these samples into a spectrophotometer, scientists can identify which wavelengths of light an organism can absorb. Additional methods for the identification of plant pigments include various types of chromatography that separate the pigments by their relative affinities to solid and mobile phases.

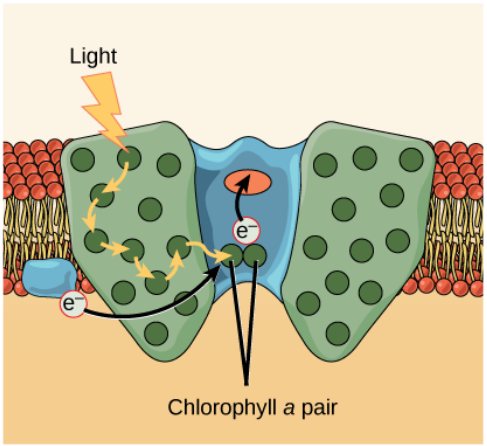
How Light-Dependent Reactions Work

The overall function of light-dependent reactions is to convert solar energy into chemical energy in the form of NADPH and ATP. This chemical energy supports the light-independent reactions and fuels the assembly of sugar molecules. The light-dependent reactions are depicted in Figure. Protein complexes and pigment molecules work together to produce NADPH and ATP.

(a) Photosystem II (P680)



(b) Photosystem I (P700)

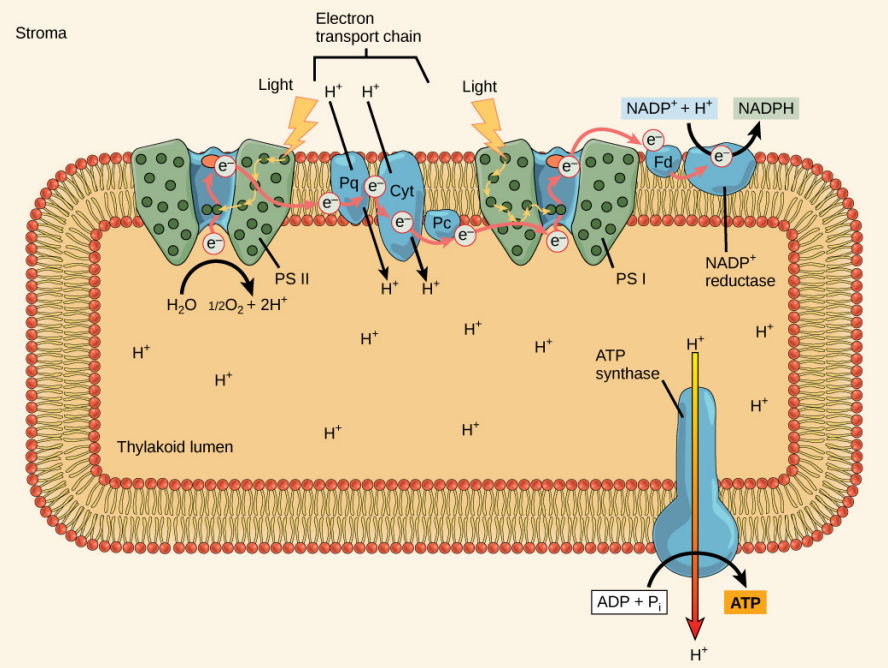


A photosystem consists of a light-harvesting complex and a reaction center. Pigments in the light-harvesting complex pass light energy to two special chlorophyll *a* molecules in the reaction center. The light excites an electron from the chlorophyll *a* pair, which passes to the primary electron acceptor. The excited electron must then be replaced. In (a) photosystem II, the electron comes from the splitting of water, which releases oxygen as a waste product. In (b) photosystem I, the electron comes from the chloroplast electron transport chain discussed below.

The actual step that converts light energy into chemical energy takes place in a multiprotein complex called a photosystem, two types of which are found embedded in the thylakoid membrane, photosystem II (PSII) and photosystem I (PSI) (Figure). The two complexes differ on the basis of what they oxidize (that is, the source of the low-energy electron supply) and what they reduce (the place to which they deliver their energized electrons).

Both photosystems have the same basic structure; a number of antenna proteins to which the chlorophyll molecules are bound surround the reaction center where the photochemistry takes place. Each photosystem is serviced by the light-harvesting complex, which passes energy from sunlight to the reaction center; it consists of multiple antenna proteins that contain a mixture of 300–400 chlorophyll *a* and *b* molecules as well as other pigments like carotenoids. The absorption of a single photon or distinct quantity or "packet" of light by any of the chlorophylls pushes that molecule into an excited state. In short, the light energy has now been captured by biological molecules but is not stored in any useful form yet. The energy is transferred from chlorophyll to chlorophyll until eventually (after about a millionth of a second), it is delivered to the reaction center. Up to this point, only energy has been transferred between molecules, not electrons.

ART CONNECTION



In the photosystem II (PSII) reaction center, energy from sunlight is used to extract electrons from water. The electrons travel through the chloroplast electron transport chain to photosystem I (PSI), which reduces $NADP^+$ to NADPH. The electron transport chain moves protons across the thylakoid membrane into the lumen. At the same time, splitting of water adds protons to the lumen, and reduction of NADPH removes protons from the stroma. The net result is a low pH in the thylakoid lumen, and a high pH in the stroma. ATP synthase uses this electrochemical gradient to make ATP.

What is the initial source of electrons for the chloroplast electron transport chain?

1. water
2. oxygen
3. carbon dioxide
4. NADPH

The reaction center contains a pair of chlorophyll *a* molecules with a special property. Those two chlorophylls can undergo oxidation upon excitation; they can actually give up an electron in a process called a photoact. It is at this step in the reaction center, this step in photosynthesis, that light energy is converted into an excited electron. All of the subsequent steps involve getting that electron onto the energy carrier NADPH for delivery to the Calvin cycle where the electron is deposited onto carbon for long-term storage in the form of a carbohydrate.PSII and PSI are two major components of the photosynthetic electron transport chain, which also includes the cytochrome complex. The cytochrome complex, an enzyme composed of two protein complexes, transfers the electrons from the carrier molecule plastoquinone (Pq) to the protein plastocyanin (Pc), thus enabling both the transfer of protons across the thylakoid membrane and the transfer of electrons from PSII to PSI.

The reaction center of PSII (called P680) delivers its high-energy electrons, one at the time, to the primary electron acceptor, and through the electron transport chain (Pq to cytochrome complex to plastocyanine) to PSI. P680's missing electron is replaced by extracting a low-energy electron from water; thus, water is split and PSII is re-reduced after every photoact. Splitting one H_2O molecule releases two electrons, two hydrogen atoms, and one atom of oxygen. Splitting two molecules is required to form one molecule of diatomic O_2 gas. About 10 percent of the oxygen is used by mitochondria in the leaf to support oxidative phosphorylation. The remainder escapes to the atmosphere where it is used by aerobic organisms to support respiration.

As electrons move through the proteins that reside between PSII and PSI, they lose energy. That energy is used to move hydrogen atoms from the stromal side of the membrane to the thylakoid lumen. Those hydrogen atoms, plus the ones produced by splitting water, accumulate in the thylakoid lumen and will be used synthesize ATP in a later step. Because the electrons have lost energy prior to their arrival at PSI, they must be re-energized by PSI, hence, another photon is absorbed by the PSI antenna. That energy is relayed to the PSI reaction center (called P700). P700 is oxidized and sends a high-energy electron to $NADP^+$ to form NADPH. Thus, PSII captures the energy to create proton gradients to make ATP, and PSI captures the energy to reduce $NADP^+$ into NADPH. The two photosystems work in concert, in part, to guarantee that the production of NADPH will roughly equal the production of ATP. Other mechanisms exist to fine tune that ratio to exactly match the chloroplast's constantly changing energy needs.

Generating an Energy Carrier: ATP

As in the intermembrane space of the mitochondria during cellular respiration, the buildup of hydrogen ions inside the thylakoid lumen creates a concentration gradient. The passive diffusion of hydrogen ions from high concentration (in the thylakoid lumen) to low concentration (in the stroma) is harnessed to create ATP, just as in the electron transport chain of cellular respiration. The ions build up energy because of diffusion and because they all have the same electrical charge, repelling each other.

To release this energy, hydrogen ions will rush through any opening, similar to water jetting through a hole in a dam. In the thylakoid, that opening is a passage through a specialized protein channel called the ATP synthase. The energy released by the hydrogen ion stream allows ATP synthase to attach a third phosphate group to ADP, which forms a molecule of ATP (Figure). The flow of hydrogen ions through ATP synthase is called chemiosmosis because the ions move from an area of high to an area of low concentration through a semi-permeable structure.

Link to Learning

Visit this site and click through the animation to view the process of photosynthesis within a leaf.

Section Summary

The pigments of the first part of photosynthesis, the light-dependent reactions, absorb energy from sunlight. A photon strikes the antenna pigments of photosystem II to initiate photosynthesis. The energy travels to the reaction center that contains chlorophyll*a* to the electron transport chain, which pumps hydrogen ions into the thylakoid interior. This action builds up a high concentration of ions. The ions flow through ATP synthase via chemiosmosis to form molecules of ATP, which are used for the formation of sugar molecules in the second stage of photosynthesis. Photosystem I absorbs a second photon, which results in the formation of an NADPH molecule, another energy and reducing power carrier for the light-independent reactions.

Art Connections

Figure What is the source of electrons for the chloroplast electron transport chain?

1. Water
2. Oxygen
3. Carbon dioxide

4. NADPH

Review Questions

Which of the following structures is *not* a component of a photosystem?

1. ATP synthase
2. antenna molecule
3. reaction center
4. primary electron acceptor

How many photons does it take to fully reduce one molecule of $NADP^+$ to NADPH?

1. 1
2. 2
3. 4
4. 8

Which complex is *not* involved in the establishment of conditions for ATP synthesis?

1. photosystem I
2. ATP synthase
3. photosystem II
4. cytochrome complex

From which component of the light-dependent reactions does NADPH form most directly?

1. photosystem II
2. photosystem I
3. cytochrome complex
4. ATP synthase

Free Response

Describe the pathway of electron transfer from photosystem II to photosystem I in light-dependent reactions.

What are the roles of ATP and NADPH in photosynthesis?

Glossary

absorption spectrum range of wavelengths of electromagnetic radiation absorbed by a given substance

antenna protein pigment molecule that directly absorbs light and transfers the energy absorbed to other pigment molecules

carotenoid photosynthetic pigment that functions to dispose of excess energy

chlorophyll *a* form of chlorophyll that absorbs violet-blue and red light and consequently has a bluish-green color; the only pigment molecule that performs the photochemistry by getting excited and losing an electron to the electron transport chain

chlorophyll *b* accessory pigment that absorbs blue and red-orange light and consequently has a yellowish-green tint

cytochrome complex group of reversibly oxidizable and reducible proteins that forms part of the electron transport chain between photosystem II and photosystem I

electromagnetic spectrum range of all possible frequencies of radiation

electron transport chain group of proteins between PSII and PSI that pass energized electrons and use the energy released by the electrons to move hydrogen ions against their concentration gradient into the thylakoid lumen

light harvesting complex complex that passes energy from sunlight to the reaction center in each photosystem; it consists of multiple antenna proteins that contain a mixture of 300–400 chlorophyll *a* and *b* molecules as well as other pigments like carotenoids

P680 reaction center of photosystem II

P700 reaction center of photosystem I

photoact ejection of an electron from a reaction center using the energy of an absorbed photon

photon distinct quantity or "packet" of light energy

photosystem group of proteins, chlorophyll, and other pigments that are used in the light-dependent reactions of photosynthesis to absorb light energy and convert it into chemical energy

photosystem I integral pigment and protein complex in thylakoid membranes that uses light energy to transport electrons from plastocyanin to $NADP^+$ (which becomes reduced to NADPH in the process)

photosystem II integral protein and pigment complex in thylakoid membranes that transports electrons from water to the electron transport chain; oxygen is a product of PSII

primary electron acceptor pigment or other organic molecule in the reaction center that accepts an energized electron from the reaction center

reaction center complex of chlorophyll molecules and other organic molecules that is assembled around a special pair of chlorophyll molecules and a primary electron acceptor; capable of undergoing oxidation and reduction

spectrophotometer instrument that can measure transmitted light and compute the absorption

wavelength distance between consecutive points of equal position (two crests or two troughs) of a wave in a graphic representation; inversely proportional to the energy of the radiation

Licensing & Attributions

CC licensed content, Shared previously

- light dependent reactions. **Authored by**: open stax. **Provided by**: opens tax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:-CmzvUct@12/The-Light-Dependent-Reactions-. **License**: *CC BY: Attribution*

STUDY GUIDE: PHOTOSYNTHESIS

Study Questions

Objective: Describe the process by which solar energy can be stored in a glucose molecule.

Use this page to check your understanding of the content.

Vocabulary

1. Stroma
2. Grana (sing: granum)
3. Thylakoid

4. Thylakoid space
5. Pigment molecules
6. Chlorophyll
7. photon
8. electron carrier
9. ATP synthase
10. chemiosmosis

Study Guide Questions

1. What is the structure of a chloroplast?
2. What is the equation for photosynthesis?
3. Where does photosynthesis take place? (Be complete in your answer. In other words, tell me ALL the specific places where photosynthetic events occur!)
4. Understand the electromagnetic spectrum, and be able to explain why an object appears the color you SEE.
5. What do you think would happen if you placed a plant in GREEN light ONLY? How would you design an experiment to test this?
6. Describe each step in the process of photosynthesis. Connect your answer to the chemical equation describing photosynthesis.
7. Clearly explain the connection between photosystem I and photosystem II. Compare and CONTRAST these photosystems.
8. What occurs in the Calvin Cycle? How is the Calvin Cycle connected to events in Photosystems II and I?
9. What is the evolutionary significance of photosynthesis?
10. Where does the sugar come from that is used during cellular respiration?
11. Where does the oxygen come from that is used during cellular respiration?

Licensing & Attributions

CC licensed content, Shared previously

- Study Guide: Photosynthesis. **Authored by**: Wendy Riggs. **Provided by**: College of the Redwoods. **Project**: Kaleidoscope. **License**: *CC BY: Attribution*

CELLULAR RESPIRATON

INTRODUCTION: CELLULAR RESPIRATION

This geothermal energy plant transforms thermal energy from deep in the ground into electrical energy, which can be easily used. (credit: modification of work by the U.S. Department of Defense)

The electrical energy plant in Figure converts energy from one form to another form that can be more easily used. This type of generating plant starts with underground thermal energy (heat) and transforms it into electrical energy that will be transported to homes and factories. Like a generating plant, plants and animals also must take in energy from the environment and convert it into a form that their cells can use. Energy enters an organism's body in one form and is converted into another form that can fuel the organism's life functions. In the process of photosynthesis, plants and other photosynthetic producers take in energy in the form of light (solar energy) and convert it into chemical energy, glucose, which stores this energy in its chemical bonds. Then, a series of metabolic pathways, collectively called cellular respiration, extracts the energy from the bonds in glucose and converts it into a form that all living things can use—both producers, such as plants, and consumers, such as animals.

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: Open Stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License**: *CC BY: Attribution*

ENERGY IN LIVING SYSTEMS

Learning Objectives:

By the end of this section, you will be able to:

- Discuss the importance of electrons in the transfer of energy in living systems
- Explain how ATP is used by the cell as an energy source

Energy production within a cell involves many coordinated chemical pathways. Most of these pathways are combinations of oxidation and reduction reactions. Oxidation and reduction occur in tandem. An oxidation reaction strips an electron from an atom in a compound, and the addition of this electron to another compound is a reduction reaction. Because oxidation and reduction usually occur together, these pairs of reactions are called oxidation reduction reactions, or redox reactions.

Electrons and Energy

The removal of an electron from a molecule, oxidizing it, results in a decrease in potential energy in the oxidized compound. The electron (sometimes as part of a hydrogen atom), does not remain unbonded, however, in the cytoplasm of a cell. Rather, the electron is shifted to a second compound, reducing the second compound. The shift of an electron from one compound to another removes some potential energy from the first compound (the oxidized compound) and increases the potential energy of the second compound (the reduced compound). The transfer of electrons between molecules is important because most of the energy stored in atoms and used to fuel cell functions is in the form of high-energy electrons. The transfer of energy in the form of electrons allows the cell to transfer and use energy in an incremental fashion—in small packages rather than in a single, destructive burst. This chapter focuses on the extraction of energy from food; you will see that as you track the path of the transfers, you are tracking the path of electrons moving through metabolic pathways.

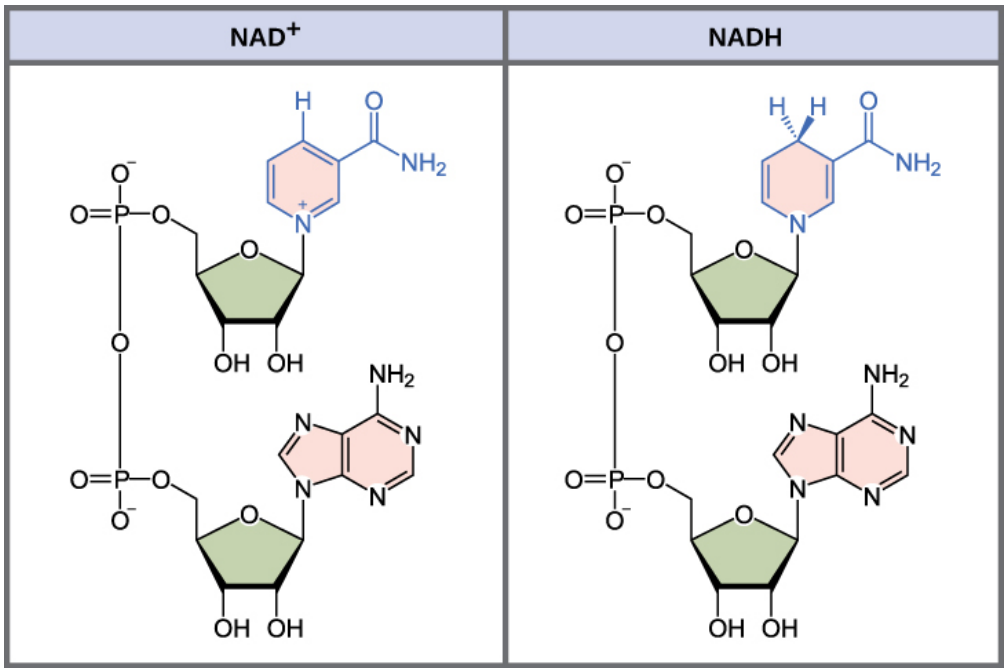
Electron Carriers

In living systems, a small class of compounds functions as electron shuttles: They bind and carry high-energy electrons between compounds in pathways. The principal electron carriers we will consider are derived from the B vitamin group and are derivatives of nucleotides. These compounds can be easily reduced (that is, they accept electrons) or oxidized (they lose electrons). Nicotinamide adenine dinucleotide (NAD) (Figure) is derived from vitamin B3, niacin. NAD^+ is the oxidized form of the molecule; NADH is the reduced form of the molecule after it has accepted two electrons and a proton (which together are the equivalent of a hydrogen atom with an extra electron).

NAD^+ can accept electrons from an organic molecule according to the general equation:
RHReducing agent + NAD+Oxidizingagent→ NADHReduced +ROxidizedRHReducing agent + NAD+Oxidizingagent→ N

When electrons are added to a compound, they are reduced. A compound that reduces another is called a reducing agent. In the above equation, RH is a reducing agent, and NAD^+ is reduced to NADH. When electrons are removed from compound, it oxidized. A compound that oxidizes another is called an oxidizing agent. In the above equation, NAD^+ is an oxidizing agent, and RH is oxidized to R.

Similarly, flavin adenine dinucleotide (FAD^+) is derived from vitamin B_2, also called riboflavin. Its reduced form is $FADH_2$. A second variation of NAD, NADP, contains an extra phosphate group. Both NAD^+ and FAD^+ are extensively used in energy extraction from sugars, and NADP plays an important role in anabolic reactions and photosynthesis.



The oxidized form of the electron carrier (NAD^+) is shown on the left and the reduced form (NADH) is shown on the right. The nitrogenous base in NADH has one more hydrogen ion and two more electrons than in NAD^+.

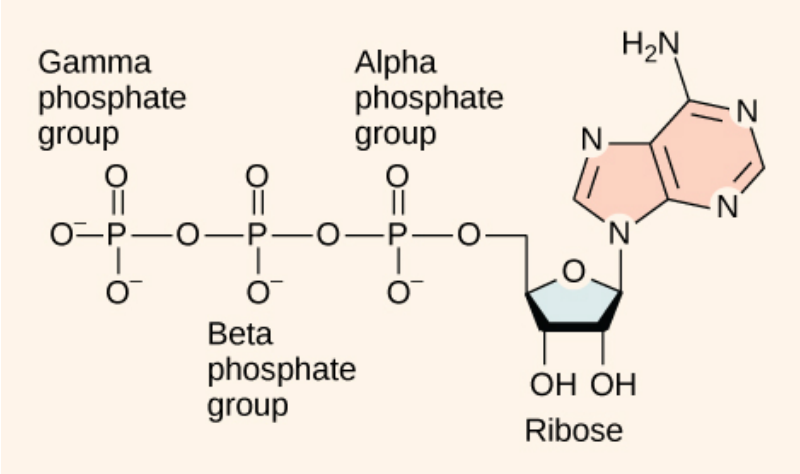
ATP in Living Systems

A living cell cannot store significant amounts of free energy. Excess free energy would result in an increase of heat in the cell, which would result in excessive thermal motion that could damage and then destroy the cell. Rather, a cell must be able to handle that energy in a way that enables the cell to store energy safely and release it for use only as needed. Living cells accomplish this by using the compound adenosine triphosphate (ATP). ATP is often called the “energy currency” of the cell, and, like currency, this versatile compound can be used to fill any energy need of the cell. How? It functions similarly to a rechargeable battery.

When ATP is broken down, usually by the removal of its terminal phosphate group, energy is released. The energy is used to do work by the cell, usually by the released phosphate binding to another molecule, activating it. For example, in the mechanical work of muscle contraction, ATP supplies the energy to move the contractile muscle proteins. Recall the active transport work of the sodium-potassium pump in cell membranes. ATP alters the structure of the integral protein that functions as the pump, changing its affinity for sodium and potassium. In this way, the cell performs work, pumping ions against their electrochemical gradients.

ATP Structure and Function

At the heart of ATP is a molecule of adenosine monophosphate (AMP), which is composed of an adenine molecule bonded to a ribose molecule and to a single phosphate group (Figure). Ribose is a five-carbon sugar found in RNA, and AMP is one of the nucleotides in RNA. The addition of a second phosphate group to this core molecule results in the formation of adenosinediphosphate (ADP); the addition of a third phosphate group forms adenosine triphosphate (ATP).



ATP (adenosine triphosphate) has three phosphate groups that can be removed by hydrolysis to form ADP (adenosine diphosphate) or AMP (adenosine monophosphate).The negative charges on the phosphate group naturally repel each other, requiring energy to bond them together and releasing energy when these bonds are broken.

The addition of a phosphate group to a molecule requires energy. Phosphate groups are negatively charged and thus repel one another when they are arranged in series, as they are in ADP and ATP. This repulsion makes the ADP and ATP molecules inherently unstable. The release of one or two phosphate groups from ATP, a process called dephosphorylation, releases energy.

Energy from ATP

Hydrolysis is the process of breaking complex macromolecules apart. During hydrolysis, water is split, or lysed, and the resulting hydrogen atom (H^+) and a hydroxyl group (OH^-) are added to the larger molecule. The hydrolysis of ATP produces ADP, together with an inorganic phosphate ion (P_i), and the release of free energy. To carry out life processes, ATP is continuously broken down into ADP, and like a rechargeable battery, ADP is continuously regenerated into ATP by the reattachment of a third phosphate group. Water, which was broken down into its hydrogen atom and hydroxyl group during ATP hydrolysis, is regenerated when a third phosphate is added to the ADP molecule, reforming ATP.

Obviously, energy must be infused into the system to regenerate ATP. Where does this energy come from? In nearly every living thing on earth, the energy comes from the metabolism of glucose. In this way, ATP is a direct link between the limited set of exergonic pathways of glucose catabolism and the multitude of endergonic pathways that power living cells.

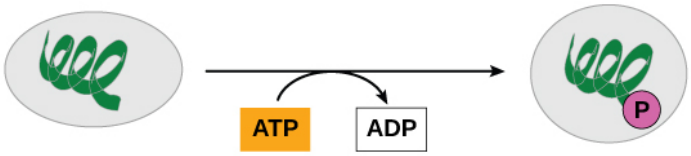
Phosphorylation

Recall that, in some chemical reactions, enzymes may bind to several substrates that react with each other on the enzyme, forming an intermediate complex. An intermediate complex is a temporary structure, and it allows one of the substrates (such as ATP) and reactants to more readily react with each other; in reactions involving ATP, ATP is one of the substrates and ADP is a product. During an endergonic chemical reaction, ATP forms an intermediate complex with the substrate and enzyme in the reaction. This intermediate complex allows the ATP to transfer its third phosphate group, with its energy, to the substrate, a process called phosphorylation. Phosphorylation refers to the addition of the phosphate (~P). This is illustrated by the following generic reaction:
A + enzyme + ATP→ [A – enzyme – ~P] → B + enzyme + ADP + phosphate ionA + enzyme + ATP→ [A – enzyme – ~P] → B + enzyme + ADP + phosphate ion

When the intermediate complex breaks apart, the energy is used to modify the substrate and convert it into a product of the reaction. The ADP molecule and a free phosphate ion are released into the medium and are available for recycling through cell metabolism.

Substrate Phosphorylation

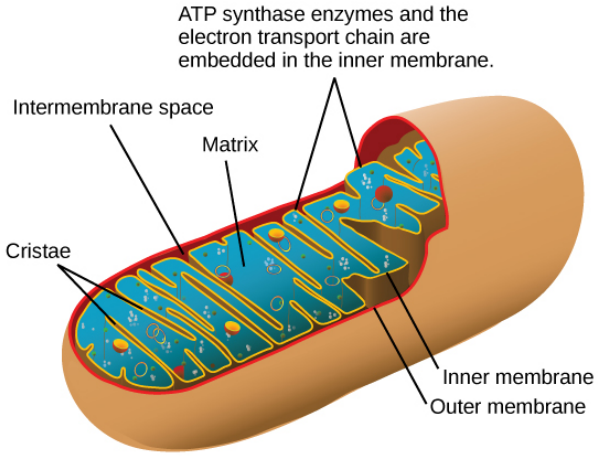
ATP is generated through two mechanisms during the breakdown of glucose. A few ATP molecules are generated (that is, regenerated from ADP) as a direct result of the chemical reactions that occur in the catabolic pathways. A phosphate group is removed from an intermediate reactant in the pathway, and the free energy of the reaction is used to add the third phosphate to an available ADP molecule, producing ATP (Figure). This very direct method of phosphorylation is called substrate-level phosphorylation.



In phosphorylation reactions, the gamma phosphate of ATP is attached to a protein.

Oxidative Phosphorylation

Most of the ATP generated during glucose catabolism, however, is derived from a much more complex process, chemiosmosis, which takes place in mitochondria (Figure) within a eukaryotic cell or the plasma membrane of a prokaryotic cell. Chemiosmosis, a process of ATP production in cellular metabolism, is used to generate 90 percent of the ATP made during glucose catabolism and is also the method used in the light reactions of photosynthesis to harness the energy of sunlight. The production of ATP using the process of chemiosmosis is called oxidative phosphorylation because of the involvement of oxygen in the process.



In eukaryotes, oxidative phosphorylation takes place in mitochondria. In prokaryotes, this process takes place in the plasma membrane. (Credit: modification of work by Mariana Ruiz Villareal)

CAREER CONNECTIONS

Mitochondrial Disease Physician

What happens when the critical reactions of cellular respiration do not proceed correctly? Mitochondrial diseases are genetic disorders of metabolism. Mitochondrial disorders can arise from mutations in nuclear or mitochondrial DNA, and they result in the production of less energy than is normal in body cells. In type 2 diabetes, for instance, the oxidation efficiency of NADH is reduced, impacting oxidative phosphorylation but not the other steps of respiration. Symptoms of mitochondrial diseases can include muscle weakness, lack of coordination, stroke-like episodes, and loss of vision and hearing. Most affected people are diagnosed in childhood, although there are some adult-onset diseases. Identifying and treating mitochondrial disorders is a specialized medical field. The educational preparation for this profession requires a college education, followed by medical school with a specialization in medical genetics. Medical geneticists can be board certified by the American Board of Medical Genetics and go on to become associated with professional organizations devoted to the study of mitochondrial diseases, such as the Mitochondrial Medicine Society and the Society for Inherited Metabolic Disease.

Section Summary

ATP functions as the energy currency for cells. It allows the cell to store energy briefly and transport it within the cell to support endergonic chemical reactions. The structure of ATP is that of an RNA nucleotide with three phosphates attached. As ATP is used for energy, a phosphate group or two are detached, and either ADP or AMP is produced. Energy derived from glucose catabolism is used to convert ADP into ATP. When ATP is used in a reaction, the third phosphate is temporarily attached to a substrate in a process called phosphorylation. The two processes of ATP regeneration that are used in conjunction with glucose catabolism are substrate-level phosphorylation and oxidative phosphorylation through the process of chemiosmosis.

Review Questions

The energy currency used by cells is _____.

1. ATP
2. ADP
3. AMP
4. adenosine

A reducing chemical reaction _____.

1. reduces the compound to a simpler form
2. adds an electron to the substrate
3. removes a hydrogen atom from the substrate
4. is a catabolic reaction

Free Response

Why is it beneficial for cells to use ATP rather than energy directly from the bonds of carbohydrates? What are the greatest drawbacks to harnessing energy directly from the bonds of several different compounds?

Glossary

chemiosmosis process in which there is a production of adenosine triphosphate (ATP) in cellular metabolism by the involvement of a proton gradient across a membrane

dephosphorylation removal of a phosphate group from a molecule

oxidative phosphorylation production of ATP using the process of chemiosmosis and oxygen

phosphorylation addition of a high-energy phosphate to a compound, usually a metabolic intermediate, a protein, or ADP

redox reaction chemical reaction that consists of the coupling of an oxidation reaction and a reduction reaction

substrate-level phosphorylation production of ATP from ADP using the excess energy from a chemical reaction and a phosphate group from a reactant

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: OpenStax College. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License**: *CC BY: Attribution*

GLYCOLYSIS

Learning Objectives

By the end of this section, you will be able to:

- Describe the overall result in terms of molecules produced in the breakdown of glucose by glycolysis
- Compare the output of glycolysis in terms of ATP molecules and NADH molecules produced

You have read that nearly all of the energy used by living cells comes to them in the bonds of the sugar, glucose. Glycolysis is the first step in the breakdown of glucose to extract energy for cellular metabolism. Nearly all living organisms carry out glycolysis as part of their metabolism. The process does not use oxygen and is therefore anaerobic. Glycolysis takes place in the cytoplasm of both prokaryotic and eukaryotic cells. Glucose enters heterotrophic cells in two ways. One method is through secondary active transport in which the transport takes place against the glucose concentration gradient. The other mechanism uses a group of integral proteins called GLUT proteins, also known as glucose transporter proteins. These transporters assist in the facilitated diffusion of glucose.

Glycolysis begins with the six carbon ring-shaped structure of a single glucose molecule and ends with two molecules of a three-carbon sugar called pyruvate. Glycolysis consists of two distinct phases. The first part of the glycolysis pathway traps the glucose molecule in the cell and uses energy to modify it so that the six-carbon sugar molecule can be split evenly into the two three-carbon molecules. The second part of glycolysis extracts energy from the molecules and stores it in the form of ATP and NADH, the reduced form of NAD.

First Half of Glycolysis (Energy-Requiring Steps)

Step 1. The first step in glycolysis (Figure) is catalyzed by hexokinase, an enzyme with broad specificity that catalyzes the phosphorylation of six-carbon sugars. Hexokinase phosphorylates glucose using ATP as the source of the phosphate, producing glucose-6-phosphate, a more reactive form of glucose. This reaction prevents the phosphorylated glucose molecule from continuing to interact with the GLUT proteins, and it can no longer leave the cell because the negatively charged phosphate will not allow it to cross the hydrophobic interior of the plasma membrane.

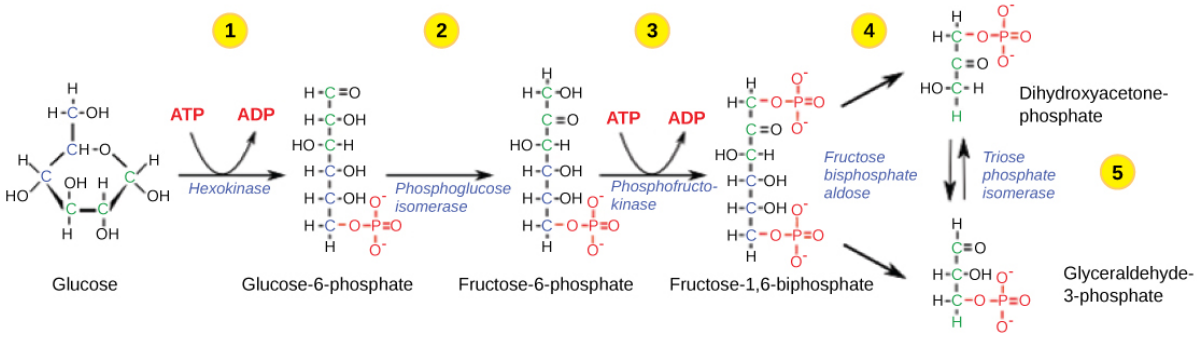
Step 2. In the second step of glycolysis, an isomerase converts glucose-6-phosphate into one of its isomers, fructose-6-phosphate. An isomerase is an enzyme that catalyzes the conversion of a molecule into one of its isomers. (This change from phosphoglucose to phosphofructose allows the eventual split of the sugar into two three-carbon molecules.).

Step 3. The third step is the phosphorylation of fructose-6-phosphate, catalyzed by the enzyme phosphofructokinase. A second ATP molecule donates a high-energy phosphate to fructose-6-phosphate, producing fructose-1,6-bisphosphate. In this pathway, phosphofructokinase is a rate-limiting enzyme. It is active when the concentration of ADP is high; it is less active when ADP levels are low and the concentration of ATP is high. Thus, if there is “sufficient” ATP in the system, the pathway slows down. This is a type of end product inhibition, since ATP is the end product of glucose catabolism.

Step 4. The newly added high-energy phosphates further destabilize fructose-1,6-bisphosphate. The fourth step in glycolysis employs an enzyme, aldolase, to cleave 1,6-bisphosphate into two three-carbon isomers: dihydroxyacetone-phosphate and glyceraldehyde-3-phosphate.

Step 5. In the fifth step, an isomerase transforms the dihydroxyacetone-phosphate into its isomer, glyceraldehyde-3-phosphate. Thus, the pathway will continue with two molecules of a single isomer. At this point

in the pathway, there is a net investment of energy from two ATP molecules in the breakdown of one glucose molecule.

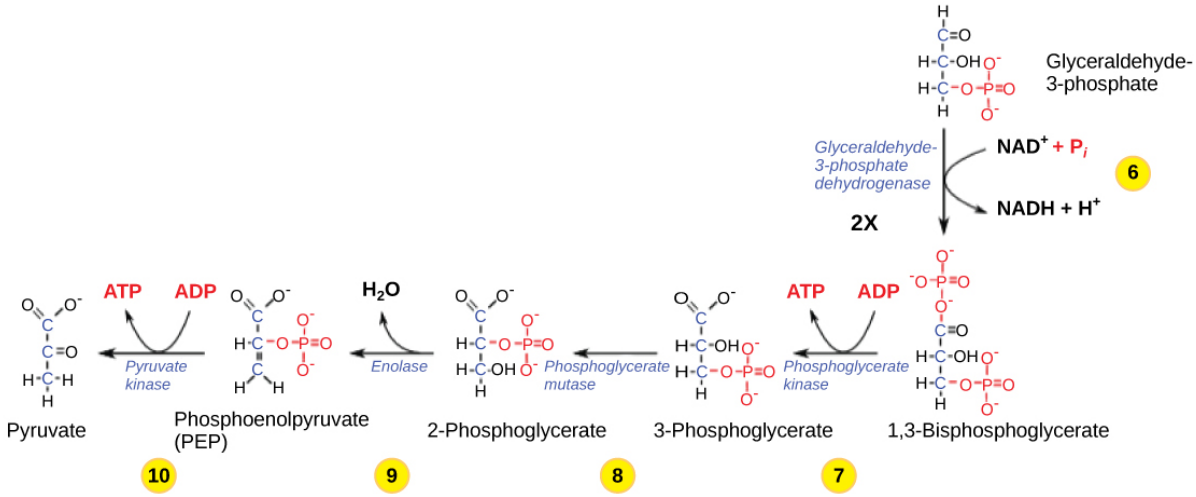


The first half of glycolysis uses two ATP molecules in the phosphorylation of glucose, which is then split into two three-carbon molecules.

Second Half of Glycolysis (Energy-Releasing Steps)

So far, glycolysis has cost the cell two ATP molecules and produced two small, three-carbon sugar molecules. Both of these molecules will proceed through the second half of the pathway, and sufficient energy will be extracted to pay back the two ATP molecules used as an initial investment and produce a profit for the cell of two additional ATP molecules and two even higher-energy NADH molecules.

Step 6. The sixth step in glycolysis (Figure) oxidizes the sugar (glyceraldehyde-3-phosphate), extracting high-energy electrons, which are picked up by the electron carrier NAD^+, producing NADH. The sugar is then phosphorylated by the addition of a second phosphate group, producing 1,3-bisphosphoglycerate. Note that the second phosphate group does not require another ATP molecule.



The second half of glycolysis involves phosphorylation without ATP investment (step 6) and produces two NADH and four ATP molecules per glucose.

Here again is a potential limiting factor for this pathway. The continuation of the reaction depends upon the availability of the oxidized form of the electron carrier, NAD^+. Thus, NADH must be continuously oxidized back into NAD^+ in order to keep this step going. If NAD^+ is not available, the second half of glycolysis slows down or stops. If oxygen is available in the system, the NADH will be oxidized readily, though indirectly, and the high-energy electrons from the hydrogen released in this process will be used to produce ATP. In an environment without oxygen, an alternate pathway (fermentation) can provide the oxidation of NADH to NAD^+.

Step 7. In the seventh step, catalyzed by phosphoglycerate kinase (an enzyme named for the reverse reaction), 1,3-bisphosphoglycerate donates a high-energy phosphate to ADP, forming one molecule of ATP. (This is an

example of substrate-level phosphorylation.) A carbonyl group on the 1,3-bisphosphoglycerate is oxidized to a carboxyl group, and 3-phosphoglycerate is formed.

Step 8. In the eighth step, the remaining phosphate group in 3-phosphoglycerate moves from the third carbon to the second carbon, producing 2-phosphoglycerate (an isomer of 3-phosphoglycerate). The enzyme catalyzing this step is a mutase (isomerase).

Step 9. Enolase catalyzes the ninth step. This enzyme causes 2-phosphoglycerate to lose water from its structure; this is a dehydration reaction, resulting in the formation of a double bond that increases the potential energy in the remaining phosphate bond and produces phosphoenolpyruvate (PEP).

Step 10. The last step in glycolysis is catalyzed by the enzyme pyruvate kinase (the enzyme in this case is named for the reverse reaction of pyruvate's conversion into PEP) and results in the production of a second ATP molecule by substrate-level phosphorylation and the compound pyruvic acid (or its salt form, pyruvate). Many enzymes in enzymatic pathways are named for the reverse reactions, since the enzyme can catalyze both forward and reverse reactions (these may have been described initially by the reverse reaction that takes place in vitro, under non-physiological conditions).

Link to Learning

Gain a better understanding of the breakdown of glucose by glycolysis by visiting this site to see the process in action.

Outcomes of Glycolysis

Glycolysis starts with glucose and ends with two pyruvate molecules, a total of four ATP molecules and two molecules of NADH. Two ATP molecules were used in the first half of the pathway to prepare the six-carbon ring for cleavage, so the cell has a net gain of two ATP molecules and 2 NADH molecules for its use. If the cell cannot catabolize the pyruvate molecules further, it will harvest only two ATP molecules from one molecule of glucose. Mature mammalian red blood cells are not capable of aerobic respiration—the process in which organisms convert energy in the presence of oxygen—and glycolysis is their sole source of ATP. If glycolysis is interrupted, these cells lose their ability to maintain their sodium-potassium pumps, and eventually, they die.

The last step in glycolysis will not occur if pyruvate kinase, the enzyme that catalyzes the formation of pyruvate, is not available in sufficient quantities. In this situation, the entire glycolysis pathway will proceed, but only two ATP molecules will be made in the second half. Thus, pyruvate kinase is a rate-limiting enzyme for glycolysis.

Section Summary

Glycolysis is the first pathway used in the breakdown of glucose to extract energy. It was probably one of the earliest metabolic pathways to evolve and is used by nearly all of the organisms on earth. Glycolysis consists of two parts: The first part prepares the six-carbon ring of glucose for cleavage into two three-carbon sugars. ATP is invested in the process during this half to energize the separation. The second half of glycolysis extracts ATP and high-energy electrons from hydrogen atoms and attaches them to NAD^+. Two ATP molecules are invested in the first half and four ATP molecules are formed by substrate phosphorylation during the second half. This produces a net gain of two ATP and two NADH molecules for the cell.

Review Questions

During the second half of glycolysis, what occurs?

1. ATP is used up.
2. Fructose is split in two.
3. ATP is made.
4. Glucose becomes fructose.

Free Response

Nearly all organisms on earth carry out some form of glycolysis. How does that fact support or not support the assertion that glycolysis is one of the oldest metabolic pathways?

Red blood cells do not perform aerobic respiration, but they do perform glycolysis. Why do all cells need an energy source, and what would happen if glycolysis were blocked in a red blood cell?

Glossary

aerobic respiration process in which organisms convert energy in the presence of oxygen

anaerobic process that does not use oxygen

glycolysis process of breaking glucose into two three-carbon molecules with the production of ATP and NADH

isomerase enzyme that converts a molecule into its isomer

pyruvate three-carbon sugar that can be decarboxylated and oxidized to make acetyl CoA, which enters the citric acid cycle under aerobic conditions; the end product of glycolysis

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: Open Stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License**: *CC BY: Attribution*

OXIDATION OF PYRUVATE AND THE CITRIC ACID CYCLE

Learning Objectives

By the end of this section, you will be able to:

- Explain how a circular pathway, such as the citric acid cycle, fundamentally differs from a linear pathway, such as glycolysis
- Describe how pyruvate, the product of glycolysis, is prepared for entry into the citric acid cycle

If oxygen is available, aerobic respiration will go forward. In eukaryotic cells, the pyruvate molecules produced at the end of glycolysis are transported into mitochondria, which are the sites of cellular respiration. There, pyruvate will be transformed into an acetyl group that will be picked up and activated by a carrier compound called coenzyme A (CoA). The resulting compound is called acetyl CoA. CoA is made from vitamin B5, pantothenic acid. Acetyl CoA can be used in a variety of ways by the cell, but its major function is to deliver the acetyl group derived from pyruvate to the next stage of the pathway in glucose catabolism.

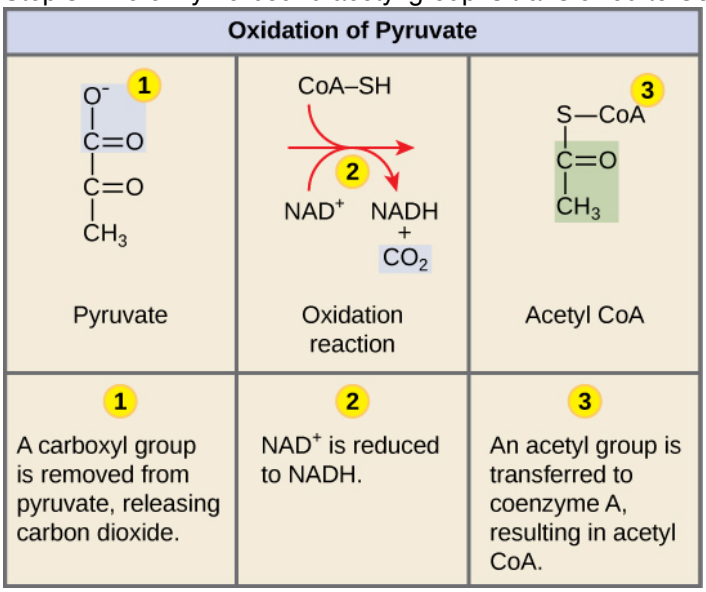
Breakdown of Pyruvate

In order for pyruvate, the product of glycolysis, to enter the next pathway, it must undergo several changes. The conversion is a three-step process (Figure).

Step 1. A carboxyl group is removed from pyruvate, releasing a molecule of carbon dioxide into the surrounding medium. The result of this step is a two-carbon hydroxyethyl group bound to the enzyme (pyruvate dehydrogenase). This is the first of the six carbons from the original glucose molecule to be removed. This step proceeds twice (remember: there are *two* pyruvate molecules produced at the end of glycolsis) for every molecule of glucose metabolized; thus, two of the six carbons will have been removed at the end of both steps.

Step 2. The hydroxyethyl group is oxidized to an acetyl group, and the electrons are picked up by NAD^+, forming NADH. The high-energy electrons from NADH will be used later to generate ATP.

Step 3. The enzyme-bound acetyl group is transferred to CoA, producing a molecule of acetyl CoA.



Upon entering the mitochondrial matrix, a multi-enzyme complex converts pyruvate into acetyl CoA. In the process, carbon dioxide is released and one molecule of NADH is formed.

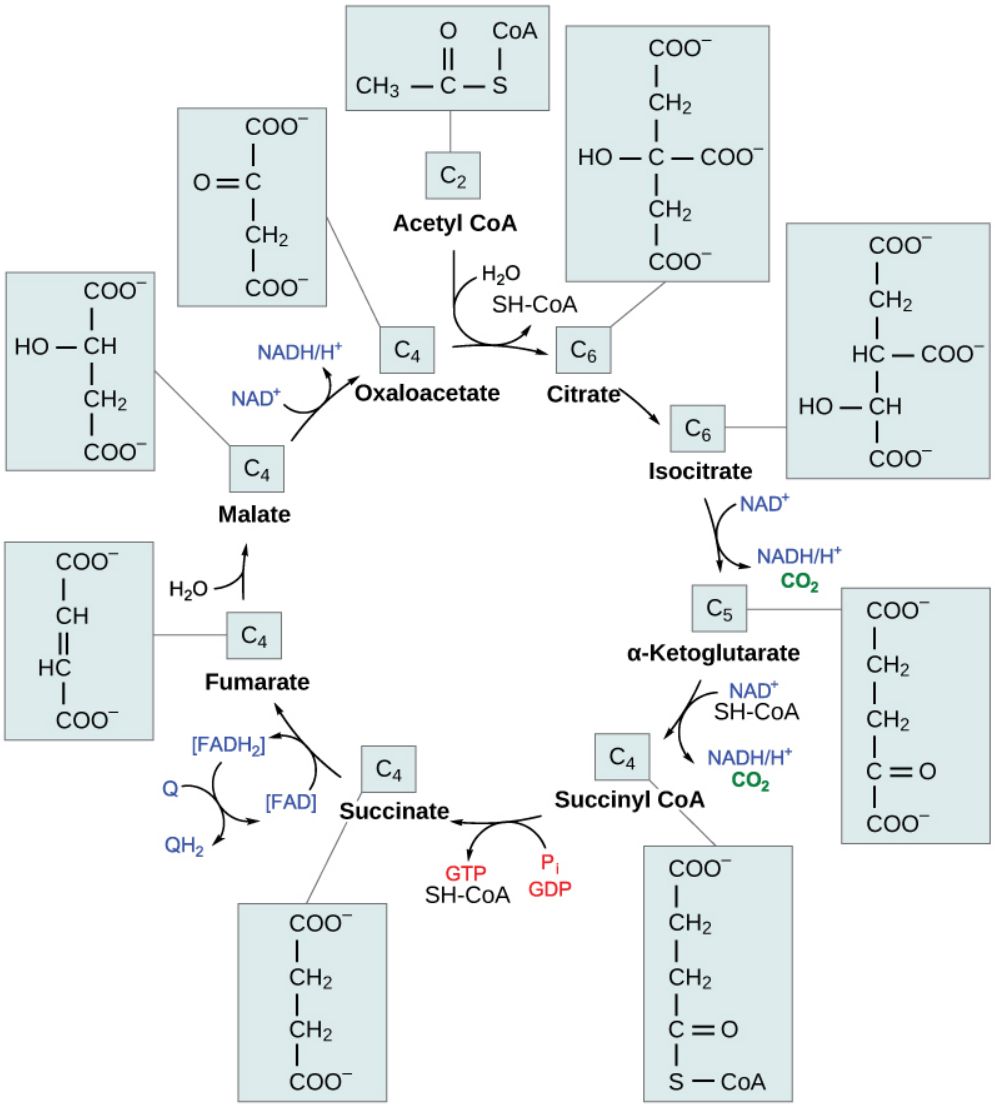
Note that during the second stage of glucose metabolism, whenever a carbon atom is removed, it is bound to two oxygen atoms, producing carbon dioxide, one of the major end products of cellular respiration.

Acetyl CoA to CO_2

In the presence of oxygen, acetyl CoA delivers its acetyl group to a four-carbon molecule, oxaloacetate, to form citrate, a six-carbon molecule with three carboxyl groups; this pathway will harvest the remainder of the extractable energy from what began as a glucose molecule. This single pathway is called by different names: the citric acid cycle (for the first intermediate formed—citric acid, or citrate—when acetate joins to the oxaloacetate), the TCA cycle (since citric acid or citrate and isocitrate are tricarboxylic acids), and the Krebs cycle, after Hans Krebs, who first identified the steps in the pathway in the 1930s in pigeon flight muscles.

Citric Acid Cycle

Like the conversion of pyruvate to acetyl CoA, the citric acid cycle takes place in the matrix of mitochondria. Almost all of the enzymes of the citric acid cycle are soluble, with the single exception of the enzyme succinate dehydrogenase, which is embedded in the inner membrane of the mitochondrion. Unlike glycolysis, the citric acid cycle is a closed loop: The last part of the pathway regenerates the compound used in the first step. The eight steps of the cycle are a series of redox, dehydration, hydration, and decarboxylation reactions that produce two carbon dioxide molecules, one GTP/ATP, and reduced forms of NADH and $FADH_2$ (Figure). This is considered an aerobic pathway because the NADH and $FADH_2$ produced must transfer their electrons to the next pathway in the system, which will use oxygen. If this transfer does not occur, the oxidation steps of the citric acid cycle also do not occur. Note that the citric acid cycle produces very little ATP directly and does not directly consume oxygen.



In the citric acid cycle, the acetyl group from acetyl CoA is attached to a four-carbon oxaloacetate molecule to form a six-carbon citrate molecule. Through a series of steps, citrate is oxidized, releasing two carbon dioxide molecules for each acetyl group fed into the cycle. In the process, three NAD^+ molecules are reduced to NADH, one FAD molecule is reduced to $FADH_2$, and one ATP or GTP (depending on the cell type) is produced (by substrate-level phosphorylation). Because the final product of the citric acid cycle is also the first reactant, the cycle runs continuously in the presence of sufficient reactants. (credit: modification of work by "Yikrazuul"/Wikimedia Commons)

Steps in the Citric Acid Cycle

Step 1. Prior to the start of the first step, a transitional phase occurs during which pyruvic acid is converted to acetyl CoA. Then, the first step of the cycle begins: This is a condensation step, combining the two-carbon acetyl group with a four-carbon oxaloacetate molecule to form a six-carbon molecule of citrate. CoA is bound to a sulfhydryl group (-SH) and diffuses away to eventually combine with another acetyl group. This step is irreversible because it is highly exergonic. The rate of this reaction is controlled by negative feedback and the amount of ATP available. If ATP levels increase, the rate of this reaction decreases. If ATP is in short supply, the rate increases.

Step 2. In step two, citrate loses one water molecule and gains another as citrate is converted into its isomer, isocitrate.

Step 3. In step three, isocitrate is oxidized, producing a five-carbon molecule, α-ketoglutarate, together with a molecule of $CO_2$and two electrons, which reduce NAD^+ to NADH. This step is also regulated by negative feedback from ATP and NADH, and a positive effect of ADP.

Steps 3 and 4. Steps three and four are both oxidation and decarboxylation steps, which release electrons that reduce NAD^+ to NADH and release carboxyl groups that form CO_2 molecules. α-Ketoglutarate is the product of step three, and a succinyl group is the product of step four. CoA binds the succinyl group to form succinyl CoA. The enzyme that catalyzes step four is regulated by feedback inhibition of ATP, succinyl CoA, and NADH.

Step 5. In step five, a phosphate group is substituted for coenzyme A, and a high-energy bond is formed. This energy is used in substrate-level phosphorylation (during the conversion of the succinyl group to succinate) to form either guanine triphosphate (GTP) or ATP. There are two forms of the enzyme, called isoenzymes, for this step, depending upon the type of animal tissue in which they are found. One form is found in tissues that use large amounts of ATP, such as heart and skeletal muscle. This form produces ATP. The second form of the enzyme is found in tissues that have a high number of anabolic pathways, such as liver. This form produces GTP. GTP is energetically equivalent to ATP; however, its use is more restricted. In particular, protein synthesis primarily uses GTP.

Step 6. Step six is a dehydration process that converts succinate into fumarate. Two hydrogen atoms are transferred to FAD, producing $FADH_2$. The energy contained in the electrons of these atoms is insufficient to reduce NAD^+ but adequate to reduce FAD. Unlike NADH, this carrier remains attached to the enzyme and transfers the electrons to the electron transport chain directly. This process is made possible by the localization of the enzyme catalyzing this step inside the inner membrane of the mitochondrion.

Step 7. Water is added to fumarate during step seven, and malate is produced. The last step in the citric acid cycle regenerates oxaloacetate by oxidizing malate. Another molecule of NADH is produced in the process.

Link to Learning

Click through each step of the citric acid cycle here.

Products of the Citric Acid Cycle

Two carbon atoms come into the citric acid cycle from each acetyl group, representing four out of the six carbons of one glucose molecule. Two carbon dioxide molecules are released on each turn of the cycle; however, these do not necessarily contain the most recently added carbon atoms. The two acetyl carbon atoms will eventually be released on later turns of the cycle; thus, all six carbon atoms from the original glucose molecule are eventually incorporated into carbon dioxide. Each turn of the cycle forms three NADH molecules and one $FADH_2$ molecule. These carriers will connect with the last portion of aerobic respiration to produce ATP molecules. One GTP or ATP is also made in each cycle. Several of the intermediate compounds

in the citric acid cycle can be used in synthesizing non-essential amino acids; therefore, the cycle is amphibolic (both catabolic and anabolic).

Section Summary

In the presence of oxygen, pyruvate is transformed into an acetyl group attached to a carrier molecule of coenzyme A. The resulting acetyl CoA can enter several pathways, but most often, the acetyl group is delivered to the citric acid cycle for further catabolism. During the conversion of pyruvate into the acetyl group, a molecule of carbon dioxide and two high-energy electrons are removed. The carbon dioxide accounts for two (conversion of two pyruvate molecules) of the six carbons of the original glucose molecule. The electrons are picked up by NAD^+, and the NADH carries the electrons to a later pathway for ATP production. At this point, the glucose molecule that originally entered cellular respiration has been completely oxidized. Chemical potential energy stored within the glucose molecule has been transferred to electron carriers or has been used to synthesize a few ATPs.

The citric acid cycle is a series of redox and decarboxylation reactions that remove high-energy electrons and carbon dioxide. The electrons temporarily stored in molecules of NADH and $FADH_2$ are used to generate ATP in a subsequent pathway. One molecule of either GTP or ATP is produced by substrate-level phosphorylation on each turn of the cycle. There is no comparison of the cyclic pathway with a linear one.

Review Questions

What is removed from pyruvate during its conversion into an acetyl group?

1. oxygen
2. ATP
3. B vitamin
4. carbon dioxide

What do the electrons added to NAD^+ do?

1. They become part of a fermentation pathway.
2. They go to another pathway for ATP production.
3. They energize the entry of the acetyl group into the citric acid cycle.
4. They are converted to NADP.

GTP or ATP is produced during the conversion of _____.

1. isocitrate into α-ketoglutarate
2. succinyl CoA into succinate
3. fumarate into malate
4. malate into oxaloacetate

How many NADH molecules are produced on each turn of the citric acid cycle?

1. one
2. two
3. three
4. four

Free Response

What is the primary difference between a circular pathway and a linear pathway?

Glossary

acetyl CoA combination of an acetyl group derived from pyruvic acid and coenzyme A, which is made from pantothenic acid (a B-group vitamin)

citric acid cycle (also, Krebs cycle) series of enzyme-catalyzed chemical reactions of central importance in all living cells

Krebs cycle (also, citric acid cycle) alternate name for the citric acid cycle, named after Hans Krebs who first identified the steps in the pathway in the 1930s in pigeon flight muscles; see citric acid cycle

TCA cycle (also, citric acid cycle) alternate name for the citric acid cycle, named after the group name for citric acid, tricarboxylic acid (TCA); see citric acid cycle

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: Open Stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License**: *CC BY: Attribution*

METABOLISM WITHOUT OXYGEN

Learning Objectives

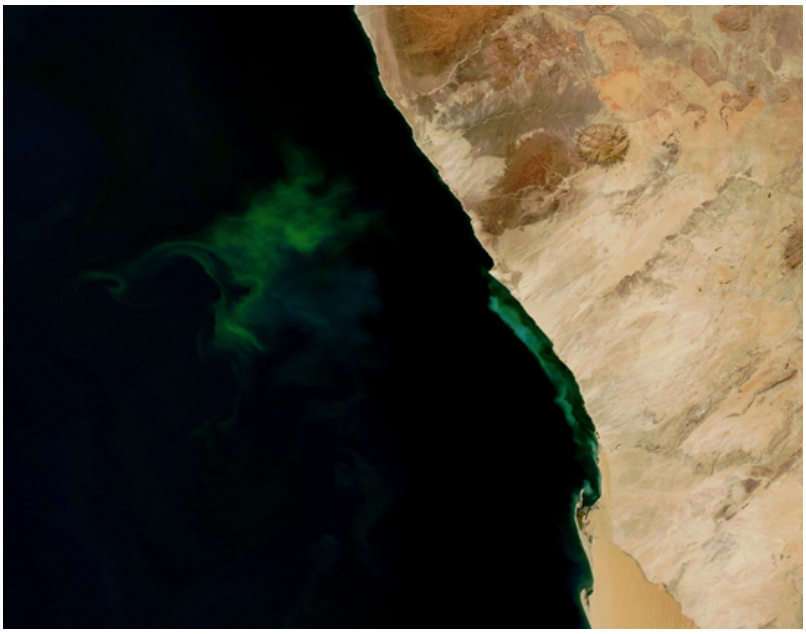
By the end of this section, you will be able to:

- Discuss the fundamental difference between anaerobic cellular respiration and fermentation
- Describe the type of fermentation that readily occurs in animal cells and the conditions that initiate that fermentation

In aerobic respiration, the final electron acceptor is an oxygen molecule, O_2. If aerobic respiration occurs, then ATP will be produced using the energy of high-energy electrons carried by NADH or $FADH_2$ to the electron transport chain. If aerobic respiration does not occur, NADH must be reoxidized to NAD^+ for reuse as an electron carrier for the glycolytic pathway to continue. How is this done? Some living systems use an organic molecule as the final electron acceptor. Processes that use an organic molecule to regenerate NAD^+ from NADH are collectively referred to as fermentation. In contrast, some living systems use an inorganic molecule as a final electron acceptor. Both methods are called anaerobic cellular respiration in which organisms convert energy for their use in the absence of oxygen.

Anaerobic Cellular Respiration

Certain prokaryotes, including some species of bacteria and Archaea, use anaerobic respiration. For example, the group of Archaea called methanogens reduces carbon dioxide to methane to oxidize NADH. These microorganisms are found in soil and in the digestive tracts of ruminants, such as cows and sheep. Similarly, sulfate-reducing bacteria and Archaea, most of which are anaerobic (Figure), reduce sulfate to hydrogen sulfide to regenerate NAD^+ from NADH.



The green color seen in these coastal waters is from an eruption of hydrogen sulfide-producing bacteria. These anaerobic, sulfate-reducing bacteria release hydrogen sulfide gas as they decompose algae in the water. (credit: modification of work by NASA/Jeff Schmaltz, MODIS Land Rapid Response Team at NASA GSFC, Visible Earth Catalog of NASA images)

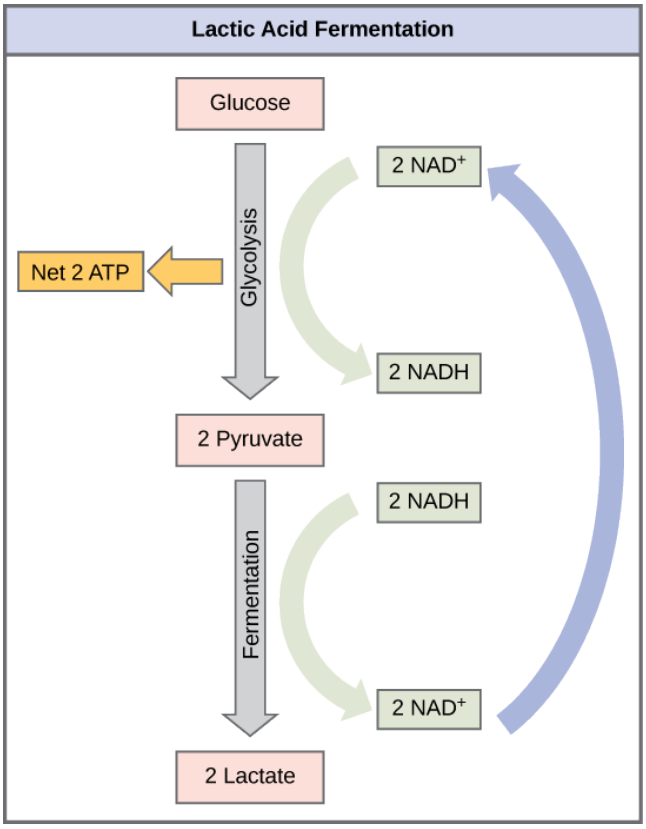
Link to Learning

Visit this site to see anaerobic cellular respiration in action.

Lactic Acid Fermentation

The fermentation method used by animals and certain bacteria, like those in yogurt, is lactic acid fermentation (Figure). This type of fermentation is used routinely in mammalian red blood cells and in skeletal muscle that has an insufficient oxygen supply to allow aerobic respiration to continue (that is, in muscles used to the point of fatigue). In muscles, lactic acid accumulation must be removed by the blood circulation and the lactate brought to the liver for further metabolism. The chemical reactions of lactic acid fermentation are the following:
Pyruvic acid+NADH↔lactic acid+NAD+Pyruvic acid+NADH↔lactic acid+NAD+

The enzyme used in this reaction is lactate dehydrogenase (LDH). The reaction can proceed in either direction, but the reaction from left to right is inhibited by acidic conditions. Such lactic acid accumulation was once believed to cause muscle stiffness, fatigue, and soreness, although more recent research disputes this hypothesis. Once the lactic acid has been removed from the muscle and circulated to the liver, it can be reconverted into pyruvic acid and further catabolized for energy.
ART CONNECTION



Lactic acid fermentation is common in muscle cells that have run out of oxygen.

Tremetol, a metabolic poison found in the white snake root plant, prevents the metabolism of lactate. When cows eat this plant, it is concentrated in the milk they produce. Humans who consume the milk become ill. Symptoms of this disease, which include vomiting, abdominal pain, and tremors, become worse after exercise. Why do you think this is the case?

Alcohol Fermentation

Another familiar fermentation process is alcohol fermentation (Figure) that produces ethanol, an alcohol. The first chemical reaction of alcohol fermentation is the following (CO_2 does not participate in the second reaction):

Pyruvic acid→CO2+acetaldehyde+NADH→ethanol+NAD+Pyruvic acid→CO2+acetaldehyde+NADH→ethanol+NAD+

The first reaction is catalyzed by pyruvate decarboxylase, a cytoplasmic enzyme, with a coenzyme of thiamine pyrophosphate (TPP, derived from vitamin B_1 and also called thiamine). A carboxyl group is removed from pyruvic acid, releasing carbon dioxide as a gas. The loss of carbon dioxide reduces the size of the molecule by one carbon, making acetaldehyde. The second reaction is catalyzed by alcohol dehydrogenase to oxidize NADH to NAD^+ and reduce acetaldehyde to ethanol. The fermentation of pyruvic acid by yeast produces the ethanol found in alcoholic beverages. Ethanol tolerance of yeast is variable, ranging from about 5 percent to 21 percent, depending on the yeast strain and environmental conditions.

Fermentation of grape juice into wine produces CO_2 as a byproduct. Fermentation tanks have valves so that the pressure inside the tanks created by the carbon dioxide produced can be released.

Other Types of Fermentation

Other fermentation methods occur in bacteria. Many prokaryotes are facultatively anaerobic. This means that they can switch between aerobic respiration and fermentation, depending on the availability of oxygen. Certain prokaryotes, like *Clostridia*, are obligate anaerobes. Obligate anaerobes live and grow in the absence of molecular oxygen. Oxygen is a poison to these microorganisms and kills them on exposure. It should be noted that all forms of fermentation, except lactic acid fermentation, produce gas. The production of particular types of gas is used as an indicator of the fermentation of specific carbohydrates, which plays a role in the laboratory identification of the bacteria. Various methods of fermentation are used by assorted organisms to ensure an adequate supply of NAD^+ for the sixth step in glycolysis. Without these pathways, that step would not occur and no ATP would be harvested from the breakdown of glucose.

Section Summary

If NADH cannot be oxidized through aerobic respiration, another electron acceptor is used. Most organisms will use some form of fermentation to accomplish the regeneration of NAD^+, ensuring the continuation of glycolysis. The regeneration of NAD^+ in fermentation is not accompanied by ATP production; therefore, the potential of NADH to produce ATP using an electron transport chain is not utilized.

Art Connections

Figure Tremetol, a metabolic poison found in the white snake root plant, prevents the metabolism of lactate. When cows eat this plant, it is concentrated in the milk they produce. Humans who consume the milk become ill. Symptoms of this disease, which include vomiting, abdominal pain, and tremors, become worse after exercise. Why do you think this is the case?

Review Questions

Which of the following fermentation methods can occur in animal skeletal muscles?

1. lactic acid fermentation
2. alcohol fermentation
3. mixed acid fermentation
4. propionic fermentation

Free Response

What is the primary difference between fermentation and anaerobic respiration?

Glossary

anaerobic cellular respiration process in which organisms convert energy for their use in the absence of oxygen

fermentation process of regenerating NAD^+ with either an inorganic or organic compound serving as the final electron acceptor, occurs in the absence; occurs in the absence of oxygen

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: Open Stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License**: *CC BY: Attribution*

CONNECTIONS OF CARBOHYDRATE, PROTEIN, AND LIPID METABOLIC PATHWAYS

Learning Objectives

By the end of this section, you will be able to:

- Discuss the ways in which carbohydrate metabolic pathways, glycolysis, and the citric acid cycle interrelate with protein and lipid metabolic pathways
- Explain why metabolic pathways are not considered closed systems

You have learned about the catabolism of glucose, which provides energy to living cells. But living things consume more than glucose for food. How does a turkey sandwich end up as ATP in your cells? This happens because all of the catabolic pathways for carbohydrates, proteins, and lipids eventually connect into glycolysis and the citric acid cycle pathways (see Figure). Metabolic pathways should be thought of as porous—that is, substances enter from other pathways, and intermediates leave for other pathways. These pathways are not closed systems. Many of the substrates, intermediates, and products in a particular pathway are reactants in other pathways.

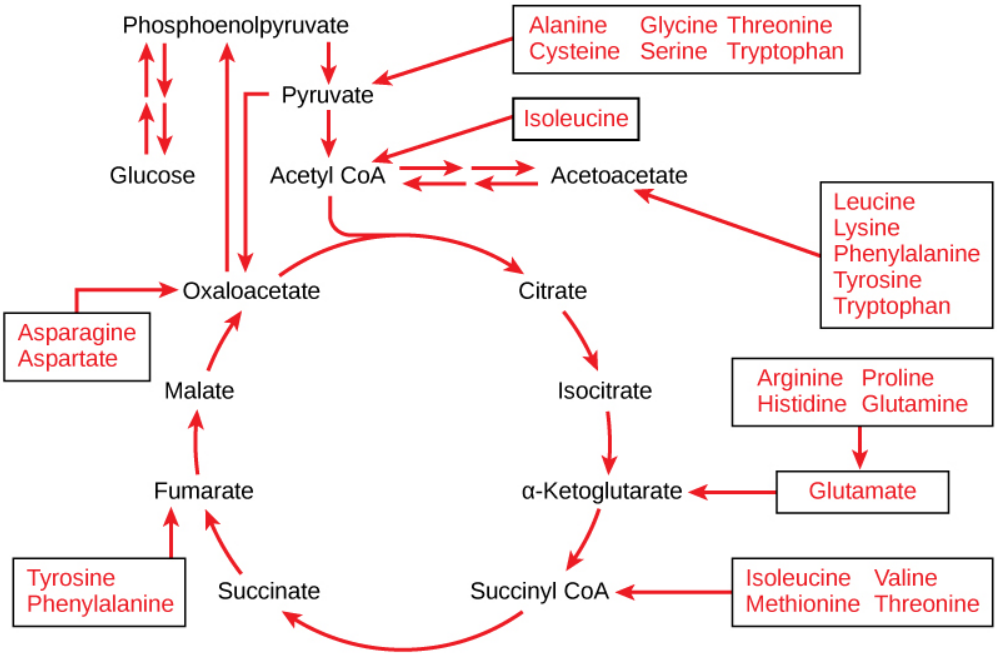
Connections of Other Sugars to Glucose Metabolism

Glycogen, a polymer of glucose, is an energy storage molecule in animals. When there is adequate ATP present, excess glucose is shunted into glycogen for storage. Glycogen is made and stored in both liver and muscle. The glycogen will be hydrolyzed into glucose monomers (G-1-P) if blood sugar levels drop. The presence of glycogen as a source of glucose allows ATP to be produced for a longer period of time during exercise. Glycogen is broken down into G-1-P and converted into G-6-P in both muscle and liver cells, and this product enters the glycolytic pathway.

Sucrose is a disaccharide with a molecule of glucose and a molecule of fructose bonded together with a glycosidic linkage. Fructose is one of the three dietary monosaccharides, along with glucose and galactose (which is part of the milk sugar, the disaccharide lactose), which are absorbed directly into the bloodstream during digestion. The catabolism of both fructose and galactose produces the same number of ATP molecules as glucose.

Connections of Proteins to Glucose Metabolism

Proteins are hydrolyzed by a variety of enzymes in cells. Most of the time, the amino acids are recycled into the synthesis of new proteins. If there are excess amino acids, however, or if the body is in a state of starvation, some amino acids will be shunted into the pathways of glucose catabolism (Figure). Each amino acid must have its amino group removed prior to entry into these pathways. The amino group is converted into ammonia. In mammals, the liver synthesizes urea from two ammonia molecules and a carbon dioxide molecule. Thus, urea is the principal waste product in mammals produced from the nitrogen originating in amino acids, and it leaves the body in urine.

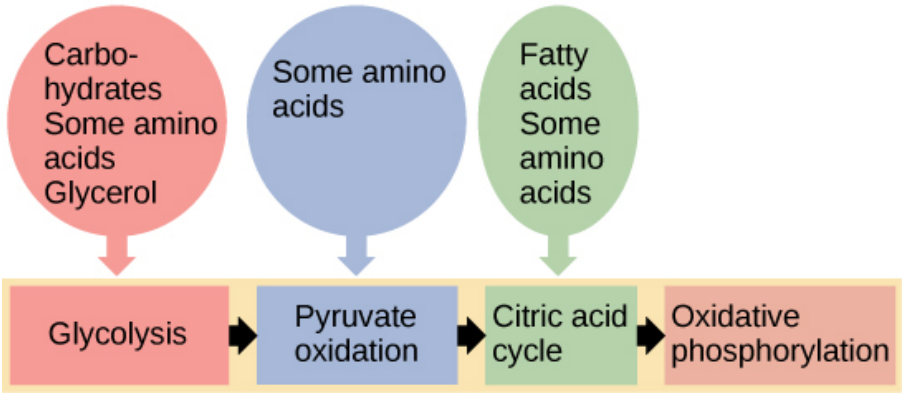


The carbon skeletons of certain amino acids (indicated in boxes) derived from proteins can feed into the citric acid cycle. (credit: modification of work by Mikael Häggström)

Connections of Lipid and Glucose Metabolisms

The lipids that are connected to the glucose pathways are cholesterol and triglycerides. Cholesterol is a lipid that contributes to cell membrane flexibility and is a precursor of steroid hormones. The synthesis of cholesterol starts with acetyl groups and proceeds in only one direction. The process cannot be reversed.

Triglycerides are a form of long-term energy storage in animals. Triglycerides are made of glycerol and three fatty acids. Animals can make most of the fatty acids they need. Triglycerides can be both made and broken down through parts of the glucose catabolism pathways. Glycerol can be phosphorylated to glycerol-3-phosphate, which continues through glycolysis. Fatty acids are catabolized in a process called beta-oxidation that takes place in the matrix of the mitochondria and converts their fatty acid chains into two carbon units of acetyl groups. The acetyl groups are picked up by CoA to form acetyl CoA that proceeds into the citric acid cycle.



Glycogen from the liver and muscles, hydrolyzed into glucose-1-phosphate, together with fats and proteins, can feed into the catabolic pathways for carbohydrates.
EVOLUTION CONNECTION

Pathways of Photosynthesis and Cellular MetabolismThe processes of photosynthesis and cellular metabolism consist of several very complex pathways. It is generally thought that the first cells arose in an aqueous environment—a “soup” of nutrients—probably on the surface of some porous clays. If these cells reproduced successfully and their numbers climbed steadily, it follows that the cells would begin to deplete the nutrients from the medium in which they lived as they shifted the nutrients into the components of their own bodies. This hypothetical situation would have resulted in natural selection favoring those organisms that could exist by using the nutrients that remained in their environment and by manipulating these nutrients into materials upon which they could survive. Selection would favor those organisms that could extract maximal value from the nutrients to which they had access.

An early form of photosynthesis developed that harnessed the sun’s energy using water as a source of hydrogen atoms, but this pathway did not produce free oxygen (anoxygenic photosynthesis). (Early photosynthesis did not produce free oxygen because it did not use water as the source of hydrogen ions; instead, it used materials like hydrogen sulfide and consequently produced sulfur). It is thought that glycolysis developed at this time and could take advantage of the simple sugars being produced, but these reactions were unable to fully extract the energy stored in the carbohydrates. The development of glycolysis probably predated the evolution of photosynthesis, as it was well suited to extract energy from materials spontaneously accumulating in the “primeval soup.” A later form of photosynthesis used water as a source of electrons and hydrogen, and generated free oxygen. Over time, the atmosphere became oxygenated, but not before the oxygen released oxidized metals in the ocean and created a “rust” layer in the sediment, permitting the dating of the rise of the first oxygenic photosynthesizers. Living things adapted to exploit this new atmosphere that allowed aerobic respiration as we know it to evolve. When the full process of oxygenic photosynthesis developed and the atmosphere became oxygenated, cells were finally able to use the oxygen expelled by photosynthesis to extract considerably more energy from the sugar molecules using the citric acid cycle and oxidative phosphorylation.

Section Summary

The breakdown and synthesis of carbohydrates, proteins, and lipids connect with the pathways of glucose catabolism. The simple sugars are galactose, fructose, glycogen, and pentose. These are catabolized during glycolysis. The amino acids from proteins connect with glucose catabolism through pyruvate, acetyl CoA, and components of the citric acid cycle. Cholesterol synthesis starts with acetyl groups, and the components of triglycerides come from glycerol-3-phosphate from glycolysis and acetyl groups produced in the mitochondria from pyruvate.

Review Questions

A major connection for sugars in glycolysis is _____.

1. glucose-6-phosphate
2. fructose-1,6-bisphosphate
3. dihydroxyacetone phosphate
4. phosphoenolpyruvate

Beta-oxidation is _____.

1. the breakdown of sugars
2. the assembly of sugars
3. the breakdown of fatty acids
4. the removal of amino groups from amino acids

Free Response

Would you describe metabolic pathways as inherently wasteful or inherently economical, and why?

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: Open Stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License**: *CC BY: Attribution*

STUDY GUIDE: CELLULAR RESPIRATION

Study Questions

Objective: Describe the process by which energy stored in a glucose molecule can be used by a cell.

Use this page to check your understanding of the content.

Study Guide Questions

1. What is the chemical equation that describes cellular respiration?
2. What are the “reactants” in the chemical equation from question 1? What are the products?
3. Where does cellular respiration take place? Be sure you know where each individual stage occurs!

4. Name and describe the purpose of the 2 electron carriers that participate in cellular respiration.
5. Be able to do “energy accounting” for each stage of cellular respiration. Account for all electron carriers and ATP molecules produced.
6. Compare and contrast the 3 stages of cellular respiration.
7. Clearly explain the importance of OXYGEN in cellular respiration.
8. Where does the water come from that is produced in cellular respiration?
9. Where does the carbon dioxide come from that is produced during cellular respiration?
10. Where does the sugar come from that is used during cellular respiration?
11. Where does the oxygen come from that is used during cellular respiration?
12. Exactly how is ATP made in the electron transport chain?
13. What role do hydrogen ions play in the process of energy production?

Licensing & Attributions

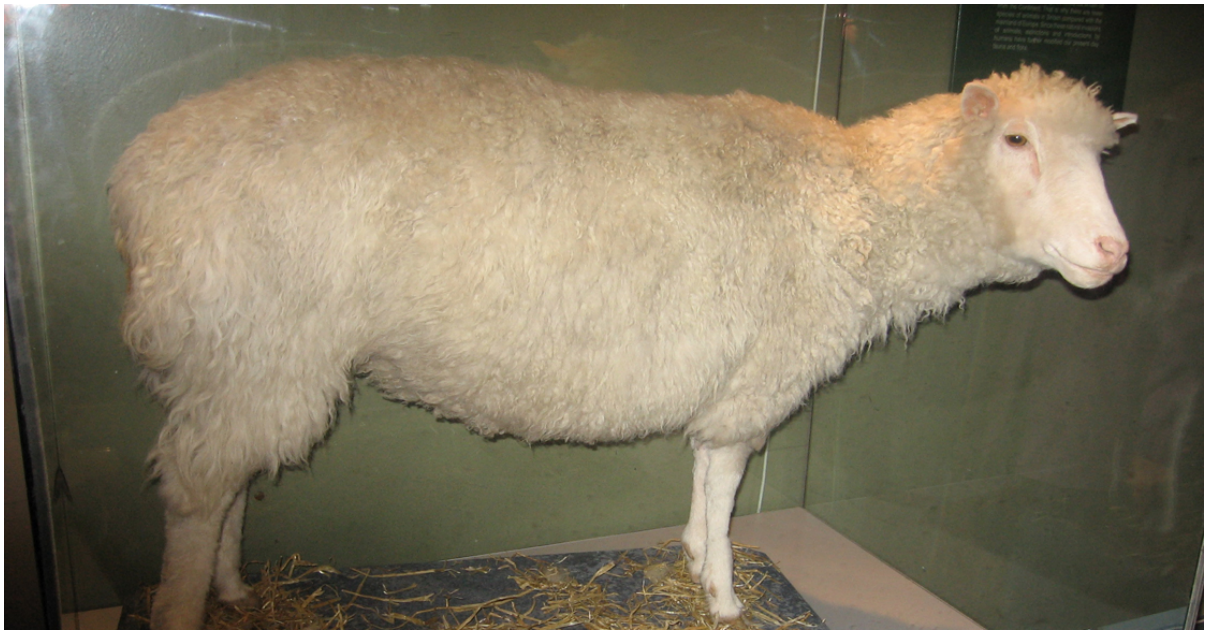
CC licensed content, Shared previously

- Study Guide: Cellular Respiration. **Authored by**: Wendy Riggs. **Provided by**: College of the Redwoods. **Project**: Kaleidoscope. **License**: *CC BY: Attribution*

DNA (STRUCTURE AND FUNCTION)

INTRODUCTION: DNA

Chapter 14. Introduction



Dolly the sheep was the first large mammal to be cloned.

The three letters “DNA” have now become synonymous with crime solving, paternity testing, human identification, and genetic testing. DNA can be retrieved from hair, blood, or saliva. Each person’s DNA is unique, and it is possible to detect differences between individuals within a species on the basis of these unique features.

DNA analysis has many practical applications beyond forensics. In humans, DNA testing is applied to numerous uses: determining paternity, tracing genealogy, identifying pathogens, archeological research, tracing disease outbreaks, and studying human migration patterns. In the medical field, DNA is used in diagnostics, new vaccine development, and cancer therapy. It is now possible to determine predisposition to diseases by looking at genes.

Each human cell has 23 pairs of chromosomes: one set of chromosomes is inherited from the mother and the other set is inherited from the father. There is also a mitochondrial genome, inherited exclusively from the mother, which can be involved in inherited genetic disorders. On each chromosome, there are thousands of genes that are responsible for determining the genotype and phenotype of the individual. A gene is defined as a sequence of DNA that codes for a functional product. The human haploid genome contains 3 billion base pairs and has between 20,000 and 25,000 functional genes.

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: OpenStax College. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License**: *CC BY: Attribution*

HISTORICAL BASIS OF OUR MODERN UNDERSTANDING OF DNA

Modern understandings of DNA have evolved from the discovery of nucleic acid to the development of the double-helix model. In the 1860s, Friedrich Miescher (Figure), a physician by profession, was the first person to isolate phosphate-rich chemicals from white blood cells or leukocytes. He named these chemicals (which would eventually be known as RNA and DNA) nuclein because they were isolated from the nuclei of the cells.

Friedrich Miescher (1844–1895) discovered nucleic acids.

Link to Learning

To see Miescher conduct an experiment step-by-step, click through this review of how he discovered the key role of DNA and proteins in the nucleus.

A half century later, British bacteriologist Frederick Griffith was perhaps the first person to show that hereditary information could be transferred from one cell to another "horizontally," rather than by descent. In 1928, he reported the first demonstration of bacterial transformation, a process in which external DNA is taken up by a cell, thereby changing morphology and physiology. He was working with *Streptococcus pneumoniae,* the bacterium that causes pneumonia. Griffith worked with two strains, rough (R) and smooth (S). The R strain is non-pathogenic (does not cause disease) and is called rough because its outer surface is a cell wall and lacks a capsule; as a result, the cell surface appears uneven under the microscope. The S strain is pathogenic (disease-causing) and has a capsule outside its cell wall. As a result, it has a smooth appearance under the microscope. Griffith injected the live R strain into mice and they survived. In another experiment, when he injected mice with the heat-killed S strain, they also survived. In a third set of experiments, a mixture of live R strain and heat-killed S strain were injected into mice, and—to his surprise—the mice died. Upon isolating the live bacteria from the dead mouse, only the S strain of bacteria was recovered. When this isolated S strain was injected into fresh mice, the mice died. Griffith concluded that something had passed from the heat-killed S strain into the live R strain and transformed it into the pathogenic S strain, and he called this the transforming principle (Figure). These experiments are now famously known as Griffith's transformation experiments.

Mouse injected with heat-killed virulant S strain lives.

Mouse injected with both heat-killed S strain and live non-virulant R strain dies.

Two strains of *S. pneumoniae* were used in Griffith's transformation experiments. The R strain is non-pathogenic. The S strain is pathogenic and causes death. When Griffith injected a mouse with the heat-killed S strain and a live R strain, the mouse died. The S strain was recovered from the dead mouse. Thus, Griffith concluded that something had passed from the heat-killed S strain to the R strain, transforming the R strain into S strain in the process. (credit "living mouse": modification of work by NIH; credit "dead mouse": modification of work by Sarah Marriage)

Scientists Oswald Avery, Colin MacLeod, and Maclyn McCarty (1944) were interested in exploring this transforming principle further. They isolated the S strain from the dead mice and isolated the proteins and nucleic acids, namely RNA and DNA, as these were possible candidates for the molecule of heredity. They conducted a systematic elimination study. They used enzymes that specifically degraded each component and then used each mixture separately to transform the R strain. They found that when DNA was degraded, the resulting mixture was no longer able to transform the bacteria, whereas all of the other combinations were able to transform the bacteria. This led them to conclude that DNA was the transforming principle.

CAREER CONNECTION

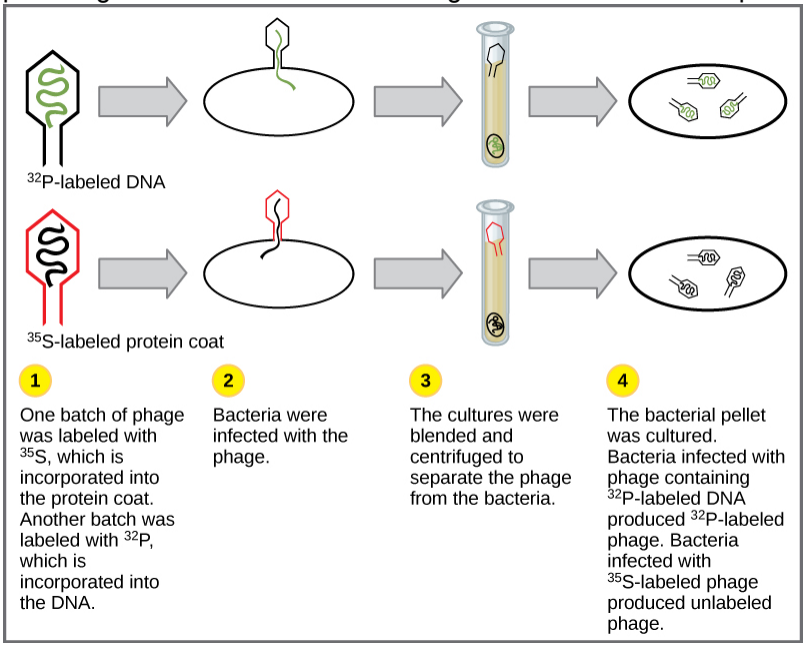
Forensic Scientists and DNA AnalysisDNA evidence was used for the first time to solve an immigration case. The story started with a teenage boy returning to London from Ghana to be with his mother. Immigration authorities at the airport were suspicious of him, thinking that he was traveling on a forged passport. After much persuasion, he was allowed to go live with his mother, but the immigration authorities did not drop the case against him. All types of evidence, including photographs, were provided to the authorities, but deportation proceedings were started nevertheless. Around the same time, Dr. Alec Jeffreys of Leicester University in the United Kingdom had invented a technique known as DNA fingerprinting. The immigration authorities approached Dr. Jeffreys for help. He took DNA samples from the mother and three of her children, plus an unrelated mother, and compared the samples with the boy's DNA. Because the biological father was not in the picture, DNA from the three children was compared with the boy's DNA. He found a match in the boy's DNA for both the mother and his three siblings. He concluded that the boy was indeed the mother's son.

Forensic scientists analyze many items, including documents, handwriting, firearms, and biological samples. They analyze the DNA content of hair, semen, saliva, and blood, and compare it with a database of DNA profiles of known criminals. Analysis includes DNA isolation, sequencing, and sequence analysis; most forensic DNA analysis involves polymerase chain reaction (PCR) amplification of short tandem repeat (STR) loci and electrophoresis to determine the length of the PCR-amplified fragment. Only mitochondrial DNA is sequenced for forensics. Forensic scientists are expected to appear at court hearings to present their findings. They are usually employed in crime labs of city and state government agencies. Geneticists experimenting with DNA techniques also work for scientific and research organizations, pharmaceutical industries, and college and university labs. Students wishing to pursue a career as a forensic scientist should have at least a bachelor's degree in chemistry, biology, or physics, and preferably some experience working in a laboratory.

Experiments conducted by Martha Chase and Alfred Hershey in 1952 provided confirmatory evidence that DNA was the genetic material and not proteins. Chase and Hershey were studying a bacteriophage, which is a virus that infects bacteria. Viruses typically have a simple structure: a protein coat, called the capsid, and a nucleic acid core that contains the genetic material, either DNA or RNA. The bacteriophage infects the host bacterial cell by attaching to its surface, and then it injects its nucleic acids inside the cell. The phage DNA makes multiple copies of itself using the host machinery, and eventually the host cell bursts, releasing a large number of bacteriophages. Hershey and Chase labeled one batch of phage with radioactive sulfur, ^{35}S, to label the protein coat. Another

batch of phage were labeled with radioactive phosphorus, ^{32}P. Because phosphorous is found in DNA, but not protein, the DNA and not the protein would be tagged with radioactive phosphorus.

Each batch of phage was allowed to infect the cells separately. After infection, the phage bacterial suspension was put in a blender, which caused the phage coat to be detached from the host cell. The phage and bacterial suspension was spun down in a centrifuge. The heavier bacterial cells settled down and formed a pellet, whereas the lighter phage particles stayed in the supernatant. In the tube that contained phage labeled with ^{35}S, the supernatant contained the radioactively labeled phage, whereas no radioactivity was detected in the pellet. In the tube that contained the phage labeled with ^{32}P, the radioactivity was detected in the pellet that contained the heavier bacterial cells, and no radioactivity was detected in the supernatant. Hershey and Chase concluded that it was the phage DNA that was injected into the cell and carried information to produce more phage particles, thus providing evidence that DNA was the genetic material and not proteins (Figure).



In Hershey and Chase's experiments, bacteria were infected with phage radiolabeled with either ^{35}S, which labels protein, or^{32}P, which labels DNA. Only ^{32}P entered the bacterial cells, indicating that DNA is the genetic material.

Around this same time, Austrian biochemist Erwin Chargaff examined the content of DNA in different species and found that the amounts of adenine, thymine, guanine, and cytosine were not found in equal quantities, and that it varied from species to species, but not between individuals of the same species. He found that the amount of adenine equals the amount of thymine, and the amount of cytosine equals the amount of guanine, or A = T and G = C. This is also known as Chargaff's rules. This finding proved immensely useful when Watson and Crick were getting ready to propose their DNA double helix model.

Section Summary

DNA was first isolated from white blood cells by Friedrich Miescher, who called it nuclein because it was isolated from nuclei. Frederick Griffith's experiments with strains of *Streptococcus pneumoniae* provided the first hint that DNA may be the transforming principle. Avery, MacLeod, and McCarty proved that DNA is required for the transformation of bacteria. Later experiments by Hershey and Chase using bacteriophage T2 proved that DNA is the genetic material. Chargaff found that the ratio of A = T and C = G, and that the percentage content of A, T, G, and C is different for different species.

Review Questions

If DNA of a particular species was analyzed and it was found that it contains 27 percent A, what would be the percentage of C?

1. 27 percent
2. 30 percent
3. 23 percent
4. 54 percent

The experiments by Hershey and Chase helped confirm that DNA was the hereditary material on the basis of the finding that:

1. radioactive phage were found in the pellet
2. radioactive cells were found in the supernatant
3. radioactive sulfur was found inside the cell
4. radioactive phosphorus was found in the cell

Free Response

Explain Griffith's transformation experiments. What did he conclude from them?

Why were radioactive sulfur and phosphorous used to label bacteriophage in Hershey and Chase's experiments?

Glossary

transformation process in which external DNA is taken up by a cell

Licensing & Attributions

CC licensed content, Shared previously

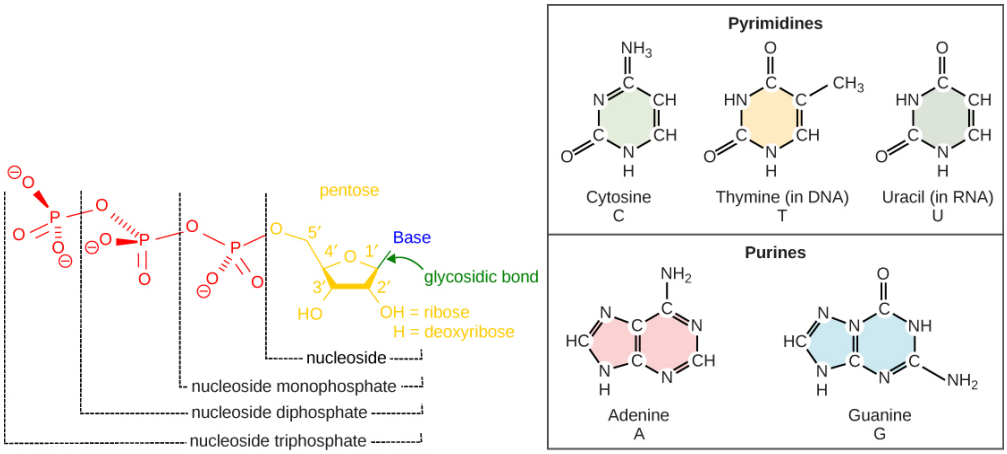
- Biology. **Authored by**: OpenStax College. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License**: *CC BY: Attribution*

DNA STRUCTURE AND SEQUENCING

14.2 DNA Structure and Sequencing

Summary

The building blocks of DNA are nucleotides. The important components of the nucleotide are a nitrogenous base, deoxyribose (5-carbon sugar), and a phosphate group (Figure). The nucleotide is named depending on the nitrogenous base. The nitrogenous base can be a purine such as adenine (A) and guanine (G), or a pyrimidine such as cytosine (C) and thymine (T).



Each nucleotide is made up of a sugar, a phosphate group, and a nitrogenous base. The sugar is deoxyribose in DNA and ribose in RNA.

The nucleotides combine with each other by covalent bonds known as phosphodiester bonds or linkages. The purines have a double ring structure with a six-membered ring fused to a five-membered ring. Pyrimidines are smaller in size; they have a single six-membered ring structure. The carbon atoms of the five-carbon sugar are numbered 1′, 2′, 3′, 4′, and 5′ (1′ is read as “one prime”). The phosphate residue is attached to the hydroxyl group of the 5′ carbon of one sugar of one nucleotide and the hydroxyl group of the 3′ carbon of the sugar of the next nucleotide, thereby forming a 5′-3′ phosphodiester bond.

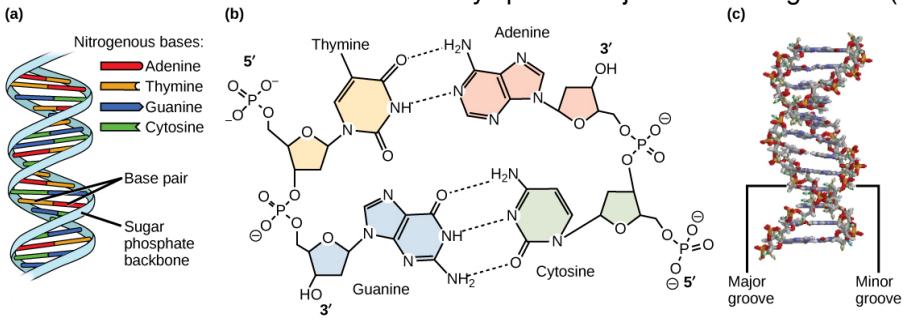
In the 1950s, Francis Crick and James Watson worked together to determine the structure of DNA at the University of Cambridge, England. Other scientists like Linus Pauling and Maurice Wilkins were also actively exploring this field. Pauling had discovered the secondary structure of proteins using X-ray crystallography. In Wilkins’ lab, researcher Rosalind Franklin was using X-ray diffraction methods to understand the structure of DNA. Watson and Crick were able to piece together the puzzle of the DNA molecule on the basis of Franklin’s data because Crick had also studied X-ray diffraction (Figure). In 1962, James Watson, Francis Crick, and Maurice Wilkins were awarded the Nobel Prize in Medicine. Unfortunately, by then Franklin had died, and Nobel prizes are not awarded posthumously.

(a) (b)

The work of pioneering scientists (a) James Watson, Francis Crick, and Maclyn McCarty led to our present day understanding of DNA. Scientist Rosalind Franklin discovered (b) the X-ray diffraction pattern of DNA, which helped to elucidate its double helix structure. (credit a: modification of work by Marjorie McCarty, Public Library of Science)

Watson and Crick proposed that DNA is made up of two strands that are twisted around each other to form a right-handed helix. Base pairing takes place between a purine and pyrimidine; namely, A pairs with T and G pairs with C. Adenine and thymine are complementary base pairs, and cytosine and guanine are also complementary base pairs. The base pairs are stabilized by hydrogen bonds; adenine and thymine form two hydrogen bonds and cytosine and guanine form three hydrogen bonds. The two strands are anti-parallel in nature; that is, the 3′ end of

one strand faces the 5′ end of the other strand. The sugar and phosphate of the nucleotides form the backbone of the structure, whereas the nitrogenous bases are stacked inside. Each base pair is separated from the other base pair by a distance of 0.34 nm, and each turn of the helix measures 3.4 nm. Therefore, ten base pairs are present per turn of the helix. The diameter of the DNA double helix is 2 nm, and it is uniform throughout. Only the pairing between a purine and pyrimidine can explain the uniform diameter. The twisting of the two strands around each other results in the formation of uniformly spaced major and minor grooves (Figure).



DNA has (a) a double helix structure and (b) phosphodiester bonds. The (c) major and minor grooves are binding sites for DNA binding proteins during processes such as transcription (the copying of RNA from DNA) and replication.

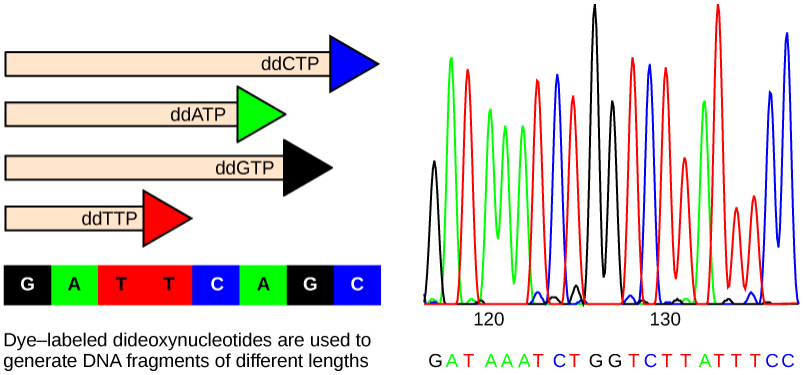
DNA Sequencing Techniques

Until the 1990s, the sequencing of DNA (reading the sequence of DNA) was a relatively expensive and long process. Using radiolabeled nucleotides also compounded the problem through safety concerns. With currently available technology and automated machines, the process is cheap, safer, and can be completed in a matter of hours. Fred Sanger developed the sequencing method used for the human genome sequencing project, which is widely used today (Figure).

Link to Learning

Visit this site to watch a video explaining the DNA sequence reading technique that resulted from Sanger's work.

The method is known as the dideoxy chain termination method. The sequencing method is based on the use of chain terminators, the dideoxynucleotides (ddNTPs). The dideoxynucleotides, or ddNTPSs, differ from the deoxynucleotides by the lack of a free 3′ OH group on the five-carbon sugar. If a ddNTP is added to a growing a DNA strand, the chain is not extended any further because the free 3′ OH group needed to add another nucleotide is not available. By using a predetermined ratio of deoxyribonucleotides to dideoxynucleotides, it is possible to generate DNA fragments of different sizes.



Dye–labeled dideoxynucleotides are used to generate DNA fragments of different lengths

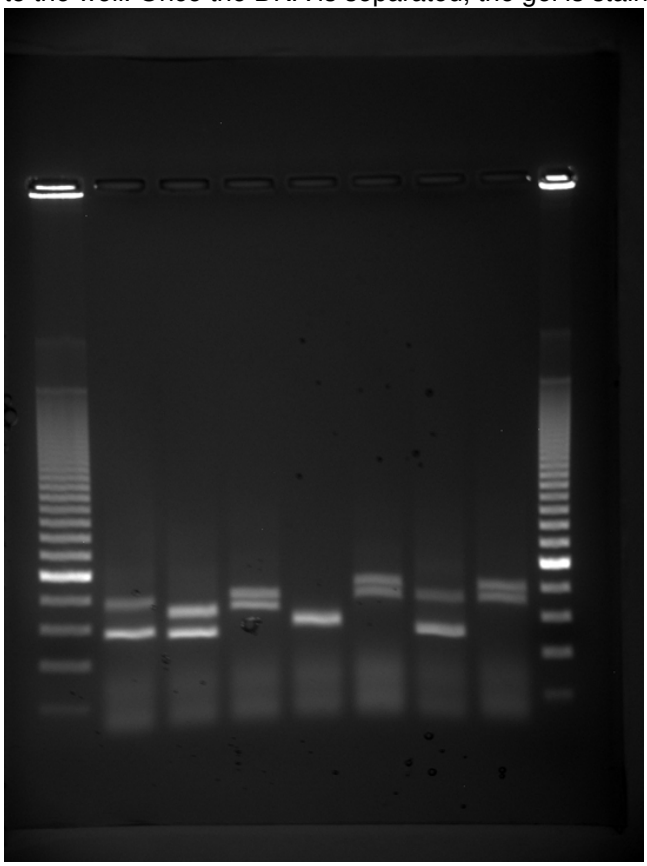
In Frederick Sanger's dideoxy chain termination method, dye-labeled dideoxynucleotides are used to generate DNA fragments that terminate at different points. The DNA is separated by capillary electrophoresis on the basis of size, and from the order of fragments formed, the DNA sequence can be read. The DNA sequence readout is shown on an electropherogram that is generated by a laser scanner.

The DNA sample to be sequenced is denatured or separated into two strands by heating it to high temperatures. The DNA is divided into four tubes in which a primer, DNA polymerase, and all four nucleotides (A, T, G, and C) are added. In addition to each of the four tubes, limited quantities of one of the four dideoxynucleotides are added to each tube respectively. The tubes are labeled as A, T, G, and C according to the ddNTP added. For detection purposes, each of the four dideoxynucleotides carries a different fluorescent label. Chain elongation continues until a fluorescent dideoxy nucleotide is incorporated, after which no further elongation takes place. After the reaction is over, electrophoresis is performed. Even a difference in length of a single base can be detected. The sequence is read from a laser scanner. For his work on DNA sequencing, Sanger received a Nobel Prize in chemistry in 1980.

Link to Learning

Sanger's genome sequencing has led to a race to sequence human genomes at a rapid speed and low cost, often referred to as the $1000 in one day sequence. Learn more by selecting the Sequencing at Speed animation here.

Gel electrophoresis is a technique used to separate DNA fragments of different sizes. Usually the gel is made of a chemical called agarose. Agarose powder is added to a buffer and heated. After cooling, the gel solution is poured into a casting tray. Once the gel has solidified, the DNA is loaded on the gel and electric current is applied. The DNA has a net negative charge and moves from the negative electrode toward the positive electrode. The electric current is applied for sufficient time to let the DNA separate according to size; the smallest fragments will be farthest from the well (where the DNA was loaded), and the heavier molecular weight fragments will be closest to the well. Once the DNA is separated, the gel is stained with a DNA-specific dye for viewing it (Figure).



DNA can be separated on the basis of size using gel electrophoresis. (credit: James Jacob, Tompkins Cortland Community College)

EVOLUTION CONNECTION

Neanderthal Genome: How Are We Related?The first draft sequence of the Neanderthal genome was recently published by Richard E. Green et al. in 2010.[1] Neanderthals are the closest ancestors of present-day humans. They were known to have lived in Europe and Western Asia before they disappeared from fossil records

approximately 30,000 years ago. Green's team studied almost 40,000-year-old fossil remains that were selected from sites across the world. Extremely sophisticated means of sample preparation and DNA sequencing were employed because of the fragile nature of the bones and heavy microbial contamination. In their study, the scientists were able to sequence some four billion base pairs. The Neanderthal sequence was compared with that of present-day humans from across the world. After comparing the sequences, the researchers found that the Neanderthal genome had 2 to 3 percent greater similarity to people living outside Africa than to people in Africa. While current theories have suggested that all present-day humans can be traced to a small ancestral population in Africa, the data from the Neanderthal genome may contradict this view. Green and his colleagues also discovered DNA segments among people in Europe and Asia that are more similar to Neanderthal sequences than to other contemporary human sequences. Another interesting observation was that Neanderthals are as closely related to people from Papua New Guinea as to those from China or France. This is surprising because Neanderthal fossil remains have been located only in Europe and West Asia. Most likely, genetic exchange took place between Neanderthals and modern humans as modern humans emerged out of Africa, before the divergence of Europeans, East Asians, and Papua New Guineans.

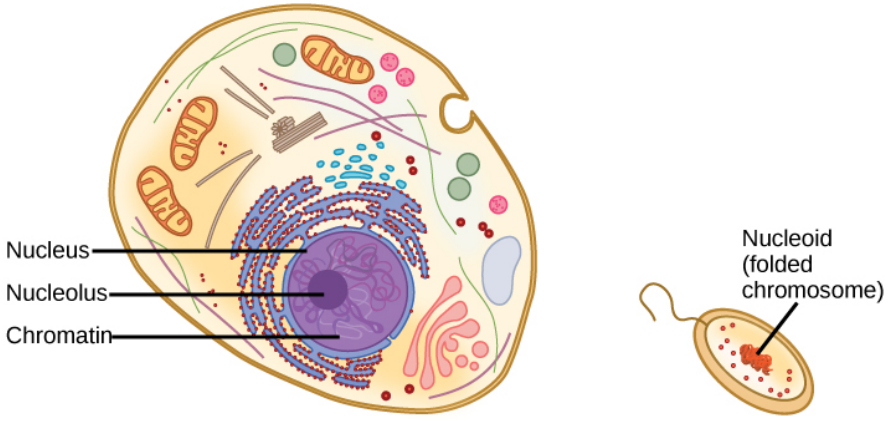
Several genes seem to have undergone changes from Neanderthals during the evolution of present-day humans. These genes are involved in cranial structure, metabolism, skin morphology, and cognitive development. One of the genes that is of particular interest is *RUNX2*, which is different in modern day humans and Neanderthals. This gene is responsible for the prominent frontal bone, bell-shaped rib cage, and dental differences seen in Neanderthals. It is speculated that an evolutionary change in *RUNX2* was important in the origin of modern-day humans, and this affected the cranium and the upper body.

Link to Learning

Watch Svante Pääbo's talk explaining the Neanderthal genome research at the 2011 annual TED (Technology, Entertainment, Design) conference.

DNA Packaging in Cells

When comparing prokaryotic cells to eukaryotic cells, prokaryotes are much simpler than eukaryotes in many of their features (Figure). Most prokaryotes contain a single, circular chromosome that is found in an area of the cytoplasm called the nucleoid.
ART CONNECTION



A eukaryote contains a well-defined nucleus, whereas in prokaryotes, the chromosome lies in the cytoplasm in an area called the nucleoid.

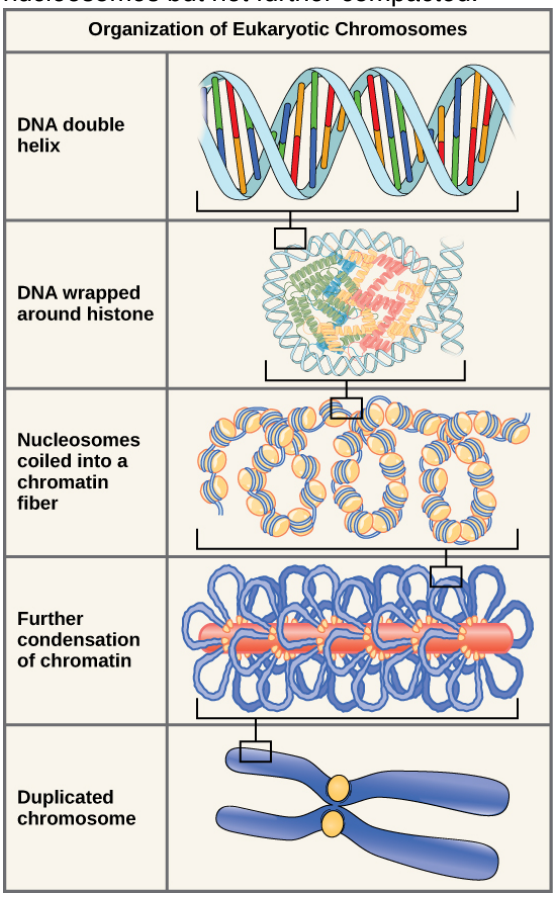
In eukaryotic cells, DNA and RNA synthesis occur in a separate compartment from protein synthesis. In prokaryotic cells, both processes occur together. What advantages might there be to separating the processes? What advantages might there be to having them occur together?

The size of the genome in one of the most well-studied prokaryotes, *E.coli,* is 4.6 million base pairs (approximately 1.1 mm, if cut and stretched out). So how does this fit inside a small bacterial cell? The DNA is

twisted by what is known as supercoiling. Supercoiling means that DNA is either under-wound (less than one turn of the helix per 10 base pairs) or over-wound (more than 1 turn per 10 base pairs) from its normal relaxed state. Some proteins are known to be involved in the supercoiling; other proteins and enzymes such as DNA gyrase help in maintaining the supercoiled structure.

Eukaryotes, whose chromosomes each consist of a linear DNA molecule, employ a different type of packing strategy to fit their DNA inside the nucleus (Figure). At the most basic level, DNA is wrapped around proteins known as histones to form structures called nucleosomes. The histones are evolutionarily conserved proteins that are rich in basic amino acids and form an octamer. The DNA (which is negatively charged because of the phosphate groups) is wrapped tightly around the histone core. This nucleosome is linked to the next one with the help of a linker DNA. This is also known as the “beads on a string” structure. This is further compacted into a 30 nm fiber, which is the diameter of the structure. At the metaphase stage, the chromosomes are at their most compact, are approximately 700 nm in width, and are found in association with scaffold proteins.

In interphase, eukaryotic chromosomes have two distinct regions that can be distinguished by staining. The tightly packaged region is known as heterochromatin, and the less dense region is known as euchromatin. Heterochromatin usually contains genes that are not expressed, and is found in the regions of the centromere and telomeres. The euchromatin usually contains genes that are transcribed, with DNA packaged around nucleosomes but not further compacted.



These figures illustrate the compaction of the eukaryotic chromosome.

Section Summary

The currently accepted model of the double-helix structure of DNA was proposed by Watson and Crick. Some of the salient features are that the two strands that make up the double helix are complementary and anti-parallel in nature. Deoxyribose sugars and phosphates form the backbone of the structure, and the nitrogenous bases are stacked inside. The diameter of the double helix, 2 nm, is uniform throughout. A purine always pairs with a pyrimidine; A pairs with T, and G pairs with C. One turn of the helix has ten base pairs. During cell division, each daughter cell receives a copy of the DNA by a process known as DNA replication. Prokaryotes are much simpler

than eukaryotes in many of their features. Most prokaryotes contain a single, circular chromosome. In general, eukaryotic chromosomes contain a linear DNA molecule packaged into nucleosomes, and have two distinct regions that can be distinguished by staining, reflecting different states of packaging and compaction.

Art Connections

Figure In eukaryotic cells, DNA and RNA synthesis occur in a separate compartment from protein synthesis. In prokaryotic cells, both processes occur together. What advantages might there be to separating the processes? What advantages might there be to having them occur together?

Review Questions

DNA double helix does not have which of the following?

1. antiparallel configuration
2. complementary base pairing
3. major and minor grooves
4. uracil

In eukaryotes, what is the DNA wrapped around?

1. single-stranded binding proteins
2. sliding clamp
3. polymerase
4. histones

Free Response

Provide a brief summary of the Sanger sequencing method.

Describe the structure and complementary base pairing of DNA.

Footnotes

1. 1

   Richard E. Green et al., “A Draft Sequence of the Neandertal Genome,” *Science* 328 (2010): 710-22.

Glossary

electrophoresis technique used to separate DNA fragments according to size

Licensing & Attributions

CC licensed content, Shared previously

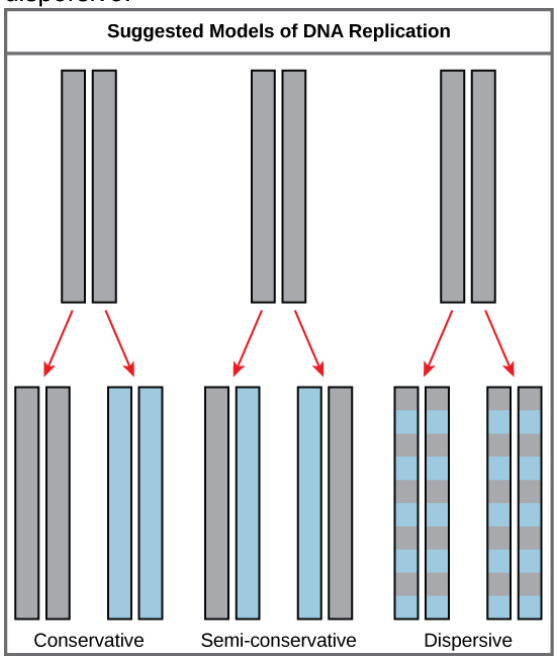
- Biology . **Authored by**: OpenStax College. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License**: *CC BY: Attribution*

BASICS OF DNA REPLICATION

14.3 Basics of DNA Replication

Summary

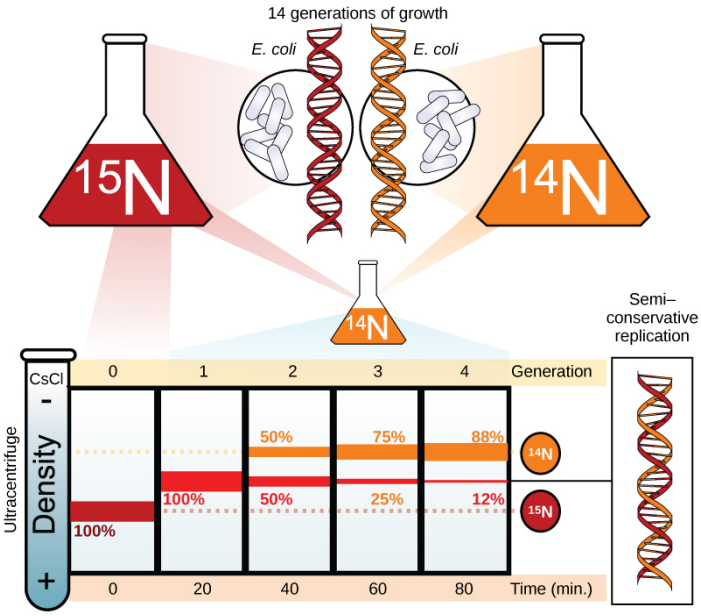
The elucidation of the structure of the double helix provided a hint as to how DNA divides and makes copies of itself. This model suggests that the two strands of the double helix separate during replication, and each strand serves as a template from which the new complementary strand is copied. What was not clear was how the replication took place. There were three models suggested (Figure): conservative, semi-conservative, and dispersive.



The three suggested models of DNA replication. Grey indicates the original DNA strands, and blue indicates newly synthesized DNA.

In conservative replication, the parental DNA remains together, and the newly formed daughter strands are together. The semi-conservative method suggests that each of the two parental DNA strands act as a template for new DNA to be synthesized; after replication, each double-stranded DNA includes one parental or “old” strand and one “new” strand. In the dispersive model, both copies of DNA have double-stranded segments of parental DNA and newly synthesized DNA interspersed.

Meselson and Stahl were interested in understanding how DNA replicates. They grew *E. coli* for several generations in a medium containing a “heavy” isotope of nitrogen (^{15}N) that gets incorporated into nitrogenous bases, and eventually into the DNA (Figure).



Meselson and Stahl experimented with *E. coli* grown first in heavy nitrogen (^{15}N) then in ^{14}N. DNA grown in ^{15}N (red band) is heavier than DNA grown in ^{14}N (orange band), and sediments to a lower level in cesium chloride solution in an ultracentrifuge. When DNA grown in ^{15}N is switched to media containing ^{14}N, after one round of cell division the DNA sediments halfway between the ^{15}N and ^{14}N levels, indicating that it now contains fifty percent ^{14}N. In subsequent cell divisions, an increasing amount of DNA contains ^{14}N only. This data supports the semi-conservative replication model. (credit: modification of work by Mariana Ruiz Villareal)

The *E. coli* culture was then shifted into medium containing ^{14}N and allowed to grow for one generation. The cells were harvested and the DNA was isolated. The DNA was centrifuged at high speeds in an ultracentrifuge. Some cells were allowed to grow for one more life cycle in ^{14}N and spun again. During the density gradient centrifugation, the DNA is loaded into a gradient (typically a salt such as cesium chloride or sucrose) and spun at high speeds of 50,000 to 60,000 rpm. Under these circumstances, the DNA will form a band according to its density in the gradient. DNA grown in ^{15}N will band at a higher density position than that grown in ^{14}N. Meselson and Stahl noted that after one generation of growth in ^{14}N after they had been shifted from ^{15}N, the single band observed was intermediate in position in between DNA of cells grown exclusively in ^{15}N and ^{14}N. This suggested either a semi-conservative or dispersive mode of replication. The DNA harvested from cells grown for two generations in ^{14}N formed two bands: one DNA band was at the intermediate position between ^{15}N and ^{14}N, and the other corresponded to the band of ^{14}N DNA. These results could only be explained if DNA replicates in a semi-conservative manner. Therefore, the other two modes were ruled out.

During DNA replication, each of the two strands that make up the double helix serves as a template from which new strands are copied. The new strand will be complementary to the parental or “old” strand. When two daughter DNA copies are formed, they have the same sequence and are divided equally into the two daughter cells.

Link to Learning

Click through this tutorial on DNA replication.

Section Summary

The model for DNA replication suggests that the two strands of the double helix separate during replication, and each strand serves as a template from which the new complementary strand is copied. In conservative replication, the parental DNA is conserved, and the daughter DNA is newly synthesized. The semi-conservative method suggests that each of the two parental DNA strands acts as template for new DNA to be synthesized;

after replication, each double-stranded DNA includes one parental or “old” strand and one “new” strand. The dispersive mode suggested that the two copies of the DNA would have segments of parental DNA and newly synthesized DNA.

Review Questions

Meselson and Stahl’s experiments proved that DNA replicates by which mode?

1. conservative
2. semi-conservative
3. dispersive
4. none of the above

If the sequence of the 5′-3′ strand is AATGCTAC, then the complementary sequence has the following sequence:

1. 3′-AATGCTAC-5′
2. 3′-CATCGTAA-5′
3. 3′-TTACGATG-5′
4. 3′-GTAGCATT-5′

Free Response

How did the scientific community learn that DNA replication takes place in a semi-conservative fashion?

- Downloads
- History
- Attribution
- More Information

This work is licensed by Rice University under a Creative Commons Attribution License (by 4.0).

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: OpenStax College. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License**: *CC BY: Attribution*

DNA REPLICATION (PROKARYOTES)

14.4 DNA Replication in Prokaryotes

Summary

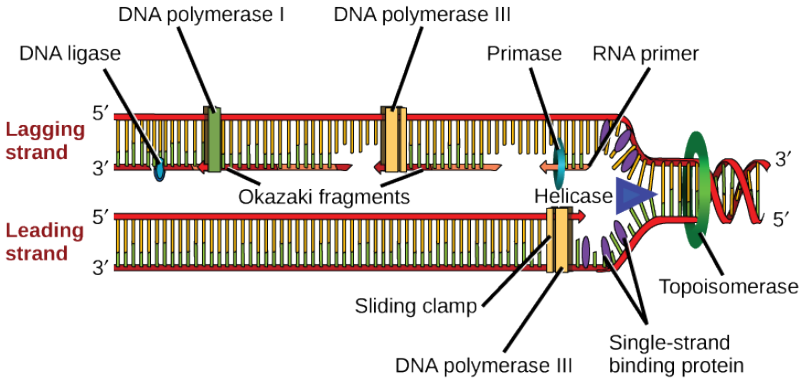
DNA replication has been extremely well studied in prokaryotes primarily because of the small size of the genome and the mutants that are available. *E. coli* has 4.6 million base pairs in a single circular chromosome and all of it gets replicated in approximately 42 minutes, starting from a single origin of replication and proceeding around the circle in both directions. This means that approximately 1000 nucleotides are added per second. The process is quite rapid and occurs without many mistakes.

DNA replication employs a large number of proteins and enzymes, each of which plays a critical role during the process. One of the key players is the enzyme DNA polymerase, also known as DNA pol, which adds nucleotides one by one to the growing DNA chain that are complementary to the template strand. The addition of nucleotides requires energy; this energy is obtained from the nucleotides that have three phosphates attached to them, similar to ATP which has three phosphate groups attached. When the bond between the phosphates is broken, the

energy released is used to form the phosphodiester bond between the incoming nucleotide and the growing chain. In prokaryotes, three main types of polymerases are known: DNA pol I, DNA pol II, and DNA pol III. It is now known that DNA pol III is the enzyme required for DNA synthesis; DNA pol I and DNA pol II are primarily required for repair.

How does the replication machinery know where to begin? It turns out that there are specific nucleotide sequences called origins of replication where replication begins. In *E. coli,* which has a single origin of replication on its one chromosome (as do most prokaryotes), it is approximately 245 base pairs long and is rich in AT sequences. The origin of replication is recognized by certain proteins that bind to this site. An enzyme called helicase unwinds the DNA by breaking the hydrogen bonds between the nitrogenous base pairs. ATP hydrolysis is required for this process. As the DNA opens up, Y-shaped structures called replication forks are formed. Two replication forks are formed at the origin of replication and these get extended bi- directionally as replication proceeds. Single-strand binding proteins coat the single strands of DNA near the replication fork to prevent the single-stranded DNA from winding back into a double helix. DNA polymerase is able to add nucleotides only in the 5′ to 3′ direction (a new DNA strand can be only extended in this direction). It also requires a free 3′-OH group to which it can add nucleotides by forming a phosphodiester bond between the 3′-OH end and the 5′ phosphate of the next nucleotide. This essentially means that it cannot add nucleotides if a free 3′-OH group is not available. Then how does it add the first nucleotide? The problem is solved with the help of a primer that provides the free 3′-OH end. Another enzyme, RNA primase, synthesizes an RNA primer that is about five to ten nucleotides long and complementary to the DNA. Because this sequence primes the DNA synthesis, it is appropriately called the primer. DNA polymerase can now extend this RNA primer, adding nucleotides one by one that are complementary to the template strand (Figure).

ART CONNECTION



A replication fork is formed when helicase separates the DNA strands at the origin of replication. The DNA tends to become more highly coiled ahead of the replication fork. Topoisomerase breaks and reforms DNA's phosphate backbone ahead of the replication fork, thereby relieving the pressure that results from this supercoiling. Single-strand binding proteins bind to the single-stranded DNA to prevent the helix from re-forming. Primase synthesizes an RNA primer. DNA polymerase III uses this primer to synthesize the daughter DNA strand. On the leading strand, DNA is synthesized continuously, whereas on the lagging strand, DNA is synthesized in short stretches called Okazaki fragments. DNA polymerase I replaces the RNA primer with DNA. DNA ligase seals the gaps between the Okazaki fragments, joining the fragments into a single DNA molecule. (credit: modification of work by Mariana Ruiz Villareal)

You isolate a cell strain in which the joining together of Okazaki fragments is impaired and suspect that a mutation has occurred in an enzyme found at the replication fork. Which enzyme is most likely to be mutated?

The replication fork moves at the rate of 1000 nucleotides per second. DNA polymerase can only extend in the 5′ to 3′ direction, which poses a slight problem at the replication fork. As we know, the DNA double helix is anti-parallel; that is, one strand is in the 5′ to 3′ direction and the other is oriented in the 3′ to 5′ direction. One strand, which is complementary to the 3′ to 5′ parental DNA strand, is synthesized continuously towards the replication fork because the polymerase can add nucleotides in this direction. This continuously synthesized strand is known as the leading strand. The other strand, complementary to the 5′ to 3′ parental DNA, is extended away from the replication fork, in small fragments known as Okazaki fragments, each requiring a primer to start the synthesis. Okazaki fragments are named after the Japanese scientist who first discovered them. The strand with the Okazaki fragments is known as the lagging strand.

The leading strand can be extended by one primer alone, whereas the lagging strand needs a new primer for each of the short Okazaki fragments. The overall direction of the lagging strand will be 3′ to 5′, and that of the

leading strand 5′ to 3′. A protein called the sliding clamp holds the DNA polymerase in place as it continues to add nucleotides. The sliding clamp is a ring-shaped protein that binds to the DNA and holds the polymerase in place. Topoisomerase prevents the over-winding of the DNA double helix ahead of the replication fork as the DNA is opening up; it does so by causing temporary nicks in the DNA helix and then resealing it. As synthesis proceeds, the RNA primers are replaced by DNA. The primers are removed by the exonuclease activity of DNA pol I, and the gaps are filled in by deoxyribonucleotides. The nicks that remain between the newly synthesized DNA (that replaced the RNA primer) and the previously synthesized DNA are sealed by the enzyme DNA ligase that catalyzes the formation of phosphodiester linkage between the 3′-OH end of one nucleotide and the 5′ phosphate end of the other fragment.

Once the chromosome has been completely replicated, the two DNA copies move into two different cells during cell division. The process of DNA replication can be summarized as follows:

1. DNA unwinds at the origin of replication.
2. Helicase opens up the DNA-forming replication forks; these are extended bidirectionally.
3. Single-strand binding proteins coat the DNA around the replication fork to prevent rewinding of the DNA.
4. Topoisomerase binds at the region ahead of the replication fork to prevent supercoiling.
5. Primase synthesizes RNA primers complementary to the DNA strand.
6. DNA polymerase starts adding nucleotides to the 3′-OH end of the primer.
7. Elongation of both the lagging and the leading strand continues.
8. RNA primers are removed by exonuclease activity.
9. Gaps are filled by DNA pol by adding dNTPs.
10. The gap between the two DNA fragments is sealed by DNA ligase, which helps in the formation of phosphodiester bonds.

Table summarizes the enzymes involved in prokaryotic DNA replication and the functions of each.

Prokaryotic DNA Replication: Enzymes and Their Function	
Enzyme/protein	Specific Function
DNA pol I	Exonuclease activity removes RNA primer and replaces with newly synthesized DNA
DNA pol II	Repair function
DNA pol III	Main enzyme that adds nucleotides in the 5′-3′ direction
Helicase	Opens the DNA helix by breaking hydrogen bonds between the nitrogenous bases
Ligase	Seals the gaps between the Okazaki fragments to create one continuous DNA strand
Primase	Synthesizes RNA primers needed to start replication
Sliding Clamp	Helps to hold the DNA polymerase in place when nucleotides are being added
Topoisomerase	Helps relieve the stress on DNA when unwinding by causing breaks and then resealing the DNA
Single-strand binding proteins (SSB)	Binds to single-stranded DNA to avoid DNA rewinding back.

Link to Learning

Review the full process of DNA replication here.

Section Summary

Replication in prokaryotes starts from a sequence found on the chromosome called the origin of replication—the point at which the DNA opens up. Helicase opens up the DNA double helix, resulting in the formation of the replication fork. Single-strand binding proteins bind to the single-stranded DNA near the replication fork to keep the fork open. Primase synthesizes an RNA primer to initiate synthesis by DNA polymerase, which can add nucleotides only in the 5′ to 3′ direction. One strand is synthesized continuously in the direction of the replication fork; this is called the leading strand. The other strand is synthesized in a direction away from the replication fork, in short stretches of DNA known as Okazaki fragments. This strand is known as the lagging strand. Once replication is completed, the RNA primers are replaced by DNA nucleotides and the DNA is sealed with DNA ligase, which creates phosphodiester bonds between the 3′-OH of one end and the 5′ phosphate of the other strand.

Art Connections

Figure You isolate a cell strain in which the joining together of Okazaki fragments is impaired and suspect that a mutation has occurred in an enzyme found at the replication fork. Which enzyme is most likely to be mutated?

Review Questions

Which of the following components is not involved during the formation of the replication fork?

1. single-strand binding proteins
2. helicase
3. origin of replication
4. ligase

Which of the following does the enzyme primase synthesize?

1. DNA primer
2. RNA primer
3. Okazaki fragments
4. phosphodiester linkage

In which direction does DNA replication take place?

1. 5′-3′
2. 3′-5′
3. 5′
4. 3′

Free Response

DNA replication is bidirectional and discontinuous; explain your understanding of those concepts.

What are Okazaki fragments and how they are formed?

If the rate of replication in a particular prokaryote is 900 nucleotides per second, how long would it take 1.2 million base pair genomes to make two copies?

Explain the events taking place at the replication fork. If the gene for helicase is mutated, what part of replication will be affected?

What is the role of a primer in DNA replication? What would happen if you forgot to add a primer in a tube containing the reaction mix for a DNA sequencing reaction?

Glossary

helicase during replication, this enzyme helps to open up the DNA helix by breaking the hydrogen bonds

lagging strand during replication, the strand that is replicated in short fragments and away from the replication fork

leading strand strand that is synthesized continuously in the 5′-3′ direction which is synthesized in the direction of the replication fork

ligase enzyme that catalyzes the formation of a phosphodiester linkage between the 3′ OH and 5′ phosphate ends of the DNA

Okazaki fragment DNA fragment that is synthesized in short stretches on the lagging strand

primase enzyme that synthesizes the RNA primer; the primer is needed for DNA pol to start synthesis of a new DNA strand

primer short stretch of nucleotides that is required to initiate replication; in the case of replication, the primer has RNA nucleotides

replication fork Y-shaped structure formed during initiation of replication

single-strand binding protein during replication, protein that binds to the single-stranded DNA; this helps in keeping the two strands of DNA apart so that they may serve as templates

sliding clamp ring-shaped protein that holds the DNA pol on the DNA strand

topoisomerase enzyme that causes underwinding or overwinding of DNA when DNA replication is taking place

- Downloads
- History
- Attribution
- More Information

This work is licensed by Rice University under a Creative Commons Attribution License (by 4.0).

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: OpenStax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License**: *CC BY: Attribution*

DNA REPLICATION (EUKARYOTES)

14.5 DNA Replication in Eukaryotes

Summary

Eukaryotic genomes are much more complex and larger in size than prokaryotic genomes. The human genome has three billion base pairs per haploid set of chromosomes, and 6 billion base pairs are replicated during the S phase of the cell cycle. There are multiple origins of replication on the eukaryotic chromosome; humans can have up to 100,000 origins of replication. The rate of replication is approximately 100 nucleotides per second, much slower than prokaryotic replication. In yeast, which is a eukaryote, special sequences known as Autonomously Replicating Sequences (ARS) are found on the chromosomes. These are equivalent to the origin of replication in *E. coli*.

The number of DNA polymerases in eukaryotes is much more than prokaryotes: 14 are known, of which five are known to have major roles during replication and have been well studied. They are known as pol *α*, pol *β*, pol *γ*, pol *δ*, and pol *ε*.

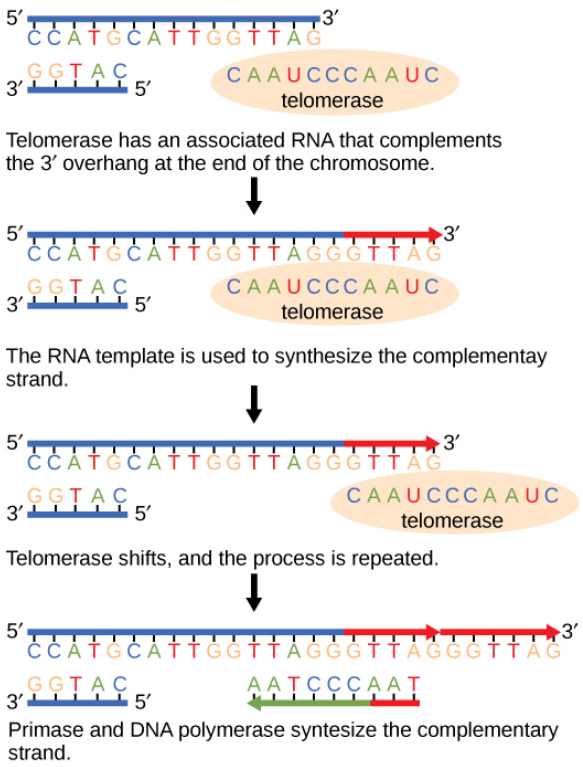
The essential steps of replication are the same as in prokaryotes. Before replication can start, the DNA has to be made available as template. Eukaryotic DNA is bound to basic proteins known as histones to form structures called nucleosomes. The chromatin (the complex between DNA and proteins) may undergo some chemical modifications, so that the DNA may be able to slide off the proteins or be accessible to the enzymes of the DNA replication machinery. At the origin of replication, a pre-replication complex is made with other initiator proteins. Other proteins are then recruited to start the replication process (Table).

A helicase using the energy from ATP hydrolysis opens up the DNA helix. Replication forks are formed at each replication origin as the DNA unwinds. The opening of the double helix causes over-winding, or supercoiling, in the DNA ahead of the replication fork. These are resolved with the action of topoisomerases. Primers are formed by the enzyme primase, and using the primer, DNA pol can start synthesis. While the leading strand is continuously synthesized by the enzyme pol *δ*, the lagging strand is synthesized by pol *ε*. A sliding clamp protein known as PCNA (Proliferating Cell Nuclear Antigen) holds the DNA pol in place so that it does not slide off the DNA. RNase H removes the RNA primer, which is then replaced with DNA nucleotides. The Okazaki fragments in the lagging strand are joined together after the replacement of the RNA primers with DNA. The gaps that remain are sealed by DNA ligase, which forms the phosphodiester bond.

Telomere replication

Unlike prokaryotic chromosomes, eukaryotic chromosomes are linear. As you've learned, the enzyme DNA pol can add nucleotides only in the 5′ to 3′ direction. In the leading strand, synthesis continues until the end of the chromosome is reached. On the lagging strand, DNA is synthesized in short stretches, each of which is initiated by a separate primer. When the replication fork reaches the end of the linear chromosome, there is no place for a primer to be made for the DNA fragment to be copied at the end of the chromosome. These ends thus remain unpaired, and over time these ends may get progressively shorter as cells continue to divide.

The ends of the linear chromosomes are known as telomeres, which have repetitive sequences that code for no particular gene. In a way, these telomeres protect the genes from getting deleted as cells continue to divide. In humans, a six base pair sequence, TTAGGG, is repeated 100 to 1000 times. The discovery of the enzyme telomerase (Figure) helped in the understanding of how chromosome ends are maintained. The telomerase enzyme contains a catalytic part and a built-in RNA template. It attaches to the end of the chromosome, and complementary bases to the RNA template are added on the 3′ end of the DNA strand. Once the 3′ end of the lagging strand template is sufficiently elongated, DNA polymerase can add the nucleotides complementary to the ends of the chromosomes. Thus, the ends of the chromosomes are replicated.



The ends of linear chromosomes are maintained by the action of the telomerase enzyme.

Telomerase is typically active in germ cells and adult stem cells. It is not active in adult somatic cells. For her discovery of telomerase and its action, Elizabeth Blackburn (Figure) received the Nobel Prize for Medicine and Physiology in 2009.

Elizabeth Blackburn, 2009 Nobel Laureate, is the scientist who discovered how telomerase works. (credit: US Embassy Sweden)

Telomerase and Aging

Cells that undergo cell division continue to have their telomeres shortened because most somatic cells do not make telomerase. This essentially means that telomere shortening is associated with aging. With the advent of modern medicine, preventative health care, and healthier lifestyles, the human life span has increased, and there is an increasing demand for people to look younger and have a better quality of life as they grow older.

In 2010, scientists found that telomerase can reverse some age-related conditions in mice. This may have potential in regenerative medicine.[1] Telomerase-deficient mice were used in these studies; these mice have

tissue atrophy, stem cell depletion, organ system failure, and impaired tissue injury responses. Telomerase reactivation in these mice caused extension of telomeres, reduced DNA damage, reversed neurodegeneration, and improved the function of the testes, spleen, and intestines. Thus, telomere reactivation may have potential for treating age-related diseases in humans.

Cancer is characterized by uncontrolled cell division of abnormal cells. The cells accumulate mutations, proliferate uncontrollably, and can migrate to different parts of the body through a process called metastasis. Scientists have observed that cancerous cells have considerably shortened telomeres and that telomerase is active in these cells. Interestingly, only after the telomeres were shortened in the cancer cells did the telomerase become active. If the action of telomerase in these cells can be inhibited by drugs during cancer therapy, then the cancerous cells could potentially be stopped from further division.

Difference between Prokaryotic and Eukaryotic Replication		
Property	Prokaryotes	Eukaryotes
Origin of replication	Single	Multiple
Rate of replication	1000 nucleotides/s	50 to 100 nucleotides/s
DNA polymerase types	5	14
Telomerase	Not present	Present
RNA primer removal	DNA pol I	RNase H
Strand elongation	DNA pol III	Pol δ, pol ε
Sliding clamp	Sliding clamp	PCNA

Section Summary

Replication in eukaryotes starts at multiple origins of replication. The mechanism is quite similar to prokaryotes. A primer is required to initiate synthesis, which is then extended by DNA polymerase as it adds nucleotides one by one to the growing chain. The leading strand is synthesized continuously, whereas the lagging strand is synthesized in short stretches called Okazaki fragments. The RNA primers are replaced with DNA nucleotides; the DNA remains one continuous strand by linking the DNA fragments with DNA ligase. The ends of the chromosomes pose a problem as polymerase is unable to extend them without a primer. Telomerase, an enzyme with an inbuilt RNA template, extends the ends by copying the RNA template and extending one end of the chromosome. DNA polymerase can then extend the DNA using the primer. In this way, the ends of the chromosomes are protected.

Review Questions

The ends of the linear chromosomes are maintained by

1. helicase
2. primase
3. DNA pol
4. telomerase

Free Response

How do the linear chromosomes in eukaryotes ensure that its ends are replicated completely?

Footnotes

1. 1 Jaskelioff et al., “Telomerase reactivation reverses tissue degeneration in aged telomerase-deficient mice,” *Nature* 469 (2011): 102-7.

Glossary

telomerase enzyme that contains a catalytic part and an inbuilt RNA template; it functions to maintain telomeres at chromosome ends

telomere DNA at the end of linear chromosomes

- Downloads
- History
- Attribution
- More Information

This work is licensed by Rice University under a Creative Commons Attribution License (by 4.0).

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: OpenStax College. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License**: *CC BY: Attribution*

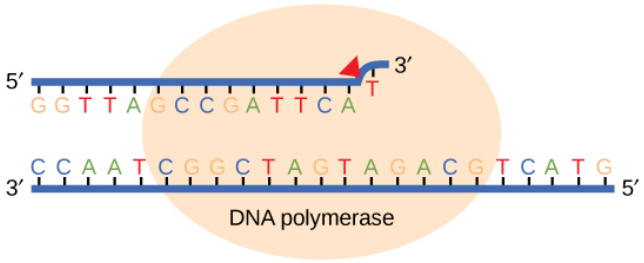
DNA REPAIR

14.6 DNA Repair

Summary

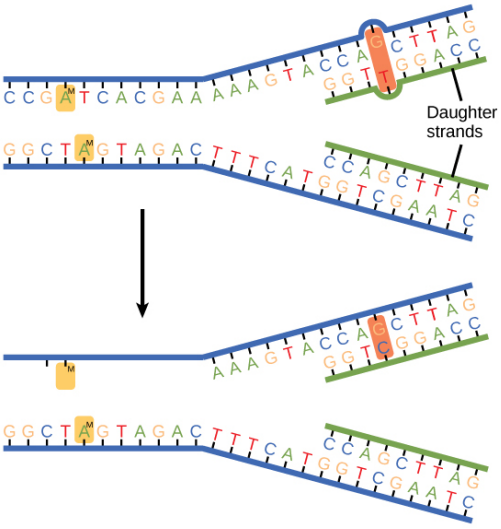
DNA replication is a highly accurate process, but mistakes can occasionally occur, such as a DNA polymerase inserting a wrong base. Uncorrected mistakes may sometimes lead to serious consequences, such as cancer. Repair mechanisms correct the mistakes. In rare cases, mistakes are not corrected, leading to mutations; in other cases, repair enzymes are themselves mutated or defective.

Most of the mistakes during DNA replication are promptly corrected by DNA polymerase by proofreading the base that has been just added (Figure). In proofreading, the DNA pol reads the newly added base before adding the next one, so a correction can be made. The polymerase checks whether the newly added base has paired correctly with the base in the template strand. If it is the right base, the next nucleotide is added. If an incorrect base has been added, the enzyme makes a cut at the phosphodiester bond and releases the wrong nucleotide. This is performed by the exonuclease action of DNA pol III. Once the incorrect nucleotide has been removed, a new one will be added again.



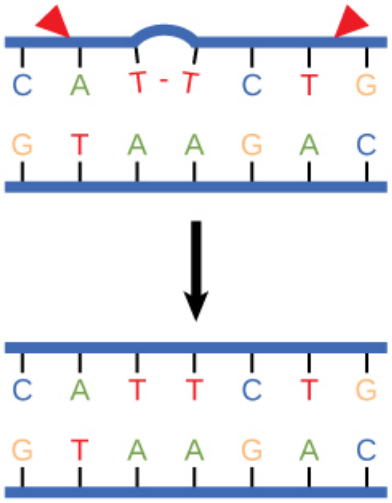
Proofreading by DNA polymerase corrects errors during replication.

Some errors are not corrected during replication, but are instead corrected after replication is completed; this type of repair is known as mismatch repair (Figure). The enzymes recognize the incorrectly added nucleotide and excise it; this is then replaced by the correct base. If this remains uncorrected, it may lead to more permanent damage. How do mismatch repair enzymes recognize which of the two bases is the incorrect one? In *E. coli*, after replication, the nitrogenous base adenine acquires a methyl group; the parental DNA strand will have methyl groups, whereas the newly synthesized strand lacks them. Thus, DNA polymerase is able to remove the wrongly incorporated bases from the newly synthesized, non-methylated strand. In eukaryotes, the mechanism is not very well understood, but it is believed to involve recognition of unsealed nicks in the new strand, as well as a short-term continuing association of some of the replication proteins with the new daughter strand after replication has completed.



In mismatch repair, the incorrectly added base is detected after replication. The mismatch repair proteins detect this base and remove it from the newly synthesized strand by nuclease action. The gap is now filled with the correctly paired base.

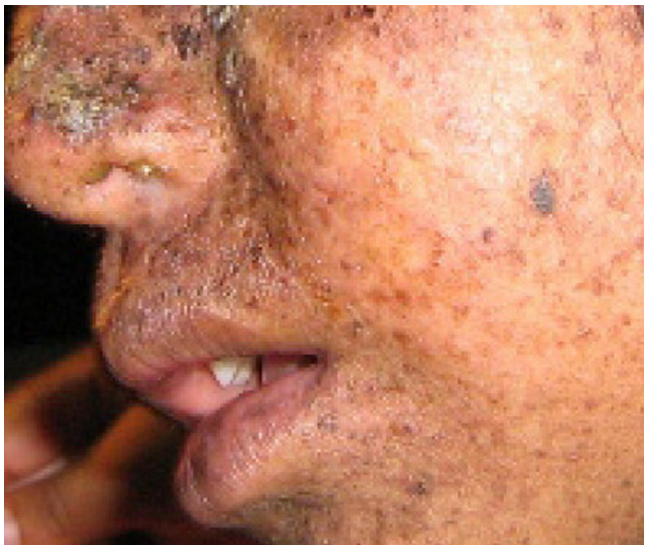
In another type of repair mechanism, nucleotide excision repair, enzymes replace incorrect bases by making a cut on both the 3′ and 5′ ends of the incorrect base (Figure). The segment of DNA is removed and replaced with the correctly paired nucleotides by the action of DNA pol. Once the bases are filled in, the remaining gap is sealed with a phosphodiester linkage catalyzed by DNA ligase. This repair mechanism is often employed when UV exposure causes the formation of pyrimidine dimers.



Nucleotide excision repairs thymine dimers. When exposed to UV, thymines lying adjacent to each other can form thymine dimers. In normal cells, they are excised and replaced.

A well-studied example of mistakes not being corrected is seen in people suffering from xeroderma pigmentosa (Figure). Affected individuals have skin that is highly sensitive to UV rays from the sun. When individuals are exposed to UV, pyrimidine dimers, especially those of thymine, are formed; people with xeroderma pigmentosa

are not able to repair the damage. These are not repaired because of a defect in the nucleotide excision repair enzymes, whereas in normal individuals, the thymine dimers are excised and the defect is corrected. The thymine dimers distort the structure of the DNA double helix, and this may cause problems during DNA replication. People with xeroderma pigmentosa may have a higher risk of contracting skin cancer than those who dont have the condition.

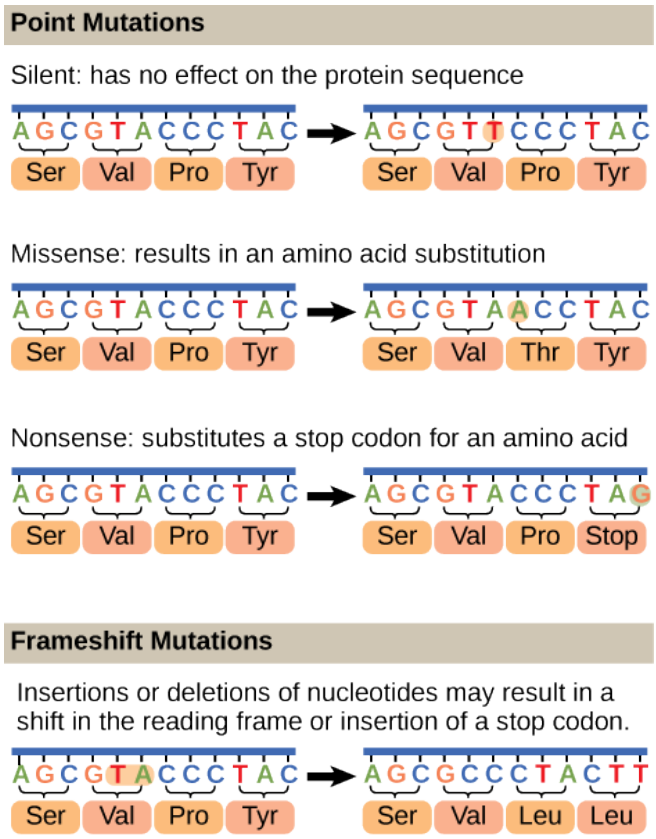


Xeroderma pigmentosa is a condition in which thymine dimerization from exposure to UV is not repaired. Exposure to sunlight results in skin lesions. (credit: James Halpern et al.)

Errors during DNA replication are not the only reason why mutations arise in DNA. Mutations, variations in the nucleotide sequence of a genome, can also occur because of damage to DNA. Such mutations may be of two types: induced or spontaneous. Induced mutations are those that result from an exposure to chemicals, UV rays, x-rays, or some other environmental agent. Spontaneous mutations occur without any exposure to any environmental agent; they are a result of natural reactions taking place within the body.

Mutations may have a wide range of effects. Some mutations are not expressed; these are known as silent mutations. Point mutations are those mutations that affect a single base pair. The most common nucleotide mutations are substitutions, in which one base is replaced by another. These can be of two types, either transitions or transversions. Transition substitution refers to a purine or pyrimidine being replaced by a base of the same kind; for example, a purine such as adenine may be replaced by the purine guanine. Transversion substitution refers to a purine being replaced by a pyrimidine, or vice versa; for example, cytosine, a pyrimidine, is replaced by adenine, a purine. Mutations can also be the result of the addition of a base, known as an insertion, or the removal of a base, also known as deletion. Sometimes a piece of DNA from one chromosome may get translocated to another chromosome or to another region of the same chromosome; this is also known as translocation. These mutation types are shown in Figure.

ART CONNECTION



Mutations can lead to changes in the protein sequence encoded by the DNA.

A frameshift mutation that results in the insertion of three nucleotides is often less deleterious than a mutation that results in the insertion of one nucleotide. Why?

Mutations in repair genes have been known to cause cancer. Many mutated repair genes have been implicated in certain forms of pancreatic cancer, colon cancer, and colorectal cancer. Mutations can affect either somatic cells or germ cells. If many mutations accumulate in a somatic cell, they may lead to problems such as the uncontrolled cell division observed in cancer. If a mutation takes place in germ cells, the mutation will be passed on to the next generation, as in the case of hemophilia and xeroderma pigmentosa.

Section Summary

DNA polymerase can make mistakes while adding nucleotides. It edits the DNA by proofreading every newly added base. Incorrect bases are removed and replaced by the correct base, and then a new base is added. Most mistakes are corrected during replication, although when this does not happen, the mismatch repair mechanism is employed. Mismatch repair enzymes recognize the wrongly incorporated base and excise it from the DNA, replacing it with the correct base. In yet another type of repair, nucleotide excision repair, the incorrect base is removed along with a few bases on the 5′ and 3′ end, and these are replaced by copying the template with the help of DNA polymerase. The ends of the newly synthesized fragment are attached to the rest of the DNA using DNA ligase, which creates a phosphodiester bond.

Most mistakes are corrected, and if they are not, they may result in a mutation defined as a permanent change in the DNA sequence. Mutations can be of many types, such as substitution, deletion, insertion, and translocation. Mutations in repair genes may lead to serious consequences such as cancer. Mutations can be induced or may occur spontaneously.

Art Connections

Figure A frameshift mutation that results in the insertion of three nucleotides is often less deleterious than a mutation that results in the insertion of one nucleotide. Why?

Review Questions

During proofreading, which of the following enzymes reads the DNA?

1. primase
2. topoisomerase
3. DNA pol
4. helicase

The initial mechanism for repairing nucleotide errors in DNA is _____.

1. mismatch repair
2. DNA polymerase proofreading
3. nucleotide excision repair
4. thymine dimers

Free Response

What is the consequence of mutation of a mismatch repair enzyme? How will this affect the function of a gene?

Glossary

induced mutation mutation that results from exposure to chemicals or environmental agents

mutation variation in the nucleotide sequence of a genome

mismatch repair type of repair mechanism in which mismatched bases are removed after replication

nucleotide excision repair type of DNA repair mechanism in which the wrong base, along with a few nucleotides upstream or downstream, are removed

proofreading function of DNA pol in which it reads the newly added base before adding the next one

point mutation mutation that affects a single base

silent mutation mutation that is not expressed

spontaneous mutation mutation that takes place in the cells as a result of chemical reactions taking place naturally without exposure to any external agent

transition substitution when a purine is replaced with a purine or a pyrimidine is replaced with another pyrimidine

transversion substitution when a purine is replaced by a pyrimidine or a pyrimidine is replaced by a purine

- Downloads
- History
- Attribution
- More Information

This work is licensed by Rice University under a Creative Commons Attribution License (by 4.0).

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: OpenStax College. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License**: *CC BY: Attribution*

STUDY GUIDE: DNA STRUCTURE

Study Questions

Objective: Describe the experiments, data, and conclusions that were instrumental in the discovery of the structure of DNA.

Use this page to check your understanding of the content.

Vocabulary

1. Nucleic acid
2. Nucleotide
3. Nitrogen base
4. Purine
5. Pyrimidine
6. Pentose sugar
7. Phosphate group

Study Guide Questions

1. Clearly describe the general structure of nucleic acids.
2. Write one sentence that clearly illustrates the relationship between “nucleotide” and “nucleic acid”.
3. Compare and contrast DNA and RNA.
4. Distinguish between the 3’ and 5’ ends of a nucleic acid.
5. Answer questions like this: If 22% of an organism’s DNA contains ADENINE nucleotides, how many THYMINE nucleotides will the DNA contain? Guanine? Cytosine?
6. Given a handful of nucleotides, be able to build a correct model of DNA.

Licensing & Attributions

CC licensed content, Shared previously

- Study Guide: DNA Structure. **Authored by**: Wendy Riggs. **Provided by**: College of the Redwoods. **Project**: Kaleidoscope. **License**: *CC BY: Attribution*

STUDY GUIDE: DNA REPLICATION

Study Questions

Objective: Relate the structure of DNA to the process of DNA replication.

Use this page to check your understanding of the content.

Study Guide Questions

1. Understand the role of the following enzymes in DNA replication: Helicase, primase, DNA polymerase, ligase
2. Given a strand of DNA and the DIRECTION replication occurs, determine which strand is the leading strand and which is the lagging strand.
3. DNA replication is _____? (Use one word to answer…or one hyphenated word?)

4. What are Okazaki fragments?
5. Clearly explain why Okazaki fragments exist.

Licensing & Attributions

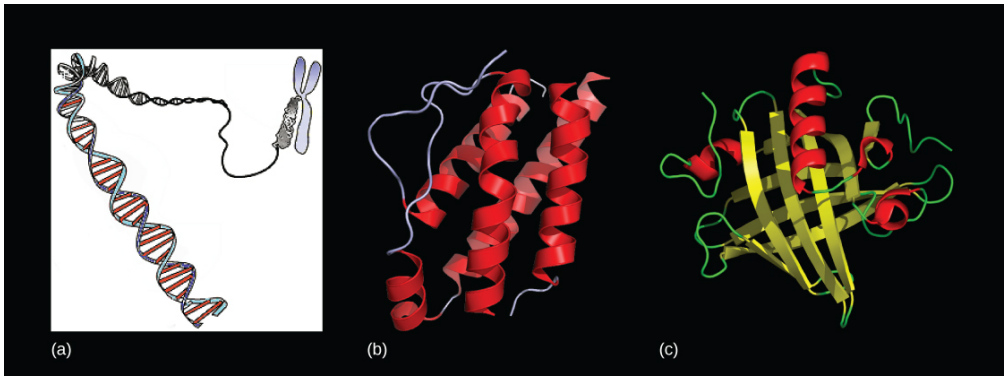
CC licensed content, Shared previously

- Study Guide: DNA Replication. **Authored by:** Wendy Riggs. **Provided by:** College of the Redwoods. **Project:** Kaleidoscope. **License:** *CC BY: Attribution*

GENES AND PROTEINS

INTRODUCTION: GENES AND PROTEINS

Chapter 15. Introduction



Genes, which are carried on (a) chromosomes, are linearly organized instructions for making the RNA and protein molecules that are necessary for all of processes of life. The (b) interleukin-2 protein and (c) alpha-2u-globulin protein are just two examples of the array of different molecular structures that are encoded by genes. (credit "chromosome: National Human Genome Research Institute; credit "interleukin-2": Ramin Herati/Created from PDB 1M47 and rendered with Pymol; credit "alpha-2u-globulin": Darren Logan/rendered with AISMIG)

Since the rediscovery of Mendel's work in 1900, the definition of the gene has progressed from an abstract unit of heredity to a tangible molecular entity capable of replication, expression, and mutation (Figure). Genes are composed of DNA and are linearly arranged on chromosomes. Genes specify the sequences of amino acids, which are the building blocks of proteins. In turn, proteins are responsible for orchestrating nearly every function of the cell. Both genes and the proteins they encode are absolutely essential to life as we know it.

- Downloads
- History
- Attribution
- More Information

This work is licensed by Rice University under a Creative Commons Attribution License (by 4.0).

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: OpenStax College. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License**: *CC BY: Attribution*

THE GENETIC CODE

Learning Objectives

By the end of this section, you will be able to:

- Explain the “central dogma” of protein synthesis
- Describe the genetic code and how the nucleotide sequence prescribes the amino acid and the protein sequence

The cellular process of transcription generates messenger RNA (mRNA), a mobile molecular copy of one or more genes with an alphabet of A, C, G, and uracil (U). Translation of the mRNA template converts nucleotide-based genetic information into a protein product. Protein sequences consist of 20 commonly occurring amino acids; therefore, it can be said that the protein alphabet consists of 20 letters (Figure 1). Each amino acid is defined by a three-nucleotide sequence called the triplet codon. Different amino acids have different chemistries (such as acidic versus basic, or polar and nonpolar) and different structural constraints. Variation in amino acid sequence gives rise to enormous variation in protein structure and function.

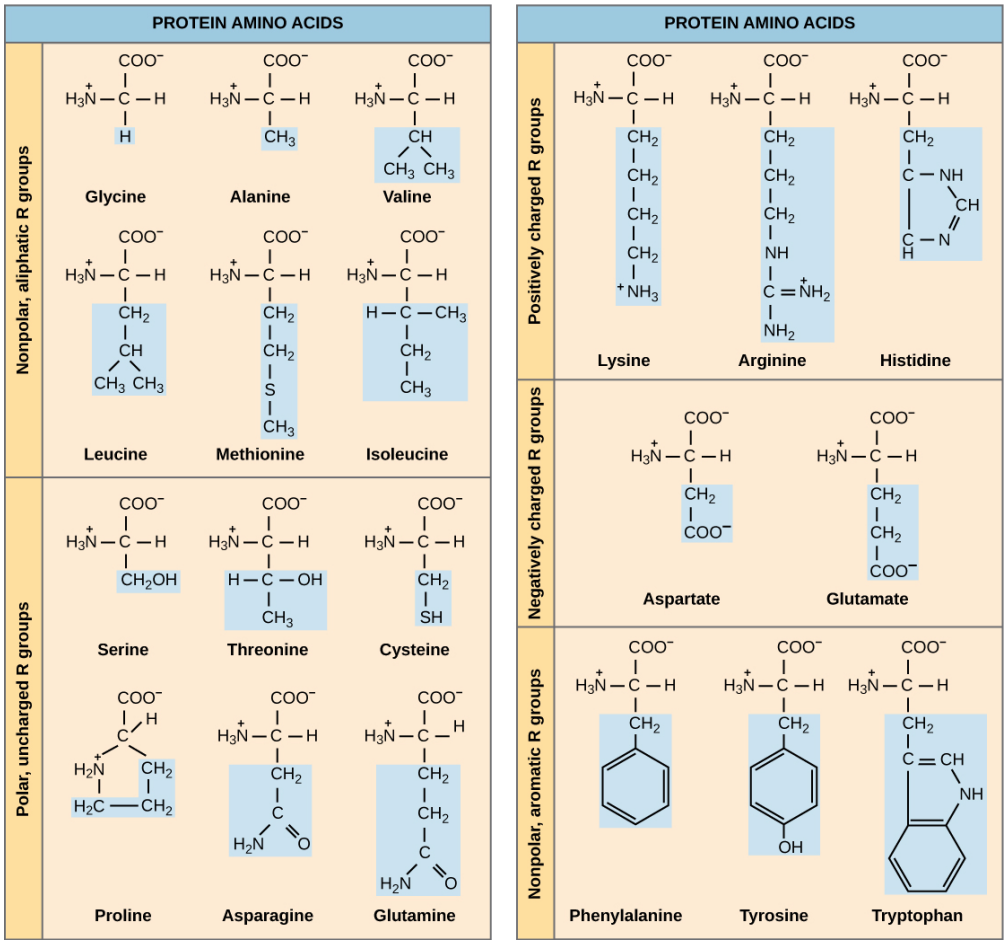


Figure 1. Structures of the 20 amino acids found in proteins are shown. Each amino acid is composed of an amino group (NH+3), a carboxyl group (COO-), and a side chain (blue). The side chain may be nonpolar, polar, or charged, as well as large or small. It is the variety of amino acid side chains that gives rise to the incredible variation of protein structure and function.

The Central Dogma: DNA Encodes RNA; RNA Encodes Protein

The flow of genetic information in cells from DNA to mRNA to protein is described by the Central Dogma (Figure 2), which states that genes specify the sequence of mRNAs, which in turn specify the sequence of proteins. The decoding of one molecule to another is performed by specific proteins and RNAs. Because the information stored in DNA is so central to cellular function, it makes intuitive sense that the cell would make mRNA copies of this information for protein synthesis, while keeping the DNA itself intact and protected. The copying of DNA to RNA is relatively straightforward, with one nucleotide being added to the mRNA strand for every nucleotide read in the DNA strand. The translation to protein is a bit more complex because three mRNA nucleotides correspond to one amino acid in the polypeptide sequence. However, the translation to protein is still systematic and colinear, such that nucleotides 1 to 3 correspond to amino acid 1, nucleotides 4 to 6 correspond to amino acid 2, and so on.

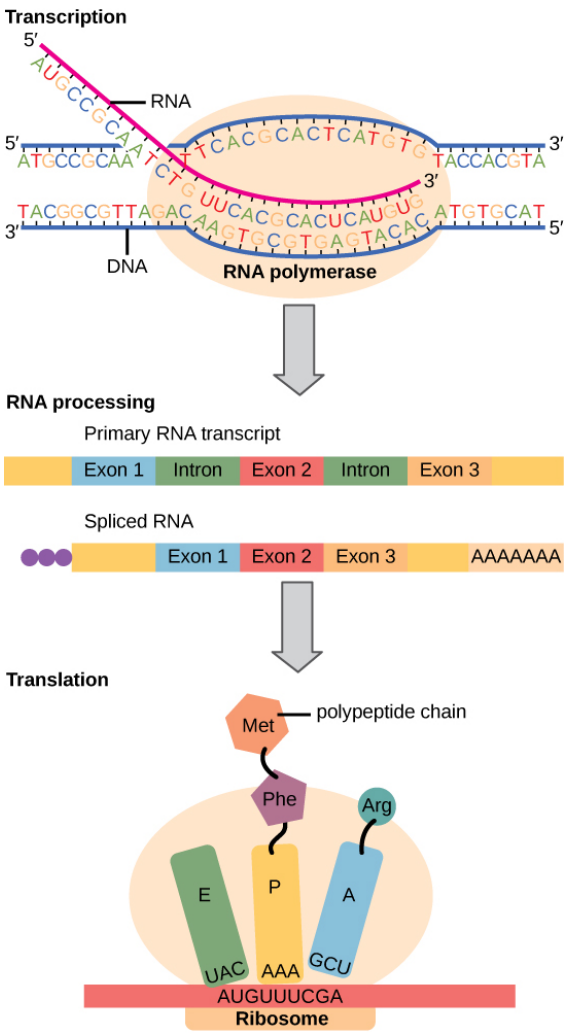


Figure 2. Instructions on DNA are transcribed onto messenger RNA. Ribosomes are able to read the genetic information inscribed on a strand of messenger RNA and use this information to string amino acids together into a protein.

The Genetic Code Is Degenerate and Universal

Given the different numbers of “letters” in the mRNA and protein “alphabets,” scientists theorized that combinations of nucleotides corresponded to single amino acids. Nucleotide doublets would not be sufficient to specify every amino acid because there are only 16 possible two-nucleotide combinations (4^2). In contrast, there are 64 possible nucleotide triplets (4^3), which is far more than the number of amino acids. Scientists theorized that amino acids were encoded by nucleotide triplets and that the genetic code was degenerate. In other words, a given amino acid could be encoded by more than one nucleotide triplet. This was later confirmed experimentally; Francis Crick and Sydney Brenner used the chemical mutagen proflavin to insert one, two, or three nucleotides into the gene of a virus. When one or two nucleotides were inserted, protein synthesis was completely abolished. When three nucleotides were inserted, the protein was synthesized and functional. This demonstrated that three nucleotides specify each amino acid. These nucleotide triplets are called codons. The insertion of one or two nucleotides completely changed the triplet reading frame, thereby altering the message for every subsequent amino acid (Figure 3). Though insertion of three nucleotides caused an extra amino acid to be inserted during translation, the integrity of the rest of the protein was maintained.

Scientists painstakingly solved the genetic code by translating synthetic mRNAs in vitro and sequencing the proteins they specified (Figure 3).

In addition to instructing the addition of a specific amino acid to a polypeptide chain, three of the 64 codons terminate protein synthesis and release the polypeptide from the translation machinery. These triplets are called nonsense codons, or stop codons. Another codon, AUG, also has a special function. In addition to specifying the amino acid methionine, it also serves as the start codon to initiate translation. The reading frame for translation is set by the AUG start codon near the 5′ end of the mRNA.

The genetic code is universal. With a few exceptions, virtually all species use the same genetic code for protein synthesis. Conservation of codons means that a purified mRNA encoding the globin protein in horses could be transferred to a tulip cell, and the tulip would synthesize horse globin. That there is only one genetic code is powerful evidence that all of life on Earth shares a common origin, especially considering that there are about 10^{84} possible combinations of 20 amino acids and 64 triplet codons.

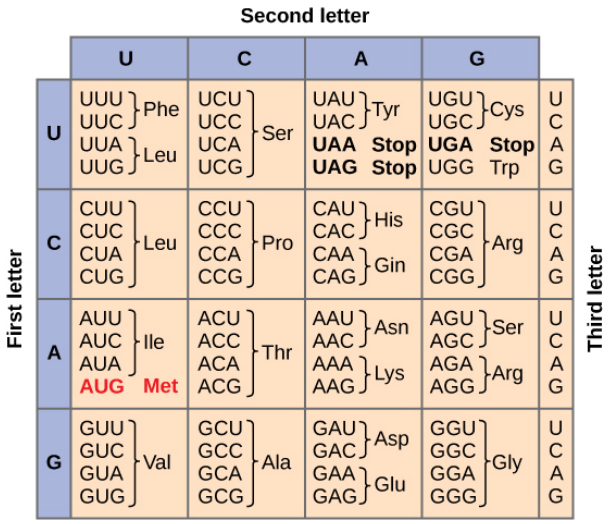


Figure 3. This figure shows the genetic code for translating each nucleotide triplet in mRNA into an amino acid or a termination signal in a nascent protein. (credit: modification of work by NIH)

Link to Learning

Transcribe a gene and translate it to protein using complementary pairing and the genetic code at this site.

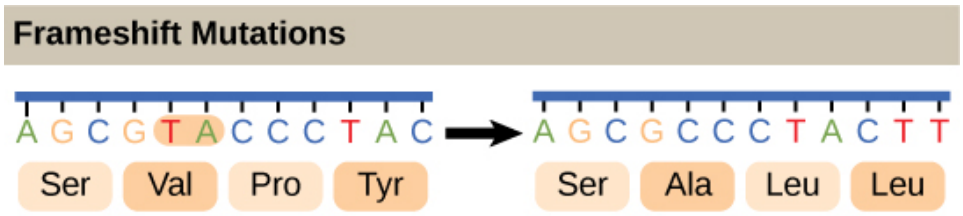


Figure 4. The deletion of two nucleotides shifts the reading frame of an mRNA and changes the entire protein message, creating a nonfunctional protein or terminating protein synthesis altogether.

Degeneracy is believed to be a cellular mechanism to reduce the negative impact of random mutations. Codons that specify the same amino acid typically only differ by one nucleotide. In addition, amino acids with chemically similar side chains are encoded by similar codons. This nuance of the genetic code ensures that a single-nucleotide substitution mutation might either specify the same amino acid but have no effect or specify a similar amino acid, preventing the protein from being rendered completely nonfunctional.

Scientific Method Connection

Which Has More DNA: A Kiwi or a Strawberry?

Question: Would a kiwifruit and strawberry that are approximately the same size (Figure) also have approximately the same amount of DNA?

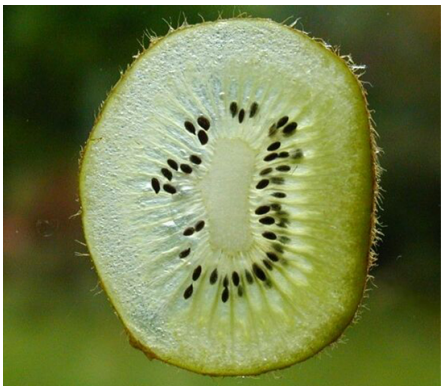


Figure 5. Do you think that a kiwi or a strawberry has more DNA per fruit? (credit "kiwi": "Kelbv"/Flickr; credit: "strawberry": Alisdair McDiarmid)

Background: Genes are carried on chromosomes and are made of DNA. All mammals are diploid, meaning they have two copies of each chromosome. However, not all plants are diploid. The common strawberry is octoploid (8*n*) and the cultivated kiwi is hexaploid (6*n*). Research the total number of chromosomes in the cells of each of these fruits and think about how this might correspond to the amount of DNA in these fruits' cell nuclei. Read about the technique of DNA isolation to understand how each step in the isolation protocol helps liberate and precipitate DNA.

Hypothesis: Hypothesize whether you would be able to detect a difference in DNA quantity from similarly sized strawberries and kiwis. Which fruit do you think would yield more DNA?

Test your hypothesis: Isolate the DNA from a strawberry and a kiwi that are similarly sized. Perform the experiment in at least triplicate for each fruit.

1. Prepare a bottle of DNA extraction buffer from 900 mL water, 50 mL dish detergent, and two teaspoons of table salt. Mix by inversion (cap it and turn it upside down a few times).
2. Grind a strawberry and a kiwifruit by hand in a plastic bag, or using a mortar and pestle, or with a metal bowl and the end of a blunt instrument. Grind for at least two minutes per fruit.
3. Add 10 mL of the DNA extraction buffer to each fruit, and mix well for at least one minute.
4. Remove cellular debris by filtering each fruit mixture through cheesecloth or porous cloth and into a funnel placed in a test tube or an appropriate container.
5. Pour ice-cold ethanol or isopropanol (rubbing alcohol) into the test tube. You should observe white, precipitated DNA.
6. Gather the DNA from each fruit by winding it around separate glass rods.

Record your observations: Because you are not quantitatively measuring DNA volume, you can record for each trial whether the two fruits produced the same or different amounts of DNA as observed by eye. If one or the other fruit produced noticeably more DNA, record this as well. Determine whether your observations are consistent with several pieces of each fruit.

Analyze your data: Did you notice an obvious difference in the amount of DNA produced by each fruit? Were your results reproducible?

Draw a conclusion: Given what you know about the number of chromosomes in each fruit, can you conclude that chromosome number necessarily correlates to DNA amount? Can you identify any drawbacks to this procedure? If you had access to a laboratory, how could you standardize your comparison and make it more quantitative?

Section Summary

The genetic code refers to the DNA alphabet (A, T, C, G), the RNA alphabet (A, U, C, G), and the polypeptide alphabet (20 amino acids). The Central Dogma describes the flow of genetic information in the cell from genes to mRNA to proteins. Genes are used to make mRNA by the process of transcription; mRNA is used to synthesize proteins by the process of translation. The genetic code is degenerate because 64 triplet codons in mRNA specify only 20 amino acids and three nonsense codons. Almost every species on the planet uses the same genetic code.

Visit this page in your course online to check your understanding.

Additional Self Test Questions

1. Imagine if there were 200 commonly occurring amino acids instead of 20. Given what you know about the genetic code, what would be the shortest possible codon length? Explain.

2. Discuss how degeneracy of the genetic code makes cells more robust to mutations.

Answers

1. For 200 commonly occurring amino acids, codons consisting of four types of nucleotides would have to be at least four nucleotides long, because $4^4 = 256$. There would be much less degeneracy in this case.

2. Codons that specify the same amino acid typically only differ by one nucleotide. In addition, amino acids with chemically similar side chains are encoded by similar codons. This nuance of the genetic code ensures that a single-nucleotide substitution mutation might either specify the same amino acid and have no effect, or may specify a similar amino acid, preventing the protein from being rendered completely nonfunctional.

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: Open Stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License**: *CC BY: Attribution*

PROKARYOTIC TRANSCRIPTION

Prokaryotic Transcription

The prokaryotes, which include bacteria and archaea, are mostly single-celled organisms that, by definition, lack membrane-bound nuclei and other organelles. A bacterial chromosome is a covalently closed circle that, unlike eukaryotic chromosomes, is not organized around histone proteins. The central region of the cell in which prokaryotic DNA resides is called the nucleoid. In addition, prokaryotes often have abundant plasmids, which are shorter circular DNA molecules that may only contain one or a few genes. Plasmids can be transferred independently of the bacterial chromosome during cell division and often carry traits such as antibiotic resistance.

Transcription in prokaryotes (and in eukaryotes) requires the DNA double helix to partially unwind in the region of mRNA synthesis. The region of unwinding is called a transcription bubble. Transcription always proceeds from the same DNA strand for each gene, which is called the template strand. The mRNA product is complementary to the template strand and is almost identical to the other DNA strand, called the nontemplate strand. The only difference is that in mRNA, all of the T nucleotides are replaced with U nucleotides. In an RNA double helix, A can bind U via two hydrogen bonds, just as in A–T pairing in a DNA double helix.

The nucleotide pair in the DNA double helix that corresponds to the site from which the first 5′ mRNA nucleotide is transcribed is called the +1 site, or the initiation site. Nucleotides preceding the initiation site are given negative numbers and are designatedupstream. Conversely, nucleotides following the initiation site are denoted with “+” numbering and are called downstreamnucleotides.

Initiation of Transcription in Prokaryotes

Prokaryotes do not have membrane-enclosed nuclei. Therefore, the processes of transcription, translation, and mRNA degradation can all occur simultaneously. The intracellular level of a bacterial protein can quickly be amplified by multiple transcription and translation events occurring concurrently on the same DNA template. Prokaryotic transcription often covers more than one gene and produces polycistronic mRNAs that specify more than one protein.

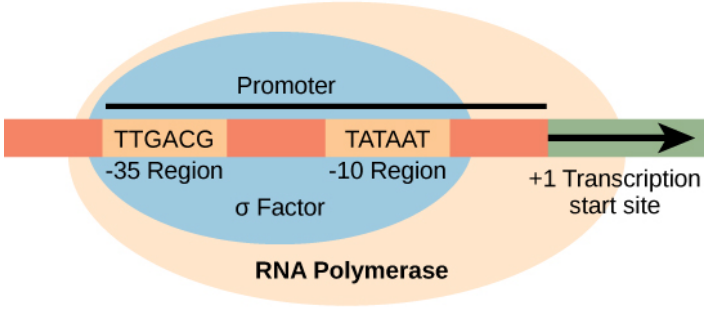
Our discussion here will exemplify transcription by describing this process in *Escherichia coli*, a well-studied bacterial species. Although some differences exist between transcription in *E. coli* and transcription in archaea, an understanding of *E. coli*transcription can be applied to virtually all bacterial species.

Prokaryotic RNA Polymerase

Prokaryotes use the same RNA polymerase to transcribe all of their genes. In *E. coli*, the polymerase is composed of five polypeptide subunits, two of which are identical. Four of these subunits, denoted α, α, β, and β' comprise the polymerase core enzyme. These subunits assemble every time a gene is transcribed, and they disassemble once transcription is complete. Each subunit has a unique role; the two α-subunits are necessary to assemble the polymerase on the DNA; the β-subunit binds to the ribonucleoside triphosphate that will become part of the nascent “recently born” mRNA molecule; and the β' binds the DNA template strand. The fifth subunit, σ, is involved only in transcription initiation. It confers transcriptional specificity such that the polymerase begins to synthesize mRNA from an appropriate initiation site. Without σ, the core enzyme would transcribe from random sites and would produce mRNA molecules that specified protein gibberish. The polymerase comprised of all five subunits is called the holoenzyme.

Prokaryotic Promoters

A promoter is a DNA sequence onto which the transcription machinery binds and initiates transcription. In most cases, promoters exist upstream of the genes they regulate. The specific sequence of a promoter is very important because it determines whether the corresponding gene is transcribed all the time, some of the time, or infrequently. Although promoters vary among prokaryotic genomes, a few elements are conserved. At the -10 and -35 regions upstream of the initiation site, there are two promoterconsensus sequences, or regions that are similar across all promoters and across various bacterial species (Figure). The -10 consensus sequence, called the -10 region, is TATAAT. The -35 sequence, TTGACA, is recognized and bound by σ. Once this interaction is made, the subunits of the core enzyme bind to the site. The A–T-rich -10 region facilitates unwinding of the DNA template, and several phosphodiester bonds are made. The transcription initiation phase ends with the production of abortive transcripts, which are polymers of approximately 10 nucleotides that are made and released.



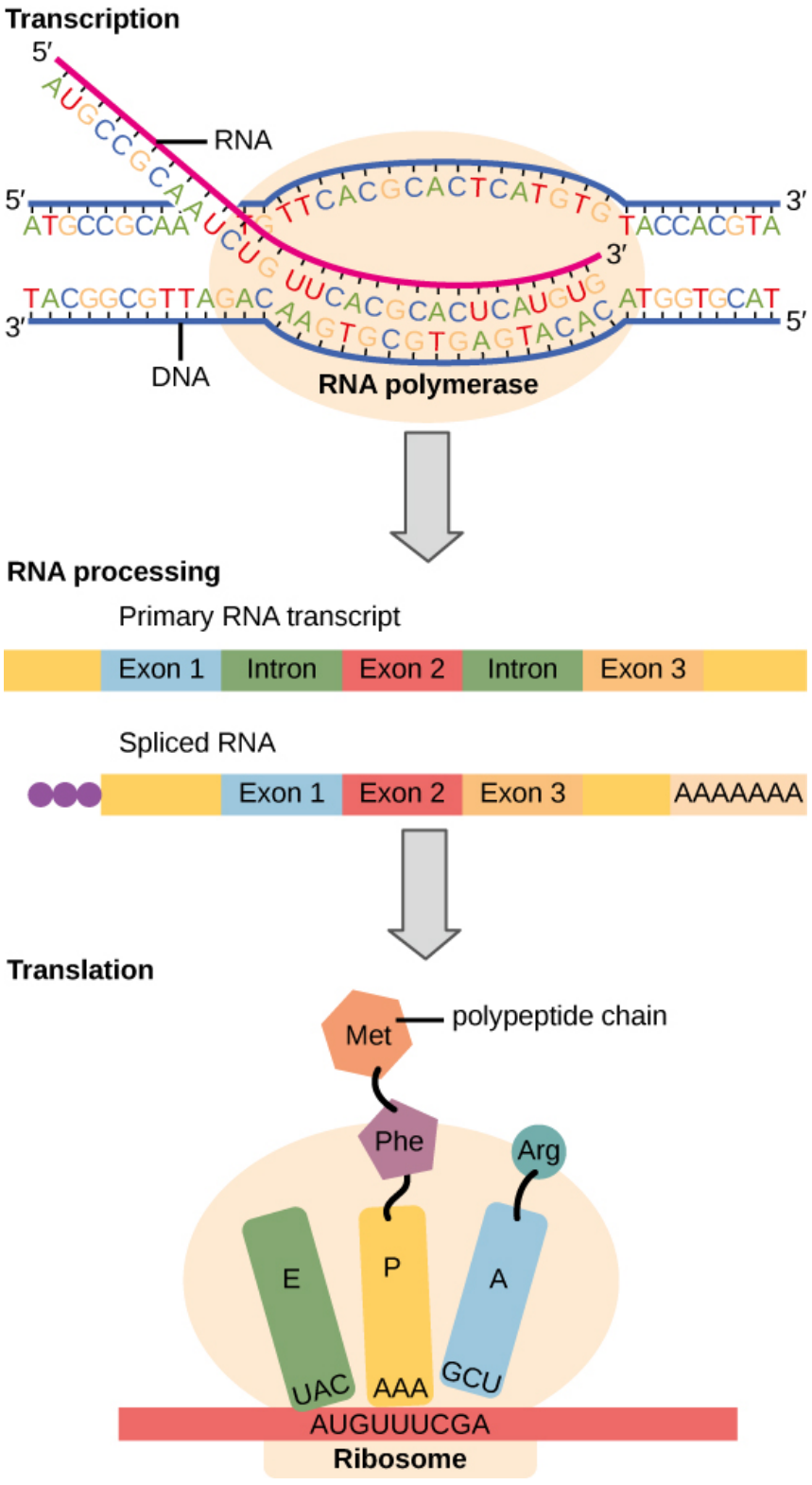
The σ subunit of prokaryotic RNA polymerase recognizes consensus sequences found in the promoter region upstream of the transcription start sight. The σ subunit dissociates from the polymerase after transcription has been initiated.

Link to Learning

View this MolecularMovies animation to see the first part of transcription and the base sequence repetition of the TATA box.

Elongation and Termination in Prokaryotes

The transcription elongation phase begins with the release of the σ subunit from the polymerase. The dissociation of σ allows the core enzyme to proceed along the DNA template, synthesizing mRNA in the 5′ to 3′ direction at a rate of approximately 40 nucleotides per second. As elongation proceeds, the DNA is continuously unwound ahead of the core enzyme and rewound behind it (Figure). The base pairing between DNA and RNA is not stable enough to maintain the stability of the mRNA synthesis components. Instead, the RNA polymerase acts as a stable linker between the DNA template and the nascent RNA strands to ensure that elongation is not interrupted prematurely.



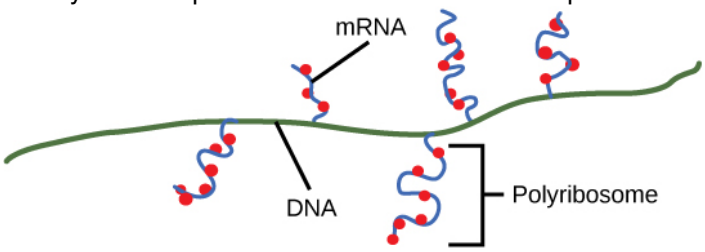
During elongation, the prokaryotic RNA polymerase tracks along the DNA template, synthesizes mRNA in the 5′ to 3′ direction, and unwinds and rewinds the DNA as it is read.

Prokaryotic Termination Signals

Once a gene is transcribed, the prokaryotic polymerase needs to be instructed to dissociate from the DNA template and liberate the newly made mRNA. Depending on the gene being transcribed, there are two kinds of termination signals. One is protein-based and the other is RNA-based. Rho-dependent termination is controlled by the rho protein, which tracks along behind the polymerase on the growing mRNA chain. Near the end of the gene, the polymerase encounters a run of G nucleotides on the DNA template and it stalls. As a result, the rho protein collides with the polymerase. The interaction with rho releases the mRNA from the transcription bubble.

Rho-independent termination is controlled by specific sequences in the DNA template strand. As the polymerase nears the end of the gene being transcribed, it encounters a region rich in C–G nucleotides. The mRNA folds back on itself, and the complementary C–G nucleotides bind together. The result is a stable hairpin that causes the polymerase to stall as soon as it begins to transcribe a region rich in A–T nucleotides. The complementary U–A region of the mRNA transcript forms only a weak interaction with the template DNA. This, coupled with the stalled polymerase, induces enough instability for the core enzyme to break away and liberate the new mRNA transcript.

Upon termination, the process of transcription is complete. By the time termination occurs, the prokaryotic transcript would already have been used to begin synthesis of numerous copies of the encoded protein because these processes can occur concurrently. The unification of transcription, translation, and even mRNA degradation is possible because all of these processes occur in the same 5′ to 3′ direction, and because there is no membranous compartmentalization in the prokaryotic cell (Figure). In contrast, the presence of a nucleus in eukaryotic cells precludes simultaneous transcription and translation.



Multiple polymerases can transcribe a single bacterial gene while numerous ribosomes concurrently translate the mRNA transcripts into polypeptides. In this way, a specific protein can rapidly reach a high concentration in the bacterial cell.

Link to Learning

Visit this BioStudio animation to see the process of prokaryotic transcription.

Section Summary

In prokaryotes, mRNA synthesis is initiated at a promoter sequence on the DNA template comprising two consensus sequences that recruit RNA polymerase. The prokaryotic polymerase consists of a core enzyme of four protein subunits and a σ protein that assists only with initiation. Elongation synthesizes mRNA in the 5′ to 3′ direction at a rate of 40 nucleotides per second. Termination liberates the mRNA and occurs either by rho protein interaction or by the formation of an mRNA hairpin.

Review Questions

Which subunit of the *E. coli* polymerase confers specificity to transcription?

1. α
2. β
3. β'
4. σ

The -10 and -35 regions of prokaryotic promoters are called consensus sequences because _____.

1. they are identical in all bacterial species
2. they are similar in all bacterial species
3. they exist in all organisms
4. they have the same function in all organisms

Free Response

If mRNA is complementary to the DNA template strand and the DNA template strand is complementary to the DNA nontemplate strand, then why are base sequences of mRNA and the DNA nontemplate strand not identical? Could they ever be?

In your own words, describe the difference between rho-dependent and rho-independent termination of transcription in prokaryotes.

Glossary

consensus DNA sequence that is used by many species to perform the same or similar functions

core enzyme prokaryotic RNA polymerase consisting of α, α, β, and β' but missing σ; this complex performs elongation

downstream nucleotides following the initiation site in the direction of mRNA transcription; in general, sequences that are toward the 3′ end relative to a site on the mRNA

hairpin structure of RNA when it folds back on itself and forms intramolecular hydrogen bonds between complementary nucleotides

holoenzyme prokaryotic RNA polymerase consisting of α, α, β, β', and σ; this complex is responsible for transcription initiation

initiation site nucleotide from which mRNA synthesis proceeds in the 5′ to 3′ direction; denoted with a “+1”

nontemplate strand strand of DNA that is not used to transcribe mRNA; this strand is identical to the mRNA except that T nucleotides in the DNA are replaced by U nucleotides in the mRNA

plasmid extrachromosomal, covalently closed, circular DNA molecule that may only contain one or a few genes; common in prokaryotes

promoter DNA sequence to which RNA polymerase and associated factors bind and initiate transcription

Rho-dependent termination in prokaryotes, termination of transcription by an interaction between RNA polymerase and the rho protein at a run of G nucleotides on the DNA template

Rho-independent termination sequence-dependent termination of prokaryotic mRNA synthesis; caused by hairpin formation in the mRNA that stalls the polymerase

TATA box conserved promoter sequence in eukaryotes and prokaryotes that helps to establish the initiation site for transcription

template strand strand of DNA that specifies the complementary mRNA molecule

transcription bubble region of locally unwound DNA that allows for transcription of mRNA

upstream nucleotides preceding the initiation site; in general, sequences toward the 5′ end relative to a site on the mRNA

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: OpenStax College. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License**: *CC BY: Attribution*

Learning Objectives

By the end of this section, you will be able to:

- List the steps in eukaryotic transcription
- Discuss the role of RNA polymerases in transcription
- Compare and contrast the three RNA polymerases
- Explain the significance of transcription factors

Prokaryotes and eukaryotes perform fundamentally the same process of transcription, with a few key differences. The most important difference between prokaryotes and eukaryotes is the latter's membrane-bound nucleus and organelles. With the genes bound in a nucleus, the eukaryotic cell must be able to transport its mRNA to the cytoplasm and must protect its mRNA from degrading before it is translated. Eukaryotes also employ three different polymerases that each transcribe a different subset of genes. Eukaryotic mRNAs are usually monogenic, meaning that they specify a single protein.

Initiation of Transcription in Eukaryotes

Unlike the prokaryotic polymerase that can bind to a DNA template on its own, eukaryotes require several other proteins, called transcription factors, to first bind to the promoter region and then help recruit the appropriate polymerase.

The Three Eukaryotic RNA Polymerases

The features of eukaryotic mRNA synthesis are markedly more complex those of prokaryotes. Instead of a single polymerase comprising five subunits, the eukaryotes have three polymerases that are each made up of 10 subunits or more. Each eukaryotic polymerase also requires a distinct set of transcription factors to bring it to the DNA template.

RNA polymerase I is located in the nucleolus, a specialized nuclear substructure in which ribosomal RNA (rRNA) is transcribed, processed, and assembled into ribosomes (Table 1). The rRNA molecules are considered structural RNAs because they have a cellular role but are not translated into protein. The rRNAs are components of the ribosome and are essential to the process of translation. RNA polymerase I synthesizes all of the rRNAs except for the 5S rRNA molecule. The "S" designation applies to "Svedberg" units, a nonadditive value that characterizes the speed at which a particle sediments during centrifugation.

Table 1. Locations, Products, and Sensitivities of the Three Eukaryotic RNA Polymerases

RNA Polymerase	Cellular Compartment	Product of Transcription	α-Amanitin Sensitivity
I	Nucleolus	All rRNAs except 5S rRNA	Insensitive
II	Nucleus	All protein-coding nuclear pre-mRNAs	Extremely sensitive

Table 1. Locations, Products, and Sensitivities of the Three Eukaryotic RNA Polymerases			
RNA Polymerase	Cellular Compartment	Product of Transcription	α-Amanitin Sensitivity
III	Nucleus	5S rRNA, tRNAs, and small nuclear RNAs	Moderately sensitive

RNA polymerase II is located in the nucleus and synthesizes all protein-coding nuclear pre-mRNAs. Eukaryotic pre-mRNAs undergo extensive processing after transcription but before translation. For clarity, this module's discussion of transcription and translation in eukaryotes will use the term "mRNAs" to describe only the mature, processed molecules that are ready to be translated. RNA polymerase II is responsible for transcribing the overwhelming majority of eukaryotic genes.

RNA polymerase III is also located in the nucleus. This polymerase transcribes a variety of structural RNAs that includes the 5S pre-rRNA, transfer pre-RNAs (pre-tRNAs), and small nuclear pre-RNAs. The tRNAs have a critical role in translation; they serve as the adaptor molecules between the mRNA template and the growing polypeptide chain. Small nuclear RNAs have a variety of functions, including "splicing" pre-mRNAs and regulating transcription factors.

A scientist characterizing a new gene can determine which polymerase transcribes it by testing whether the gene is expressed in the presence of a particular mushroom poison, α-amanitin (Table 1). Interestingly, α-amanitin produced by *Amanita phalloides*, the Death Cap mushroom, affects the three polymerases very differently. RNA polymerase I is completely insensitive to α-amanitin, meaning that the polymerase can transcribe DNA in vitro in the presence of this poison. In contrast, RNA polymerase II is extremely sensitive to α-amanitin, and RNA polymerase III is moderately sensitive. Knowing the transcribing polymerase can clue a researcher into the general function of the gene being studied. Because RNA polymerase II transcribes the vast majority of genes, we will focus on this polymerase in our subsequent discussions about eukaryotic transcription factors and promoters.

Structure of an RNA Polymerase II Promoter

Eukaryotic promoters are much larger and more complex than prokaryotic promoters, but both have a TATA box. For example, in the mouse thymidine kinase gene, the TATA box is located at approximately -30 relative to the initiation (+1) site (Figure 1). For this gene, the exact TATA box sequence is TATAAAA, as read in the 5′ to 3′ direction on the nontemplate strand. This sequence is not identical to the *E. coli* TATA box, but it conserves the A–T rich element. The thermostability of A–T bonds is low and this helps the DNA template to locally unwind in preparation for transcription.

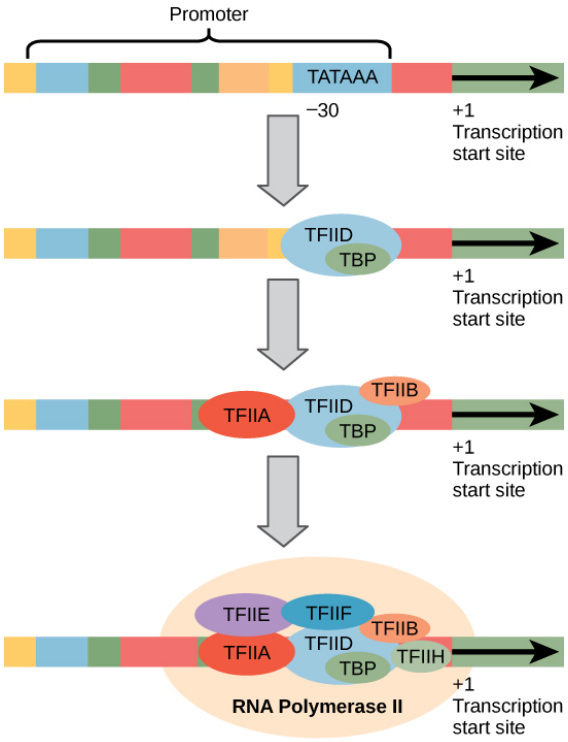


Figure 1. A generalized promoter of a gene transcribed by RNA polymerase II is shown. Transcription factors recognize the promoter. RNA polymerase II then binds and forms the transcription initiation complex.

Art Connection

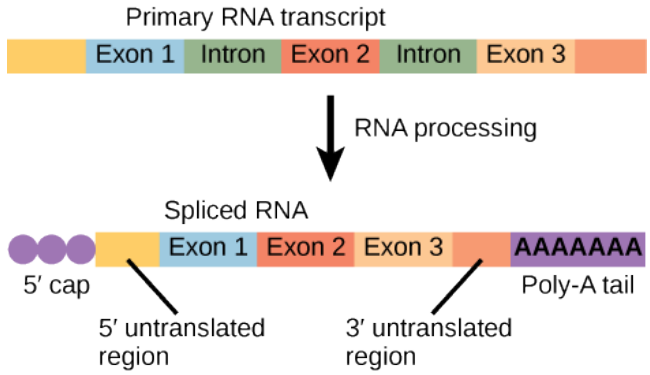


Figure 2. Eukaryotic mRNA contains introns that must be spliced out. A 5′ cap and 3′ poly-A tail are also added.

A scientist splices a eukaryotic promoter in front of a bacterial gene and inserts the gene in a bacterial chromosome. Would you expect the bacteria to transcribe the gene?

The mouse genome includes one gene and two pseudogenes for cytoplasmic thymidine kinase. Pseudogenes are genes that have lost their protein-coding ability or are no longer expressed by the cell. These pseudogenes are copied from mRNA and incorporated into the chromosome. For example, the mouse thymidine kinase promoter also has a conserved CAAT box (GGCCAATCT) at approximately -80. This sequence is essential and is involved in binding transcription factors. Further upstream of the TATA box, eukaryotic promoters may also contain one or more GC-rich boxes (GGCG) or octamer boxes (ATTTGCAT). These elements bind cellular factors that increase the efficiency of transcription initiation and are often identified in more “active” genes that are constantly being expressed by the cell.

Transcription Factors for RNA Polymerase II

The complexity of eukaryotic transcription does not end with the polymerases and promoters. An army of basal transcription factors, enhancers, and silencers also help to regulate the frequency with which pre-mRNA is synthesized from a gene. Enhancers and silencers affect the efficiency of transcription but are not necessary for transcription to proceed. Basal transcription factors are crucial in the formation of a preinitiation complex on the DNA template that subsequently recruits RNA polymerase II for transcription initiation.

The names of the basal transcription factors begin with "TFII" (this is the transcription factor for RNA polymerase II) and are specified with the letters A–J. The transcription factors systematically fall into place on the DNA template, with each one further stabilizing the preinitiation complex and contributing to the recruitment of RNA polymerase II.

The processes of bringing RNA polymerases I and III to the DNA template involve slightly less complex collections of transcription factors, but the general theme is the same. Eukaryotic transcription is a tightly regulated process that requires a variety of proteins to interact with each other and with the DNA strand. Although the process of transcription in eukaryotes involves a greater metabolic investment than in prokaryotes, it ensures that the cell transcribes precisely the pre-mRNAs that it needs for protein synthesis.

Evolution Connection

The Evolution of Promoters

The evolution of genes may be a familiar concept. Mutations can occur in genes during DNA replication, and the result may or may not be beneficial to the cell. By altering an enzyme, structural protein, or some other factor, the process of mutation can transform functions or physical features. However, eukaryotic promoters and other gene regulatory sequences may evolve as well. For instance, consider a gene that, over many generations, becomes more valuable to the cell. Maybe the gene encodes a structural protein that the cell needs to synthesize in abundance for a certain function. If this is the case, it would be beneficial to the cell for that gene's promoter to recruit transcription factors more efficiently and increase gene expression.

Scientists examining the evolution of promoter sequences have reported varying results. In part, this is because it is difficult to infer exactly where a eukaryotic promoter begins and ends. Some promoters occur within genes; others are located very far upstream, or even downstream, of the genes they are regulating. However, when researchers limited their examination to human core promoter sequences that were defined experimentally as sequences that bind the preinitiation complex, they found that promoters evolve even faster than protein-coding genes.

It is still unclear how promoter evolution might correspond to the evolution of humans or other higher organisms. However, the evolution of a promoter to effectively make more or less of a given gene product is an intriguing alternative to the evolution of the genes themselves.

Promoter Structures for RNA Polymerases I and III

In eukaryotes, the conserved promoter elements differ for genes transcribed by RNA polymerases I, II, and III. RNA polymerase I transcribes genes that have two GC-rich promoter sequences in the -45 to +20 region. These sequences alone are sufficient for transcription initiation to occur, but promoters with additional sequences in the region from -180 to -105 upstream of the initiation site will further enhance initiation. Genes that are transcribed by RNA polymerase III have upstream promoters or promoters that occur within the genes themselves.

Eukaryotic Elongation and Termination

Following the formation of the preinitiation complex, the polymerase is released from the other transcription factors, and elongation is allowed to proceed as it does in prokaryotes with the polymerase synthesizing pre-

mRNA in the 5′ to 3′ direction. As discussed previously, RNA polymerase II transcribes the major share of eukaryotic genes, so this section will focus on how this polymerase accomplishes elongation and termination.

Although the enzymatic process of elongation is essentially the same in eukaryotes and prokaryotes, the DNA template is more complex. When eukaryotic cells are not dividing, their genes exist as a diffuse mass of DNA and proteins called chromatin. The DNA is tightly packaged around charged histone proteins at repeated intervals. These DNA–histone complexes, collectively called nucleosomes, are regularly spaced and include 146 nucleotides of DNA wound around eight histones like thread around a spool.

For polynucleotide synthesis to occur, the transcription machinery needs to move histones out of the way every time it encounters a nucleosome. This is accomplished by a special protein complex called FACT, which stands for "facilitates chromatin transcription." This complex pulls histones away from the DNA template as the polymerase moves along it. Once the pre-mRNA is synthesized, the FACT complex replaces the histones to recreate the nucleosomes.

The termination of transcription is different for the different polymerases. Unlike in prokaryotes, elongation by RNA polymerase II in eukaryotes takes place 1,000–2,000 nucleotides beyond the end of the gene being transcribed. This pre-mRNA tail is subsequently removed by cleavage during mRNA processing. On the other hand, RNA polymerases I and III require termination signals. Genes transcribed by RNA polymerase I contain a specific 18-nucleotide sequence that is recognized by a termination protein. The process of termination in RNA polymerase III involves an mRNA hairpin similar to rho-independent termination of transcription in prokaryotes.

Section Summary

Transcription in eukaryotes involves one of three types of polymerases, depending on the gene being transcribed. RNA polymerase II transcribes all of the protein-coding genes, whereas RNA polymerase I transcribes rRNA genes, and RNA polymerase III transcribes rRNA, tRNA, and small nuclear RNA genes. The initiation of transcription in eukaryotes involves the binding of several transcription factors to complex promoter sequences that are usually located upstream of the gene being copied. The mRNA is synthesized in the 5′ to 3′ direction, and the FACT complex moves and reassembles nucleosomes as the polymerase passes by. Whereas RNA polymerases I and III terminate transcription by protein- or RNA hairpin-dependent methods, RNA polymerase II transcribes for 1,000 or more nucleotides beyond the gene template and cleaves the excess during pre-mRNA processing.

Visit this page in your course online to check your understanding.

Additional Self Check Question

1. A scientist splices a eukaryotic promoter in front of a bacterial gene and inserts the gene in a bacterial chromosome. Would you expect the bacteria to transcribe the gene?

Answer

1. No. Prokaryotes use different promoters than eukaryotes.

Glossary

CAAT box: (GGCCAATCT) essential eukaryotic promoter sequence involved in binding transcription factors

FACT: complex that "facilitates chromatin transcription" by disassembling nucleosomes ahead of a transcribing RNA polymerase II and reassembling them after the polymerase passes by

GC-rich box: (GGCG) nonessential eukaryotic promoter sequence that binds cellular factors to increase the efficiency of transcription; may be present several times in a promoter

Octamer box: (ATTTGCAT) nonessential eukaryotic promoter sequence that binds cellular factors to increase the efficiency of transcription; may be present several times in a promoter

preinitiation complex: cluster of transcription factors and other proteins that recruit RNA polymerase II for transcription of a DNA template

small nuclear RNA: molecules synthesized by RNA polymerase III that have a variety of functions, including splicing pre-mRNAs and regulating transcription factors

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by:** Open Stax. **Located at:** http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License:** *CC BY: Attribution*

RNA PROCESSING IN EUKARYOTES

RNA Processing in Eukaryotes

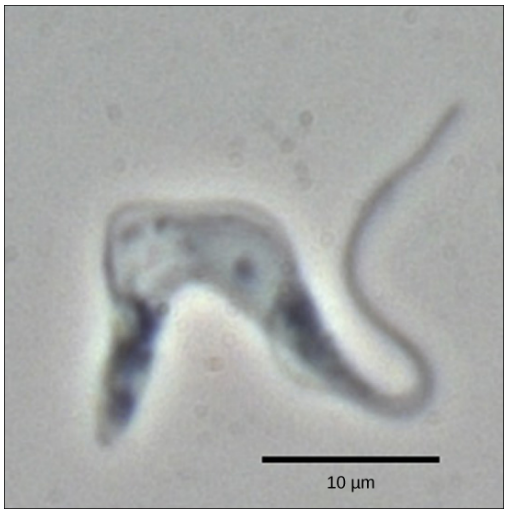
After transcription, eukaryotic pre-mRNAs must undergo several processing steps before they can be translated. Eukaryotic (and prokaryotic) tRNAs and rRNAs also undergo processing before they can function as components in the protein synthesis machinery.

mRNA Processing

The eukaryotic pre-mRNA undergoes extensive processing before it is ready to be translated. The additional steps involved in eukaryotic mRNA maturation create a molecule with a much longer half-life than a prokaryotic mRNA. Eukaryotic mRNAs last for several hours, whereas the typical *E. coli* mRNA lasts no more than five seconds.

Pre-mRNAs are first coated in RNA-stabilizing proteins; these protect the pre-mRNA from degradation while it is processed and exported out of the nucleus. The three most important steps of pre-mRNA processing are the addition of stabilizing and signaling factors at the 5′ and 3′ ends of the molecule, and the removal of intervening sequences that do not specify the appropriate amino acids. In rare cases, the mRNA transcript can be “edited” after it is transcribed.
EVOLUTION CONNECTION

RNA Editing in TrypanosomesThe trypanosomes are a group of protozoa that include the pathogen *Trypanosoma brucei*, which causes sleeping sickness in humans (Figure). Trypanosomes, and virtually all other eukaryotes, have organelles called mitochondria that supply the cell with chemical energy. Mitochondria are organelles that express their own DNA and are believed to be the remnants of a symbiotic relationship between a eukaryote and an engulfed prokaryote. The mitochondrial DNA of trypanosomes exhibit an interesting exception to The Central Dogma: their pre-mRNAs do not have the correct information to specify a functional protein. Usually, this is because the mRNA is missing several U nucleotides. The cell performs an additional RNA processing step called RNA editing to remedy this.



Trypanosoma brucei is the causative agent of sleeping sickness in humans. The mRNAs of this pathogen must be modified by the addition of nucleotides before protein synthesis can occur. (credit: modification of work by Torsten Ochsenreiter)

Other genes in the mitochondrial genome encode 40- to 80-nucleotide guide RNAs. One or more of these molecules interacts by complementary base pairing with some of the nucleotides in the pre-mRNA transcript. However, the guide RNA has more A nucleotides than the pre-mRNA has U nucleotides to bind with. In these regions, the guide RNA loops out. The 3′ ends of guide RNAs have a long poly-U tail, and these U bases are inserted in regions of the pre-mRNA transcript at which the guide RNAs are looped. This process is entirely mediated by RNA molecules. That is, guide RNAs—rather than proteins—serve as the catalysts in RNA editing.

RNA editing is not just a phenomenon of trypanosomes. In the mitochondria of some plants, almost all pre-mRNAs are edited. RNA editing has also been identified in mammals such as rats, rabbits, and even humans. What could be the evolutionary reason for this additional step in pre-mRNA processing? One possibility is that the mitochondria, being remnants of ancient prokaryotes, have an equally ancient RNA-based method for regulating gene expression. In support of this hypothesis, edits made to pre-mRNAs differ depending on cellular conditions. Although speculative, the process of RNA editing may be a holdover from a primordial time when RNA molecules, instead of proteins, were responsible for catalyzing reactions.

5′ Capping

While the pre-mRNA is still being synthesized, a 7-methylguanosine cap is added to the 5′ end of the growing transcript by a phosphate linkage. This moiety (functional group) protects the nascent mRNA from degradation. In addition, factors involved in protein synthesis recognize the cap to help initiate translation by ribosomes.

3′ Poly-A Tail

Once elongation is complete, the pre-mRNA is cleaved by an endonuclease between an AAUAAA consensus sequence and a GU-rich sequence, leaving the AAUAAA sequence on the pre-mRNA. An enzyme called poly-A polymerase then adds a string of approximately 200 A residues, called the poly-A tail. This modification further protects the pre-mRNA from degradation and signals the export of the cellular factors that the transcript needs to the cytoplasm.

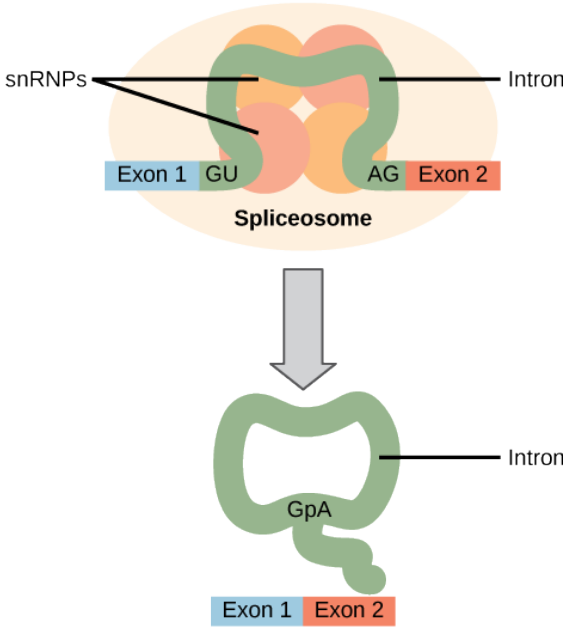
Pre-mRNA Splicing

Eukaryotic genes are composed of exons, which correspond to protein-coding sequences (*ex*-on signifies that they are*ex*pressed), and *int*ervening sequences called introns (*int*-ron denotes their *int*ervening role), which may be involved in gene regulation but are removed from the pre-mRNA during processing. Intron sequences in mRNA do not encode functional proteins.

The discovery of introns came as a surprise to researchers in the 1970s who expected that pre-mRNAs would specify protein sequences without further processing, as they had observed in prokaryotes. The genes of higher eukaryotes very often contain one or more introns. These regions may correspond to regulatory sequences; however, the biological significance of having many introns or having very long introns in a gene is unclear. It is possible that introns slow down gene expression because it takes longer to transcribe pre-mRNAs with lots of introns. Alternatively, introns may be nonfunctional sequence remnants left over from the fusion of ancient genes throughout evolution. This is supported by the fact that separate exons often encode separate protein subunits or domains. For the most part, the sequences of introns can be mutated without ultimately affecting the protein product.

All of a pre-mRNA's introns must be completely and precisely removed before protein synthesis. If the process errs by even a single nucleotide, the reading frame of the rejoined exons would shift, and the resulting protein would be dysfunctional. The process of removing introns and reconnecting exons is called splicing (Figure). Introns are removed and degraded while the pre-mRNA is still in the nucleus. Splicing occurs by a sequence-specific mechanism that ensures introns will be removed and exons rejoined with the accuracy and precision of a single nucleotide. The splicing of pre-mRNAs is conducted by complexes of proteins and RNA molecules called spliceosomes.

ART CONNECTION



Pre-mRNA splicing involves the precise removal of introns from the primary RNA transcript. The splicing process is catalyzed by protein complexes called spliceosomes that are composed of proteins and RNA molecules called snRNAs. Spliceosomes recognize sequences at the 5′ and 3′ end of the intron.

Errors in splicing are implicated in cancers and other human diseases. What kinds of mutations might lead to splicing errors? Think of different possible outcomes if splicing errors occur.

Note that more than 70 individual introns can be present, and each has to undergo the process of splicing—in addition to 5′ capping and the addition of a poly-A tail—just to generate a single, translatable mRNA molecule.

Link to Learning

See how introns are removed during RNA splicing at this website.

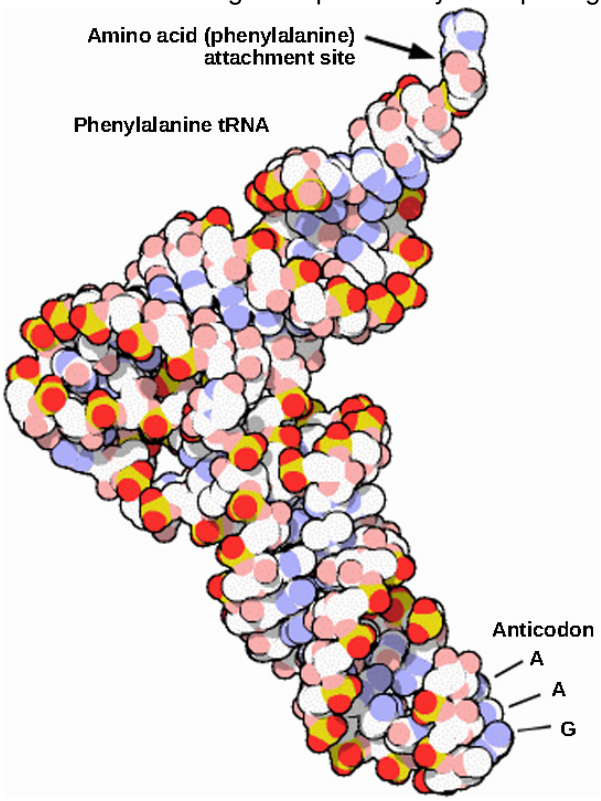
Processing of tRNAs and rRNAs

The tRNAs and rRNAs are structural molecules that have roles in protein synthesis; however, these RNAs are not themselves translated. Pre-rRNAs are transcribed, processed, and assembled into ribosomes in the nucleolus.

Pre-tRNAs are transcribed and processed in the nucleus and then released into the cytoplasm where they are linked to free amino acids for protein synthesis.

Most of the tRNAs and rRNAs in eukaryotes and prokaryotes are first transcribed as a long precursor molecule that spans multiple rRNAs or tRNAs. Enzymes then cleave the precursors into subunits corresponding to each structural RNA. Some of the bases of pre-rRNAs are methylated; that is, a $–CH_3$ moiety (methyl functional group) is added for stability. Pre-tRNA molecules also undergo methylation. As with pre-mRNAs, subunit excision occurs in eukaryotic pre-RNAs destined to become tRNAs or rRNAs.

Mature rRNAs make up approximately 50 percent of each ribosome. Some of a ribosome's RNA molecules are purely structural, whereas others have catalytic or binding activities. Mature tRNAs take on a three-dimensional structure through intramolecular hydrogen bonding to position the amino acid binding site at one end and the anticodon at the other end (Figure). The anticodon is a three-nucleotide sequence in a tRNA that interacts with an mRNA codon through complementary base pairing.



This is a space-filling model of a tRNA molecule that adds the amino acid phenylalanine to a growing polypeptide chain. The anticodon AAG binds the Codon UUC on the mRNA. The amino acid phenylalanine is attached to the other end of the tRNA.

Section Summary

Eukaryotic pre-mRNAs are modified with a 5′ methylguanosine cap and a poly-A tail. These structures protect the mature mRNA from degradation and help export it from the nucleus. Pre-mRNAs also undergo splicing, in which introns are removed and exons are reconnected with single-nucleotide accuracy. Only finished mRNAs that have undergone 5′ capping, 3′ polyadenylation, and intron splicing are exported from the nucleus to the cytoplasm. Pre-rRNAs and pre-tRNAs may be processed by intramolecular cleavage, splicing, methylation, and chemical conversion of nucleotides. Rarely, RNA editing is also performed to insert missing bases after an mRNA has been synthesized.

Art Connections

Figure Errors in splicing are implicated in cancers and other human diseases. What kinds of mutations might lead to splicing errors? Think of different possible outcomes if splicing errors occur.

Review Questions

Which pre-mRNA processing step is important for initiating translation?

1. poly-A tail
2. RNA editing
3. splicing
4. 7-methylguanosine cap

What processing step enhances the stability of pre-tRNAs and pre-rRNAs?

1. methylation
2. nucleotide modification
3. cleavage
4. splicing

Glossary

7-methylguanosine cap modification added to the 5′ end of pre-mRNAs to protect mRNA from degradation and assist translation

anticodon three-nucleotide sequence in a tRNA molecule that corresponds to an mRNA codon

exon sequence present in protein-coding mRNA after completion of pre-mRNA splicing

intron non–protein-coding intervening sequences that are spliced from mRNA during processing

poly-A tail modification added to the 3′ end of pre-mRNAs to protect mRNA from degradation and assist mRNA export from the nucleus

RNA editing direct alteration of one or more nucleotides in an mRNA that has already been synthesized

splicing process of removing introns and reconnecting exons in a pre-mRNA

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: OpenStax College. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License**: *CC BY: Attribution*

RIBOSOMES AND PROTEIN SYNTHESIS

Learning Objectives

By the end of this section, you will be able to:

- Describe the different steps in protein synthesis
- Discuss the role of ribosomes in protein synthesis

The synthesis of proteins consumes more of a cell's energy than any other metabolic process. In turn, proteins account for more mass than any other component of living organisms (with the exception of water), and proteins perform virtually every function of a cell. The process of translation, or protein synthesis, involves the decoding of an mRNA message into a polypeptide product. Amino acids are covalently strung together by interlinking peptide bonds in lengths ranging from approximately 50 amino acid residues to more than 1,000. Each individual amino acid has an amino group (NH_2) and a carboxyl (COOH) group. Polypeptides are formed when the amino group of one amino acid forms an amide (i.e., peptide) bond with the carboxyl group of another amino acid (Figure 1). This reaction is catalyzed by ribosomes and generates one water molecule.

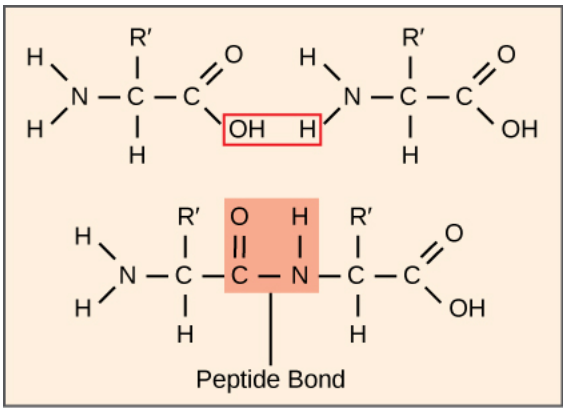


Figure 1 A peptide bond links the carboxyl end of one amino acid with the amino end of another, expelling one water molecule. For simplicity in this image, only the functional groups involved in the peptide bond are shown. The R and R' designations refer to the rest of each amino acid structure.

The Protein Synthesis Machinery

In addition to the mRNA template, many molecules and macromolecules contribute to the process of translation. The composition of each component may vary across species; for instance, ribosomes may consist of different numbers of rRNAs and polypeptides depending on the organism. However, the general structures and functions of the protein synthesis machinery are comparable from bacteria to human cells. Translation requires the input of an mRNA template, ribosomes, tRNAs, and various enzymatic factors.

Link to Learning

Click through the steps of this PBS interactive to see protein synthesis in action.

Ribosomes

Even before an mRNA is translated, a cell must invest energy to build each of its ribosomes. In *E. coli*, there are between 10,000 and 70,000 ribosomes present in each cell at any given time. A ribosome is a complex macromolecule composed of structural and catalytic rRNAs, and many distinct polypeptides. In eukaryotes, the nucleolus is completely specialized for the synthesis and assembly of rRNAs.

Ribosomes exist in the cytoplasm in prokaryotes and in the cytoplasm and rough endoplasmic reticulum in eukaryotes. Mitochondria and chloroplasts also have their own ribosomes in the matrix and stroma, which look more similar to prokaryotic ribosomes (and have similar drug sensitivities) than the ribosomes just outside their outer membranes in the cytoplasm. Ribosomes dissociate into large and small subunits when they are not synthesizing proteins and reassociate during the initiation of translation. In *E. coli*, the small subunit is described as 30S, and the large subunit is 50S, for a total of 70S (recall that Svedberg units are not additive). Mammalian ribosomes have a small 40S subunit and a large 60S subunit, for a total of 80S. The small subunit is responsible for binding the mRNA template, whereas the large subunit sequentially binds tRNAs. Each mRNA molecule is simultaneously translated by many ribosomes, all synthesizing protein in the same direction: reading the mRNA from 5′ to 3′ and synthesizing the polypeptide from the N terminus to the C terminus. The complete mRNA/poly-ribosome structure is called a polysome.

tRNAs

The tRNAs are structural RNA molecules that were transcribed from genes by RNA polymerase III. Depending on the species, 40 to 60 types of tRNAs exist in the cytoplasm. Serving as adaptors, specific tRNAs bind to sequences on the mRNA template and add the corresponding amino acid to the polypeptide chain. Therefore, tRNAs are the molecules that actually "translate" the language of RNA into the language of proteins.

Of the 64 possible mRNA codons—or triplet combinations of A, U, G, and C—three specify the termination of protein synthesis and 61 specify the addition of amino acids to the polypeptide chain. Of these 61, one codon (AUG) also encodes the initiation of translation. Each tRNA anticodon can base pair with one of the mRNA codons and add an amino acid or terminate translation, according to the genetic code. For instance, if the sequence CUA occurred on an mRNA template in the proper reading frame, it would bind a tRNA expressing the complementary sequence, GAU, which would be linked to the amino acid leucine.

As the adaptor molecules of translation, it is surprising that tRNAs can fit so much specificity into such a small package. Consider that tRNAs need to interact with three factors: 1) they must be recognized by the correct aminoacyl synthetase (see below); 2) they must be recognized by ribosomes; and 3) they must bind to the correct sequence in mRNA.

Aminoacyl tRNA Synthetases

The process of pre-tRNA synthesis by RNA polymerase III only creates the RNA portion of the adaptor molecule. The corresponding amino acid must be added later, once the tRNA is processed and exported to the cytoplasm. Through the process of tRNA "charging," each tRNA molecule is linked to its correct amino acid by a group of enzymes called aminoacyl tRNA synthetases. At least one type of aminoacyl tRNA synthetase exists for each of the 20 amino acids; the exact number of aminoacyl tRNA synthetases varies by species. These enzymes first bind and hydrolyze ATP to catalyze a high-energy bond between an amino acid and adenosine monophosphate (AMP); a pyrophosphate molecule is expelled in this reaction. The activated amino acid is then transferred to the tRNA, and AMP is released.

The Mechanism of Protein Synthesis

As with mRNA synthesis, protein synthesis can be divided into three phases: initiation, elongation, and termination. The process of translation is similar in prokaryotes and eukaryotes. Here we'll explore how translation occurs in *E. coli*, a representative prokaryote, and specify any differences between prokaryotic and eukaryotic translation.

Initiation of Translation

Protein synthesis begins with the formation of an initiation complex. In *E. coli*, this complex involves the small 30S ribosome, the mRNA template, three initiation factors (IFs; IF-1, IF-2, and IF-3), and a special initiator tRNA, called tRNAMetf. The initiator tRNA interacts with the start codon AUG (or rarely, GUG), links to a formylated methionine called fMet, and can also bind IF-2. Formylated methionine is inserted by fMet–tRNAMetf at the beginning of every polypeptide chain synthesized by *E. coli*, but it is usually clipped off after translation is complete. When an in-frame AUG is encountered during translation elongation, a non-formylated methionine is inserted by a regular Met-tRNA^{Met}.

In *E. coli* mRNA, a sequence upstream of the first AUG codon, called the Shine-Dalgarno sequence (AGGAGG), interacts with the rRNA molecules that compose the ribosome. This interaction anchors the 30S ribosomal subunit at the correct location on the mRNA template. Guanosine triphosphate (GTP), which is a purine nucleotide triphosphate, acts as an energy source during translation—both at the start of elongation and during the ribosome's translocation.

In eukaryotes, a similar initiation complex forms, comprising mRNA, the 40S small ribosomal subunit, IFs, and nucleoside triphosphates (GTP and ATP). The charged initiator tRNA, called Met-tRNA$_i$, does not bind fMet in eukaryotes, but is distinct from other Met-tRNAs in that it can bind IFs.

Instead of depositing at the Shine-Dalgarno sequence, the eukaryotic initiation complex recognizes the 7-methylguanosine cap at the 5′ end of the mRNA. A cap-binding protein (CBP) and several other IFs assist the movement of the ribosome to the 5′ cap. Once at the cap, the initiation complex tracks along the mRNA in the 5′ to 3′ direction, searching for the AUG start codon. Many eukaryotic mRNAs are translated from the first AUG, but this is not always the case. According to Kozak's rules, the nucleotides around the AUG indicate whether it is the correct start codon. Kozak's rules state that the following consensus sequence must appear around the AUG of

vertebrate genes: 5′-gccRccAUGG-3′. The R (for purine) indicates a site that can be either A or G, but cannot be C or U. Essentially, the closer the sequence is to this consensus, the higher the efficiency of translation.

Once the appropriate AUG is identified, the other proteins and CBP dissociate, and the 60S subunit binds to the complex of Met-tRNA$_i$, mRNA, and the 40S subunit. This step completes the initiation of translation in eukaryotes.

Translation, Elongation, and Termination

In prokaryotes and eukaryotes, the basics of elongation are the same, so we will review elongation from the perspective of *E. coli*. The 50S ribosomal subunit of *E. coli* consists of three compartments: the A (aminoacyl) site binds incoming charged aminoacyl tRNAs. The P (peptidyl) site binds charged tRNAs carrying amino acids that have formed peptide bonds with the growing polypeptide chain but have not yet dissociated from their corresponding tRNA. The E (exit) site releases dissociated tRNAs so that they can be recharged with free amino acids. There is one exception to this assembly line of tRNAs: in *E. coli*, fMet–tRNAMetf is capable of entering the P site directly without first entering the A site. Similarly, the eukaryotic Met-tRNA$_i$, with help from other proteins of the initiation complex, binds directly to the P site. In both cases, this creates an initiation complex with a free A site ready to accept the tRNA corresponding to the first codon after the AUG.

During translation elongation, the mRNA template provides specificity. As the ribosome moves along the mRNA, each mRNA codon comes into register, and specific binding with the corresponding charged tRNA anticodon is ensured. If mRNA were not present in the elongation complex, the ribosome would bind tRNAs nonspecifically.

Elongation proceeds with charged tRNAs entering the A site and then shifting to the P site followed by the E site with each single-codon “step” of the ribosome. Ribosomal steps are induced by conformational changes that advance the ribosome by three bases in the 3′ direction. The energy for each step of the ribosome is donated by an elongation factor that hydrolyzes GTP. Peptide bonds form between the amino group of the amino acid attached to the A-site tRNA and the carboxyl group of the amino acid attached to the P-site tRNA. The formation of each peptide bond is catalyzed by peptidyl transferase, an RNA-based enzyme that is integrated into the 50S ribosomal subunit. The energy for each peptide bond formation is derived from GTP hydrolysis, which is catalyzed by a separate elongation factor. The amino acid bound to the P-site tRNA is also linked to the growing polypeptide chain. As the ribosome steps across the mRNA, the former P-site tRNA enters the E site, detaches from the amino acid, and is expelled (Figure 2). Amazingly, the *E. coli* translation apparatus takes only 0.05 seconds to add each amino acid, meaning that a 200-amino acid protein can be translated in just 10 seconds.

Art Connection

Translation begins when an initiator tRNA anticodon recognizes a codon on mRNA. The large ribosomal subunit joins the small subunit, and a second tRNA is recruited. As the mRNA moves relative to the ribosome, the polypeptide chain is formed. Entry of a release factor into the A site terminates translation and the components dissociate.

Many antibiotics inhibit bacterial protein synthesis. For example, tetracycline blocks the A site on the bacterial ribosome, and chloramphenicol blocks peptidyl transfer. What specific effect would you expect each of these antibiotics to have on protein synthesis?

Tetracycline would directly affect:

1. tRNA binding to the ribosome
2. ribosome assembly
3. growth of the protein chain

Chloramphenicol would directly affect

1. tRNA binding to the ribosome
2. ribosome assembly
3. growth of the protein chain

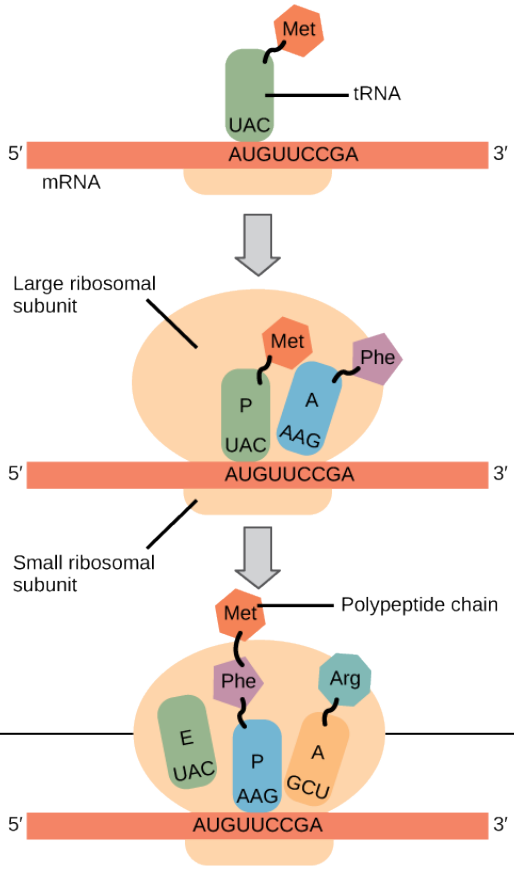


Figure 2. Translation

Termination of translation occurs when a nonsense codon (UAA, UAG, or UGA) is encountered. Upon aligning with the A site, these nonsense codons are recognized by release factors in prokaryotes and eukaryotes that instruct peptidyl transferase to add a water molecule to the carboxyl end of the P-site amino acid. This reaction forces the P-site amino acid to detach from its tRNA, and the newly made protein is released. The small and large ribosomal subunits dissociate from the mRNA and from each other; they are recruited almost immediately into another translation initiation complex. After many ribosomes have completed translation, the mRNA is degraded so the nucleotides can be reused in another transcription reaction.

Protein Folding, Modification, and Targeting

During and after translation, individual amino acids may be chemically modified, signal sequences may be appended, and the new protein “folds” into a distinct three-dimensional structure as a result of intramolecular interactions. A signal sequence is a short tail of amino acids that directs a protein to a specific cellular compartment. These sequences at the amino end or the carboxyl end of the protein can be thought of as the protein’s “train ticket” to its ultimate destination. Other cellular factors recognize each signal sequence and help transport the protein from the cytoplasm to its correct compartment. For instance, a specific sequence at the amino terminus will direct a protein to the mitochondria or chloroplasts (in plants). Once the protein reaches its cellular destination, the signal sequence is usually clipped off.

Many proteins fold spontaneously, but some proteins require helper molecules, called chaperones, to prevent them from aggregating during the complicated process of folding. Even if a protein is properly specified by its corresponding mRNA, it could take on a completely dysfunctional shape if abnormal temperature or pH conditions prevent it from folding correctly.

Section Summary

The players in translation include the mRNA template, ribosomes, tRNAs, and various enzymatic factors. The small ribosomal subunit forms on the mRNA template either at the Shine-Dalgarno sequence (prokaryotes) or the 5′ cap (eukaryotes). Translation begins at the initiating AUG on the mRNA, specifying methionine. The formation of peptide bonds occurs between sequential amino acids specified by the mRNA template according to the genetic code. Charged tRNAs enter the ribosomal A site, and their amino acid bonds with the amino acid at the P site. The entire mRNA is translated in three-nucleotide "steps" of the ribosome. When a nonsense codon is encountered, a release factor binds and dissociates the components and frees the new protein. Folding of the protein occurs during and after translation.

Visit this page in your course online to check your understanding.

Additional Self Check Questions

1. Many antibiotics inhibit bacterial protein synthesis. For example, tetracycline blocks the A site on the bacterial ribosome, and chloramphenicol blocks peptidyl transfer. What specific effect would you expect each of these antibiotics to have on protein synthesis?

    Tetracycline would directly affect:

        A. tRNA binding to the ribosome
        B. ribosome assembly
        C. growth of the protein chain

    Chloramphenicol would directly affect

        A. tRNA binding to the ribosome
        B. ribosome assembly
        C. growth of the protein chain

2. Transcribe and translate the following DNA sequence (nontemplate strand):

    5′-ATGGCCGGTTATTAAGCA-3′

3. Explain how single nucleotide changes can have vastly different effects on protein function.

Answers

1. Tetracycline: a; Chloramphenicol: c.

2. The mRNA would be: 5′-AUGGCCGGUUAUUAAGCA-3′. The protein would be: MAGY. Even though there are six codons, the fifth codon corresponds to a stop, so the sixth codon would not be translated.

3. Nucleotide changes in the third position of codons may not change the amino acid and would have no effect on the protein. Other nucleotide changes that change important amino acids or create or delete start or stop codons would have severe effects on the amino acid sequence of the protein.

Glossary

aminoacyl tRNA synthetase: enzyme that "charges" tRNA molecules by catalyzing a bond between the tRNA and a corresponding amino acid

initiator tRNA: in prokaryotes, called tRNAMetf ; in eukaryotes, called $tRNA_i$; a tRNA that interacts with a start codon, binds directly to the ribosome P site, and links to a special methionine to begin a polypeptide chain

Kozak's rules: determines the correct initiation AUG in a eukaryotic mRNA; the following consensus sequence must appear around the AUG: 5'-GCC(**purine**)CCAUG**G**-3'; the bolded bases are most important

peptidyl transferase: RNA-based enzyme that is integrated into the 50S ribosomal subunit and catalyzes the formation of peptide bonds

polysome: mRNA molecule simultaneously being translated by many ribosomes all going in the same direction

Shine-Dalgarno sequence: (AGGAGG); initiates prokaryotic translation by interacting with rRNA molecules comprising the 30S ribosome

signal sequence: short tail of amino acids that directs a protein to a specific cellular compartment

start codon: AUG (or rarely, GUG) on an mRNA from which translation begins; always specifies methionine

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: Open Stax. **Located at:** http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License**: *CC BY: Attribution*

STUDY GUIDE: PROTEIN SYNTHESIS

Study Questions

Objective: Illustrate the molecular processes described in the central dogma.

Use this page to check your understanding of the content.

Vocabulary

1. Transcription
2. Translation
3. Intron
4. Exon
5. Codon
6. Anticodon

Study Guide Questions

1. What is the central dogma of biology?
2. Compare and contrast mRNA, rRNA, and tRNA.
3. What is DNA polymerase? What is RNA polymerase?
4. How are mRNA transcripts modified before they exit the nucleus?
5. Compare and contrast transcription and DNA replication.
6. Compare and contrast transcription and translation.
7. Compare and contrast mRNA, rRNA, and tRNA.
8. Compare and contrast introns and exons.
9. What is a codon? Compare this to an anticodon.
10. Be able to use an amino acid/mRNA codon chart.
11. Given a DNA template, be able to determine the mRNA base sequence, the tRNA anticodons, and the amino acids coded for (like in activity!)
12. Know the structure of the ribosome.
13. Know the significance of the E site, the P site and the A site in the ribosome.

14. Know the structure of a tRNA molecule.

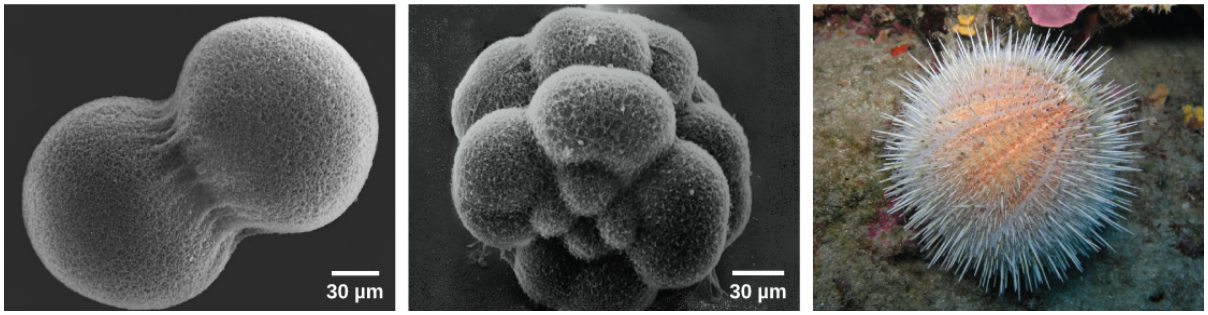
Licensing & Attributions

CC licensed content, Shared previously

- Study Guide: Protein Synthesis. **Authored by**: Wendy Riggs. **Provided by**: College of the Redwoods. **Project**: Kaleidoscope. **License**: *CC BY: Attribution*

THE CELL CYCLE AND MEIOSIS

INTRODUCTION - CELL REPRODUCTION (MITOSIS)



A sea urchin begins life as a single cell that (a) divides to form two cells, visible by scanning electron microscopy. After four rounds of cell division, (b) there are 16 cells, as seen in this SEM image. After many rounds of cell division, the individual develops into a complex, multicellular organism, as seen in this (c) mature sea urchin. (credit a: modification of work by Evelyn Spiegel, Louisa Howard; credit b: modification of work by Evelyn Spiegel, Louisa Howard; credit c: modification of work by Marco Busdraghi; scale-bar data from Matt Russell)

A human, as well as every sexually reproducing organism, begins life as a fertilized egg (embryo) or zygote. Trillions of cell divisions subsequently occur in a controlled manner to produce a complex, multicellular human. In other words, that original single cell is the ancestor of every other cell in the body. Once a being is fully grown, cell reproduction is still necessary to repair or regenerate tissues. For example, new blood and skin cells are constantly being produced. All multicellular organisms use cell division for growth and the maintenance and repair of cells and tissues. Cell division is tightly regulated, and the occasional failure of regulation can have life-threatening consequences. Single-celled organisms use cell division as their method of reproduction.

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: OpenStax College. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License**: *CC BY: Attribution*

CELL DIVISION

Learning Objectives

By the end of this section, you will be able to:

- Describe the structure of prokaryotic and eukaryotic genomes

- Distinguish between chromosomes, genes, and traits
- Describe the mechanisms of chromosome compaction

The continuity of life from one cell to another has its foundation in the reproduction of cells by way of the cell cycle. The cell cycle is an orderly sequence of events that describes the stages of a cell's life from the division of a single parent cell to the production of two new daughter cells. The mechanisms involved in the cell cycle are highly regulated.

Genomic DNA

Before discussing the steps a cell must undertake to replicate, a deeper understanding of the structure and function of a cell's genetic information is necessary. A cell's DNA, packaged as a double-stranded DNA molecule, is called its genome. In prokaryotes, the genome is composed of a single, double-stranded DNA molecule in the form of a loop or circle (Figure 1). The region in the cell containing this genetic material is called a nucleoid. Some prokaryotes also have smaller loops of DNA called plasmids that are not essential for normal growth. Bacteria can exchange these plasmids with other bacteria, sometimes receiving beneficial new genes that the recipient can add to their chromosomal DNA. Antibiotic resistance is one trait that often spreads through a bacterial colony through plasmid exchange.

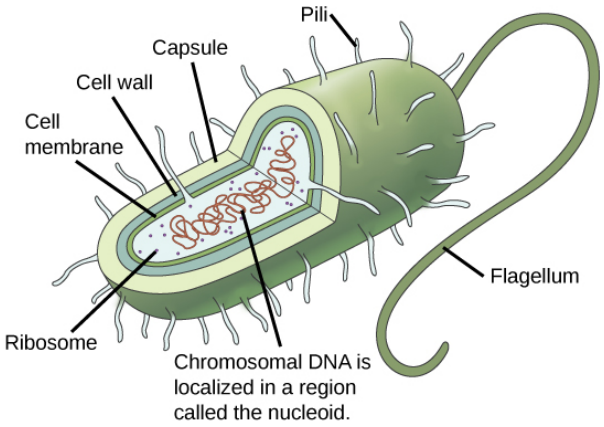


Figure 1 Prokaryotes, including bacteria and archaea, have a single, circular chromosome located in a central region called the nucleoid.

In

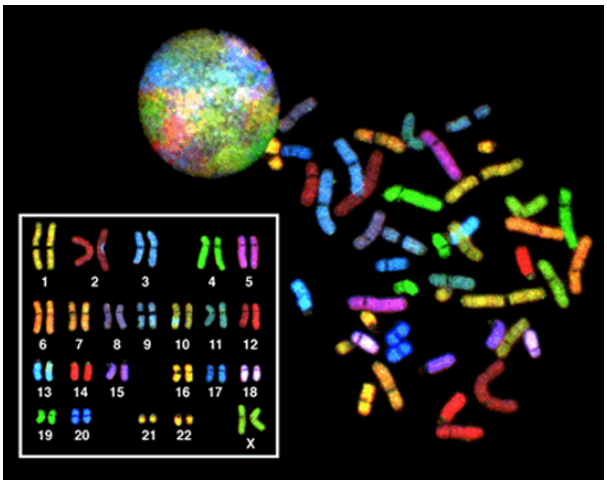


Figure 2 There are 23 pairs of homologous chromosomes in a female human somatic cell. The condensed chromosomes are viewed within the nucleus (top), removed from a cell in mitosis and spread out on a slide (right), and artificially arranged according to length (left); an arrangement like this is called a karyotype. In this image, the chromosomes were exposed to fluorescent stains for differentiation of the different chromosomes. A method of staining called "chromosome painting" employs fluorescent dyes that highlight chromosomes in different colors. (credit: National Human Genome Project/NIH)

eukaryotes, the genome consists of several double-stranded linear DNA molecules (Figure 2). Each species of eukaryotes has a characteristic number of chromosomes in the nuclei of its cells. Human body cells have 46 chromosomes, while human gametes (sperm or eggs) have 23 chromosomes each. A typical body cell, or

somatic cell, contains two matched sets of chromosomes, a configuration known as diploid. The letter *n* is used to represent a single set of chromosomes; therefore, a diploid organism is designated 2*n*. Human cells that contain one set of chromosomes are called gametes, or sex cells; these are eggs and sperm, and are designated *1n*, or haploid.

Matched pairs of chromosomes in a diploid organism are called homologous ("same knowledge") chromosomes. Homologous chromosomes are the same length and have specific nucleotide segments called genes in exactly the same location, or locus. Genes, the functional units of chromosomes, determine specific characteristics by coding for specific proteins. Traits are the variations of those characteristics. For example, hair color is a characteristic with traits that are blonde, brown, or black.

Each copy of a homologous pair of chromosomes originates from a different parent; therefore, the genes themselves are not identical. The variation of individuals within a species is due to the specific combination of the genes inherited from both parents. Even a slightly altered sequence of nucleotides within a gene can result in an alternative trait. For example, there are three possible gene sequences on the human chromosome that code for blood type: sequence A, sequence B, and sequence O. Because all diploid human cells have two copies of the chromosome that determines blood type, the blood type (the trait) is determined by which two versions of the marker gene are inherited. It is possible to have two copies of the same gene sequence on both homologous chromosomes, with one on each (for example, AA, BB, or OO), or two different sequences, such as AB.

Minor variations of traits, such as blood type, eye color, and handedness, contribute to the natural variation found within a species. However, if the entire DNA sequence from any pair of human homologous chromosomes is compared, the difference is less than one percent. The sex chromosomes, X and Y, are the single exception to the rule of homologous chromosome uniformity: Other than a small amount of homology that is necessary to accurately produce gametes, the genes found on the X and Y chromosomes are different.

Eukaryotic Chromosomal Structure and Compaction

If the DNA from all 46 chromosomes in a human cell nucleus was laid out end to end, it would measure approximately two meters; however, its diameter would be only 2 nm. Considering that the size of a typical human cell is about 10 μm (100,000 cells lined up to equal one meter), DNA must be tightly packaged to fit in the cell's nucleus. At the same time, it must also be readily accessible for the genes to be expressed. During some stages of the cell cycle, the long strands of DNA are condensed into compact chromosomes. There are a number of ways that chromosomes are compacted.

In the first level of compaction, short stretches of the DNA double helix wrap around a core of eight histone proteins at regular intervals along the entire length of the chromosome (Figure 3). The DNA-histone complex is called chromatin. The beadlike, histone DNA complex is called a nucleosome, and DNA connecting the nucleosomes is called linker DNA. A DNA molecule in this form is about seven times shorter than the double helix without the histones, and the beads are about 10 nm in diameter, in contrast with the 2-nm diameter of a DNA double helix. The next level of compaction occurs as the nucleosomes and the linker DNA between them are coiled into a 30-nm chromatin fiber. This coiling further shortens the chromosome so that it is now about 50 times shorter than the extended form. In the third level of packing, a variety of fibrous proteins is used to pack the chromatin. These fibrous proteins also ensure that each chromosome in a non-dividing cell occupies a particular area of the nucleus that does not overlap with that of any other chromosome (see the top image in Figure 3).

DNA replicates in the S phase of interphase. After replication, the chromosomes are composed of two linked sister chromatids. When fully compact, the pairs of identically packed chromosomes are bound to each other by cohesin proteins. The connection between the sister chromatids is closest in a region called the centromere. The conjoined sister chromatids, with a diameter of about 1 μm, are visible under a light microscope. The centromeric region is highly condensed and thus will appear as a constricted area.

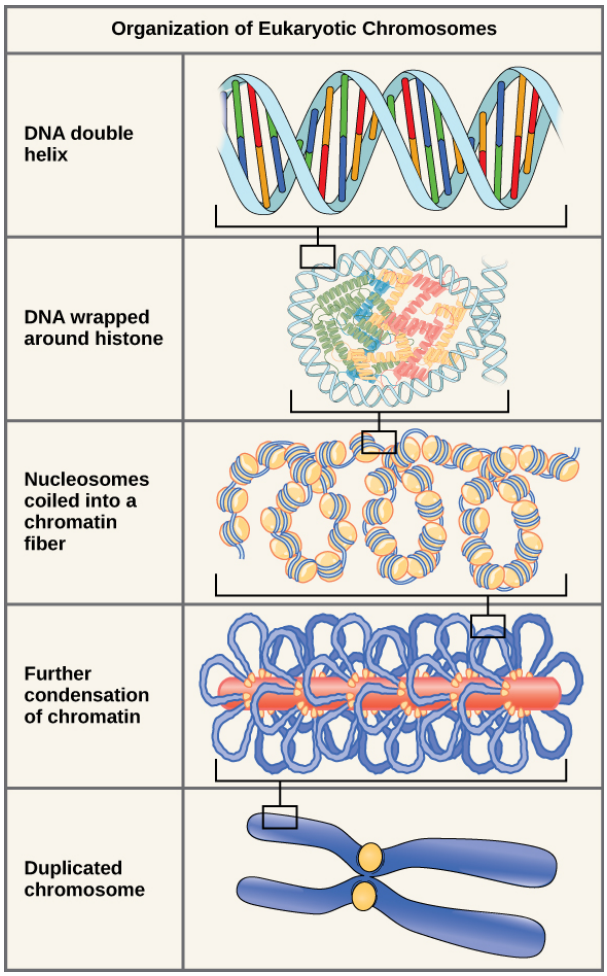


Figure 3 Double-stranded DNA wraps around histone proteins to form nucleosomes that have the appearance of "beads on a string." The nucleosomes are coiled into a 30-nm chromatin fiber. When a cell undergoes mitosis, the chromosomes condense even further.

Link to Learning

This animation illustrates the different levels of chromosome packing.

Section Summary

Prokaryotes have a single circular chromosome composed of double-stranded DNA, whereas eukaryotes have multiple, linear chromosomes composed of chromatin surrounded by a nuclear membrane. The 46 chromosomes of human somatic cells are composed of 22 pairs of autosomes (matched pairs) and a pair of sex chromosomes, which may or may not be matched. This is the 2*n* or diploid state. Human gametes have 23 chromosomes or one complete set of chromosomes; a set of chromosomes is complete with either one of the sex chromosomes. This is the *n* or haploid state. Genes are segments of DNA that code for a specific protein. An organism's traits are determined by the genes inherited from each parent. Duplicated chromosomes are composed of two sister chromatids. Chromosomes are compacted using a variety of mechanisms during certain stages of the cell cycle. Several classes of protein are involved in the organization and packing of the chromosomal DNA into a highly

condensed structure. The condensing complex compacts chromosomes, and the resulting condensed structure is necessary for chromosomal segregation during mitosis.

Visit this page in your course online to check your understanding.

Additional Self Check Questions

1. Compare and contrast a human somatic cell to a human gamete.

2. What is the relationship between a genome, chromosomes, and genes?

3. Eukaryotic chromosomes are thousands of times longer than a typical cell. Explain how chromosomes can fit inside a eukaryotic nucleus.

Answers

1. Human somatic cells have 46 chromosomes: 22 pairs and 2 sex chromosomes that may or may not form a pair. This is the 2*n* or diploid condition. Human gametes have 23 chromosomes, one each of 23 unique chromosomes, one of which is a sex chromosome. This is the *n* or haploid condition.
2. The genome consists of the sum total of an organism's chromosomes. Each chromosome contains hundreds and sometimes thousands of genes, segments of DNA that code for a polypeptide or RNA, and a large amount of DNA with no known function.
3. The DNA double helix is wrapped around histone proteins to form structures called nucleosomes. Nucleosomes and the linker DNA in between them are coiled into a 30-nm fiber. During cell division, chromatin is further condensed by packing proteins.

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: Open Stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License**: *CC BY: Attribution*

THE CELL CYCLE

Learning Objectives

By the end of this section, you will be able to:

- Describe the three stages of interphase
- Discuss the behavior of chromosomes during karyokinesis
- Explain how the cytoplasmic content is divided during cytokinesis
- Define the quiescent G_0 phase

The cell cycle is an ordered series of events involving cell growth and cell division that produces two new daughter cells. Cells on the path to cell division proceed through a series of precisely timed and carefully regulated stages of growth, DNA replication, and division that produces two identical (clone) cells. The cell cycle has two major phases: interphase and the mitotic phase (Figure 1). During interphase, the cell grows and DNA is replicated. During the mitotic phase, the replicated DNA and cytoplasmic contents are separated, and the cell divides.

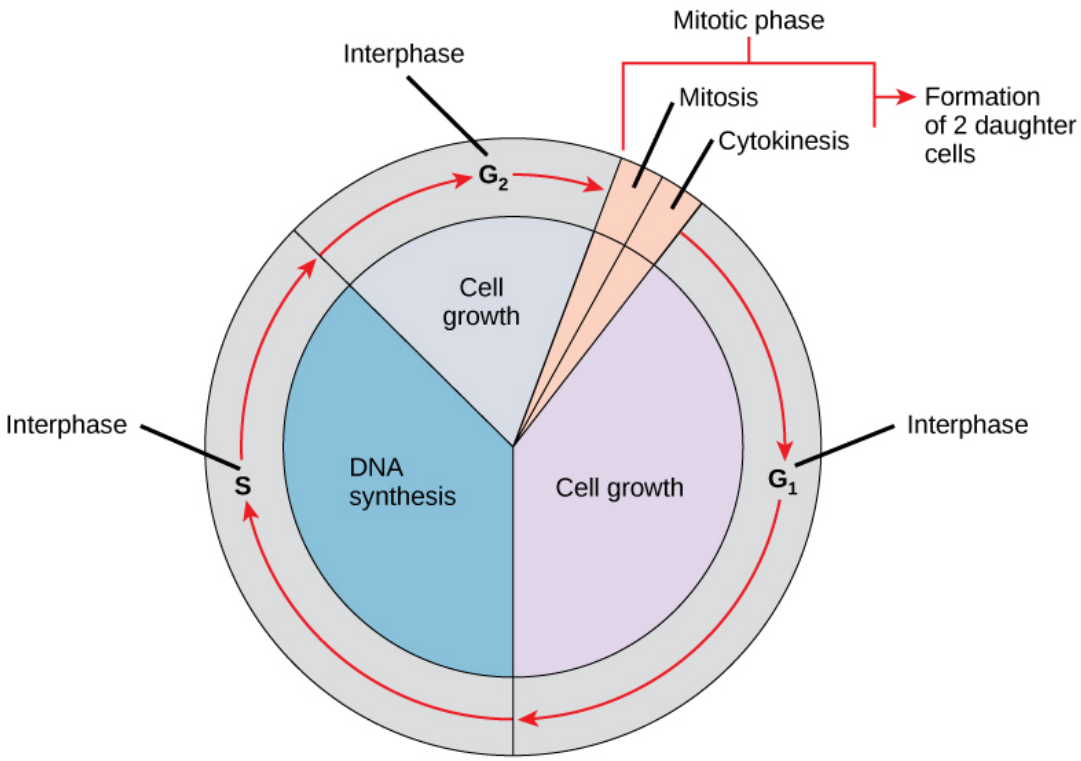


Figure 1 The cell cycle consists of interphase and the mitotic phase. During interphase, the cell grows and the nuclear DNA is duplicated. Interphase is followed by the mitotic phase. During the mitotic phase, the duplicated chromosomes are segregated and distributed into daughter nuclei. The cytoplasm is usually divided as well, resulting in two daughter cells.

Interphase

During interphase, the cell undergoes normal growth processes while also preparing for cell division. In order for a cell to move from interphase into the mitotic phase, many internal and external conditions must be met. The three stages of interphase are called G_1, S, and G_2.

G_1 Phase (First Gap)

The first stage of interphase is called the G_1 phase (first gap) because, from a microscopic aspect, little change is visible. However, during the G_1 stage, the cell is quite active at the biochemical level. The cell is accumulating the building blocks of chromosomal DNA and the associated proteins as well as accumulating sufficient energy reserves to complete the task of replicating each chromosome in the nucleus.

S Phase (Synthesis of DNA)

Throughout interphase, nuclear DNA remains in a semi-condensed chromatin configuration. In the S phase, DNA replication can proceed through the mechanisms that result in the formation of identical pairs of DNA molecules—sister chromatids—that are firmly attached to the centromeric region. The centrosome is duplicated during the S phase. The two centrosomes will give rise to the mitotic spindle, the apparatus that orchestrates the movement of chromosomes during mitosis. At the center of each animal cell, the centrosomes of animal cells are associated with a pair of rod-like objects, the centrioles, which are at right angles to each other. Centrioles help organize cell division. Centrioles are not present in the centrosomes of other eukaryotic species, such as plants and most fungi.

G_2 Phase (Second Gap)

In the G_2 phase, the cell replenishes its energy stores and synthesizes proteins necessary for chromosome manipulation. Some cell organelles are duplicated, and the cytoskeleton is dismantled to provide resources for the mitotic phase. There may be additional cell growth during G_2. The final preparations for the mitotic phase must be completed before the cell is able to enter the first stage of mitosis.

The Mitotic Phase

The mitotic phase is a multistep process during which the duplicated chromosomes are aligned, separated, and move into two new, identical daughter cells. The first portion of the mitotic phase is called karyokinesis, or nuclear division. The second portion of the mitotic phase, called cytokinesis, is the physical separation of the cytoplasmic components into the two daughter cells.

Link to Learning

Revisit the stages of mitosis at this site.

Karyokinesis (Mitosis)

Karyokinesis, also known as mitosis, is divided into a series of phases—prophase, prometaphase, metaphase, anaphase, and telophase—that result in the division of the cell nucleus (Figure 2). Karyokinesis is also called mitosis.

Art Connection

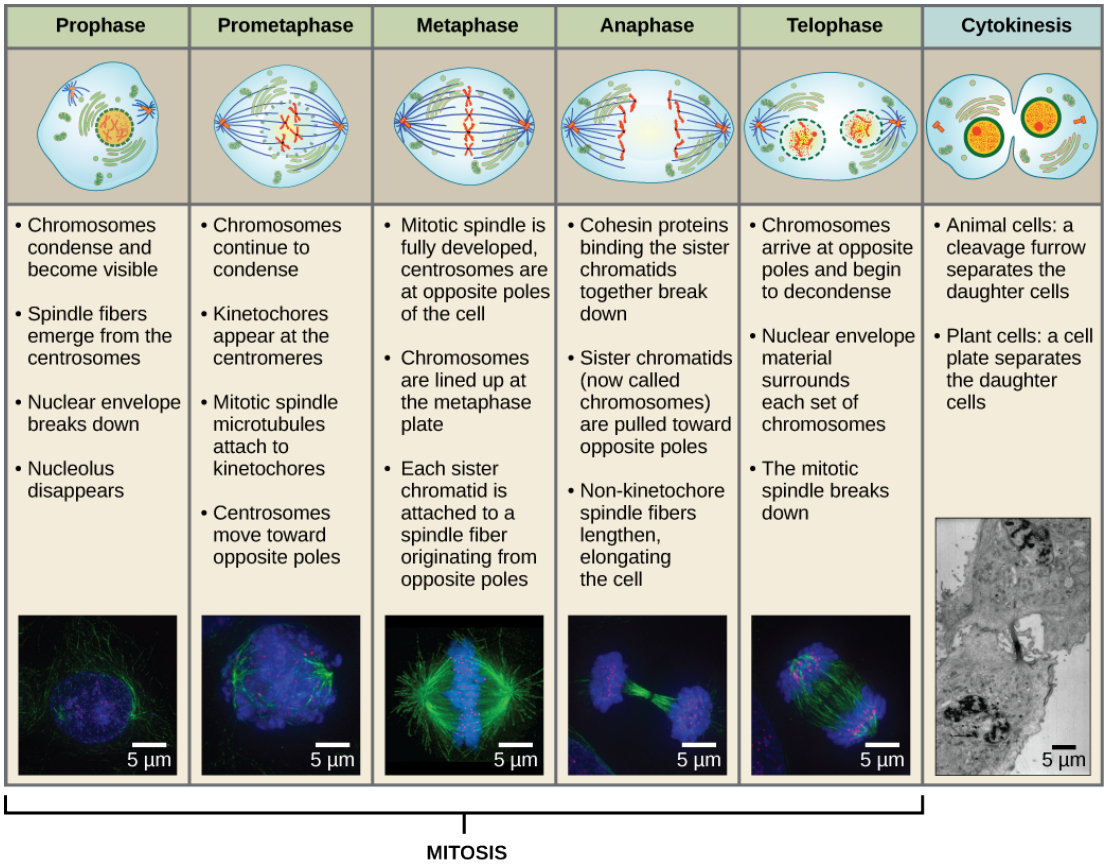


Figure 2. Karyokinesis (or mitosis) is divided into five stages—prophase, prometaphase, metaphase, anaphase, and telophase. The pictures at the bottom were taken by fluorescence microscopy (hence, the black background) of cells artificially stained by fluorescent dyes: blue fluorescence indicates DNA (chromosomes) and green fluorescence indicates microtubules (spindle apparatus). (credit "mitosis drawings": modification of work by Mariana Ruiz Villareal; credit "micrographs": modification of work by Roy van Heesbeen; credit "cytokinesis micrograph": Wadsworth Center/New York State Department of Health; scale-bar data from Matt Russell)

Which of the following is the correct order of events in mitosis?

1. Sister chromatids line up at the metaphase plate. The kinetochore becomes attached to the mitotic spindle. The nucleus reforms and the cell divides. Cohesin proteins break down and the sister chromatids separate.
2. The kinetochore becomes attached to the mitotic spindle. Cohesin proteins break down and the sister chromatids separate. Sister chromatids line up at the metaphase plate. The nucleus reforms and the cell divides.
3. The kinetochore becomes attached to the cohesin proteins. Sister chromatids line up at the metaphase plate. The kinetochore breaks down and the sister chromatids separate. The nucleus reforms and the cell divides.
4. The kinetochore becomes attached to the mitotic spindle. Sister chromatids line up at the metaphase plate. Cohesin proteins break down and the sister chromatids separate. The nucleus reforms and the cell divides.

During prophase, the "first phase," the nuclear envelope starts to dissociate into small vesicles, and the membranous organelles (such as the Golgi complex or Golgi apparatus, and endoplasmic reticulum), fragment and disperse toward the periphery of the cell. The nucleolus disappears (disperses). The centrosomes begin to move to opposite poles of the cell. Microtubules that will form the mitotic spindle extend between the

centrosomes, pushing them farther apart as the microtubule fibers lengthen. The sister chromatids begin to coil more tightly with the aid of condensin proteins and become visible under a light microscope.

During prometaphase, the “first change phase,” many processes that were begun in prophase continue to advance. The remnants of the nuclear envelope fragment. The mitotic spindle continues to develop as more microtubules assemble and stretch across the length of the former nuclear area. Chromosomes become more condensed and discrete. Each sister chromatid develops a protein structure called a kinetochore in the centromeric region (Figure 3). The proteins of the kinetochore attract and bind mitotic spindle microtubules. As the spindle microtubules extend from the centrosomes, some of these microtubules come into contact with and firmly bind to the kinetochores. Once a mitotic fiber attaches to a chromosome, the chromosome will be oriented until the kinetochores of sister chromatids face the opposite poles. Eventually, all the sister chromatids will be attached via their kinetochores to microtubules from opposing poles. Spindle microtubules that do not engage the chromosomes are called polar microtubules. These microtubules overlap each other midway between the two poles and contribute to cell elongation. Astral microtubules are located near the poles, aid in spindle orientation, and are required for the regulation of mitosis.

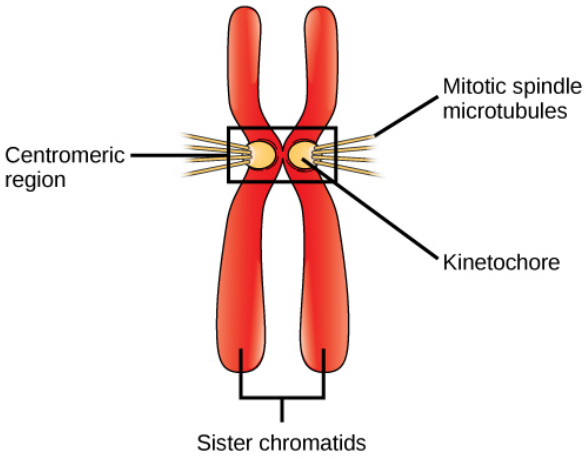


Figure 3. During prometaphase, mitotic spindle microtubules from opposite poles attach to each sister chromatid at the kinetochore. In anaphase, the connection between the sister chromatids breaks down, and the microtubules pull the chromosomes toward opposite poles.

During metaphase, the “change phase,” all the chromosomes are aligned in a plane called the metaphase plate, or the equatorial plane, midway between the two poles of the cell. The sister chromatids are still tightly attached to each other by cohesin proteins. At this time, the chromosomes are maximally condensed.

During anaphase, the “upward phase,” the cohesin proteins degrade, and the sister chromatids separate at the centromere. Each chromatid, now called a chromosome, is pulled rapidly toward the centrosome to which its microtubule is attached. The cell becomes visibly elongated (oval shaped) as the polar microtubules slide against each other at the metaphase plate where they overlap.

During telophase, the “distance phase,” the chromosomes reach the opposite poles and begin to decondense (unravel), relaxing into a chromatin configuration. The mitotic spindles are depolymerized into tubulin monomers that will be used to assemble cytoskeletal components for each daughter cell. Nuclear envelopes form around the chromosomes, and nucleosomes appear within the nuclear area.

Cytokinesis

Cytokinesis, or “cell motion,” is the second main stage of the mitotic phase during which cell division is completed via the physical separation of the cytoplasmic components into two daughter cells. Division is not complete until the cell components have been apportioned and completely separated into the two daughter cells. Although the stages of mitosis are similar for most eukaryotes, the process of cytokinesis is quite different for eukaryotes that have cell walls, such as plant cells.

In cells such as animal cells that lack cell walls, cytokinesis follows the onset of anaphase. A contractile ring composed of actin filaments forms just inside the plasma membrane at the former metaphase plate. The actin filaments pull the equator of the cell inward, forming a fissure. This fissure, or "crack," is called the cleavage furrow. The furrow deepens as the actin ring contracts, and eventually the membrane is cleaved in two (Figure 4).

In plant cells, a new cell wall must form between the daughter cells. During interphase, the Golgi apparatus accumulates enzymes, structural proteins, and glucose molecules prior to breaking into vesicles and dispersing throughout the dividing cell. During telophase, these Golgi vesicles are transported on microtubules to form a phragmoplast (a vesicular structure) at the metaphase plate. There, the vesicles fuse and coalesce from the center toward the cell walls; this structure is called a cell plate. As more vesicles fuse, the cell plate enlarges until it merges with the cell walls at the periphery of the cell. Enzymes use the glucose that has accumulated between the membrane layers to build a new cell wall. The Golgi membranes become parts of the plasma membrane on either side of the new cell wall (Figure 4).

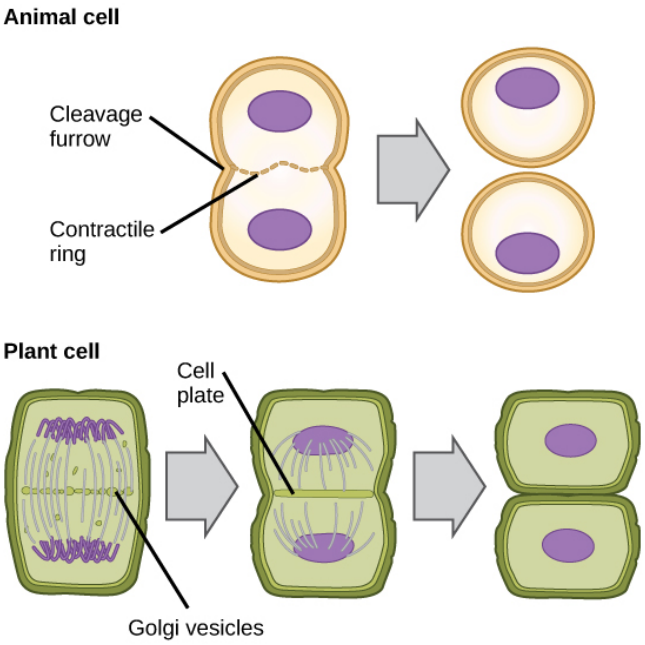


Figure 4 During cytokinesis in animal cells, a ring of actin filaments forms at the metaphase plate. The ring contracts, forming a cleavage furrow, which divides the cell in two. In plant cells, Golgi vesicles coalesce at the former metaphase plate, forming a phragmoplast. A cell plate formed by the fusion of the vesicles of the phragmoplast grows from the center toward the cell walls, and the membranes of the vesicles fuse to form a plasma membrane that divides the cell in two.

G_0 Phase

Not all cells adhere to the classic cell cycle pattern in which a newly formed daughter cell immediately enters the preparatory phases of interphase, closely followed by the mitotic phase. Cells in G_0 phase are not actively preparing to divide. The cell is in a quiescent (inactive) stage that occurs when cells exit the cell cycle. Some cells enter G_0 temporarily until an external signal triggers the onset of G_1. Other cells that never or rarely divide, such as mature cardiac muscle and nerve cells, remain in G_0 permanently.

Scientific Method Connection

Determine the Time Spent in Cell Cycle Stages

Problem: How long does a cell spend in interphase compared to each stage of mitosis?

Background: A prepared microscope slide of blastula cross-sections will show cells arrested in various stages of the cell cycle. It is not visually possible to separate the stages of interphase from each other, but the mitotic stages are readily identifiable. If 100 cells are examined, the number of cells in each identifiable cell cycle stage will give an estimate of the time it takes for the cell to complete that stage.

Problem Statement: Given the events included in all of interphase and those that take place in each stage of mitosis, estimate the length of each stage based on a 24-hour cell cycle. Before proceeding, state your hypothesis.

Test your hypothesis: Test your hypothesis by doing the following:

1. Place a fixed and stained microscope slide of whitefish blastula cross-sections under the scanning objective of a light microscope.

2. Locate and focus on one of the sections using the scanning objective of your microscope. Notice that the section is a circle composed of dozens of closely packed individual cells.
3. Switch to the low-power objective and refocus. With this objective, individual cells are visible.

4. Switch to the high-power objective and slowly move the slide left to right, and up and down to view all the cells in the section (Figure 5). As you scan, you will notice that most of the cells are not undergoing mitosis but are in the interphase period of the cell cycle.

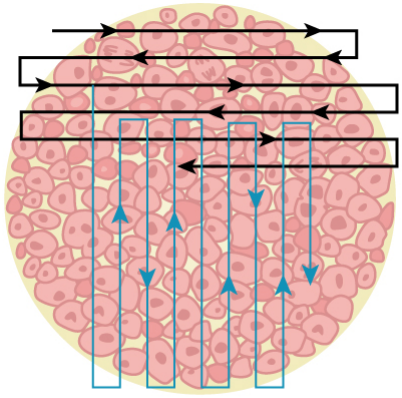


Figure 5: Slowly scan whitefish blastula cells with the high-power objective as illustrated in image (a) to identify their mitotic stage. (b) A microscopic image of the scanned cells is shown. (credit "micrograph": modification of work by Linda Flora; scale-bar data from Matt Russell)

5. Practice identifying the various stages of the cell cycle, using the drawings of the stages as a guide (Figure 2).
6. Once you are confident about your identification, begin to record the stage of each cell you encounter as you scan left to right, and top to bottom across the blastula section.
7. Keep a tally of your observations and stop when you reach 100 cells identified.
8. The larger the sample size (total number of cells counted), the more accurate the results. If possible, gather and record group data prior to calculating percentages and making estimates.

Record your observations: Make a table similar to Table 1 in which you record your observations.

Results of Cell Stage Identification			
Phase or Stage	**Individual Totals**	**Group Totals**	**Percent**
Interphase			
Prophase			
Metaphase			
Anaphase			

Results of Cell Stage Identification			
Phase or Stage	Individual Totals	Group Totals	Percent
Telophase			
Cytokinesis			
Totals	100	100	100 percent

Table 1

Analyze your data/report your results: To find the length of time whitefish blastula cells spend in each stage, multiply the percent (recorded as a decimal) by 24 hours. Make a table similar to Table to illustrate your data.

Estimate of Cell Stage Length		
Phase or Stage	Percent (as Decimal)	Time in Hours
Interphase		
Prophase		
Metaphase		
Anaphase		
Telophase		
Cytokinesis		

Draw a conclusion: Did your results support your estimated times? Were any of the outcomes unexpected? If so, discuss which events in that stage might contribute to the calculated time.

Section Summary

The cell cycle is an orderly sequence of events. Cells on the path to cell division proceed through a series of precisely timed and carefully regulated stages. In eukaryotes, the cell cycle consists of a long preparatory period, called interphase. Interphase is divided into G_1, S, and G_2 phases. The mitotic phase begins with karyokinesis (mitosis), which consists of five stages: prophase, prometaphase, metaphase, anaphase, and telophase. The final stage of the mitotic phase is cytokinesis, during which the cytoplasmic components of the daughter cells are separated either by an actin ring (animal cells) or by cell plate formation (plant cells).

Visit this page in your course online to check your understanding.

Additional Self Check Questions

1. Which of the following is the correct order of events in mitosis?

A. Sister chromatids line up at the metaphase plate. The kinetochore becomes attached to the mitotic spindle. The nucleus reforms and the cell divides. Cohesin proteins break down and the sister chromatids separate.
B. The kinetochore becomes attached to the mitotic spindle. Cohesin proteins break down and the sister chromatids separate. Sister chromatids line up at the metaphase plate. The nucleus reforms and the cell divides.
C. The kinetochore becomes attached to the cohesin proteins. Sister chromatids line up at the metaphase plate. The kinetochore breaks down and the sister chromatids separate. The nucleus reforms and the cell divides.
D. The kinetochore becomes attached to the mitotic spindle. Sister chromatids line up at the metaphase plate. Cohesin proteins break down and the sister chromatids separate. The nucleus reforms and the cell divides.

2. Briefly describe the events that occur in each phase of interphase.

3. Chemotherapy drugs such as vincristine and colchicine disrupt mitosis by binding to tubulin (the subunit of microtubules) and interfering with microtubule assembly and disassembly. Exactly what mitotic structure is targeted by these drugs and what effect would that have on cell division?

4. Describe the similarities and differences between the cytokinesis mechanisms found in animal cells versus those in plant cells.

5. List some reasons why a cell that has just completed cytokinesis might enter the G_0 phase instead of the G_1 phase.

6. What cell cycle events will be affected in a cell that produces mutated (non-functional) cohesin protein?

Answers

1. D. The kinetochore becomes attached to the mitotic spindle. Sister chromatids line up at the metaphase plate. Cohesin proteins break down and the sister chromatids separate. The nucleus reforms and the cell divides.

2.During G_1, the cell increases in size, the genomic DNA is assessed for damage, and the cell stockpiles energy reserves and the components to synthesize DNA. During the S phase, the chromosomes, the centrosomes, and the centrioles (animal cells) duplicate. During the G_2 phase, the cell recovers from the S phase, continues to grow, duplicates some organelles, and dismantles other organelles.
3. The mitotic spindle is formed of microtubules. Microtubules are polymers of the protein tubulin; therefore, it is the mitotic spindle that is disrupted by these drugs. Without a functional mitotic spindle, the chromosomes will not be sorted or separated during mitosis. The cell will arrest in mitosis and die.4. There are very few similarities between animal cell and plant cell cytokinesis. In animal cells, a ring of actin fibers is formed around the periphery of the cell at the former metaphase plate (cleavage furrow). The actin ring contracts inward, pulling the plasma membrane toward the center of the cell until the cell is pinched in two. In plant cells, a new cell wall must be formed between the daughter cells. Due to the rigid cell walls of the parent cell, contraction of the middle of the cell is not possible. Instead, a phragmoplast first forms. Subsequently, a cell plate is formed in the center of the cell at the former metaphase plate. The cell plate is formed from Golgi vesicles that contain enzymes, proteins, and glucose. The vesicles fuse and the enzymes build a new cell wall from the proteins and glucose. The cell plate grows toward and eventually fuses with the cell wall of the parent cell.5. Many cells temporarily enter G_0 until they reach maturity. Some cells are only triggered to enter G_1 when the organism needs to increase that particular cell type. Some cells only reproduce following an injury to the tissue. Some cells never divide once they reach maturity.6. If cohesion is not functional, chromosomes are not packaged after DNA replication in the S phase of interphase. It is likely that the proteins of the centromeric

region, such as the kinetochore, would not form. Even if the mitotic spindle fibers could attach to the chromatids without packing, the chromosomes would not be sorted or separated during mitosis.

Glossary

anaphase: stage of mitosis during which sister chromatids are separated from each other

cell cycle: ordered series of events involving cell growth and cell division that produces two new daughter cells

cell plate: structure formed during plant cell cytokinesis by Golgi vesicles, forming a temporary structure (phragmoplast) and fusing at the metaphase plate; ultimately leads to the formation of cell walls that separate the two daughter cells

centriole: rod-like structure constructed of microtubules at the center of each animal cell centrosome

cleavage furrow: constriction formed by an actin ring during cytokinesis in animal cells that leads to cytoplasmic division

condensin: proteins that help sister chromatids coil during prophase

cytokinesis: division of the cytoplasm following mitosis that forms two daughter cells.

G_0 phase: distinct from the G_1 phase of interphase; a cell in G_0 is not preparing to divide

G_1 phase: (also, first gap) first phase of interphase centered on cell growth during mitosis

G_2 phase: (also, second gap) third phase of interphase during which the cell undergoes final preparations for mitosis

interphase: period of the cell cycle leading up to mitosis; includes G_1, S, and G_2 phases (the interim period between two consecutive cell divisions

karyokinesis: mitotic nuclear division

kinetochore: protein structure associated with the centromere of each sister chromatid that attracts and binds spindle microtubules during prometaphase

metaphase plate: equatorial plane midway between the two poles of a cell where the chromosomes align during metaphase

metaphase: stage of mitosis during which chromosomes are aligned at the metaphase plate

mitosis: (also, karyokinesis) period of the cell cycle during which the duplicated chromosomes are separated into identical nuclei; includes prophase, prometaphase, metaphase, anaphase, and telophase

mitotic phase: period of the cell cycle during which duplicated chromosomes are distributed into two nuclei and cytoplasmic contents are divided; includes karyokinesis (mitosis) and cytokinesis

mitotic spindle: apparatus composed of microtubules that orchestrates the movement of chromosomes during mitosis

prometaphase: stage of mitosis during which the nuclear membrane breaks down and mitotic spindle fibers attach to kinetochores

prophase: stage of mitosis during which chromosomes condense and the mitotic spindle begins to form

quiescent: refers to a cell that is performing normal cell functions and has not initiated preparations for cell division

S phase: second, or synthesis, stage of interphase during which DNA replication occurs

telophase: stage of mitosis during which chromosomes arrive at opposite poles, decondense, and are surrounded by a new nuclear envelope

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: Open Stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biologyt. **License**: *CC BY: Attribution*

CONTROL OF THE CELL CYCLE

Learning Objectives

By the end of this section, you will be able to:

- Understand how the cell cycle is controlled by mechanisms both internal and external to the cell
- Explain how the three internal control checkpoints occur at the end of G_1, at the G_2/M transition, and during metaphase
- Describe the molecules that control the cell cycle through positive and negative regulation

The length of the cell cycle is highly variable, even within the cells of a single organism. In humans, the frequency of cell turnover ranges from a few hours in early embryonic development, to an average of two to five days for epithelial cells, and to an entire human lifetime spent in G_0 by specialized cells, such as cortical neurons or cardiac muscle cells. There is also variation in the time that a cell spends in each phase of the cell cycle. When fast-dividing mammalian cells are grown in culture (outside the body under optimal growing conditions), the length of the cycle is about 24 hours. In rapidly dividing human cells with a 24-hour cell cycle, the G_1 phase lasts approximately nine hours, the S phase lasts 10 hours, the G_2 phase lasts about four and one-half hours, and the M phase lasts approximately one-half hour. In early embryos of fruit flies, the cell cycle is completed in about eight minutes. The timing of events in the cell cycle is controlled by mechanisms that are both internal and external to the cell.

Regulation of the Cell Cycle by External Events

Both the initiation and inhibition of cell division are triggered by events external to the cell when it is about to begin the replication process. An event may be as simple as the death of a nearby cell or as sweeping as the release of growth-promoting hormones, such as human growth hormone (HGH). A lack of HGH can inhibit cell division, resulting in dwarfism, whereas too much HGH can result in gigantism. Crowding of cells can also inhibit cell division. Another factor that can initiate cell division is the size of the cell; as a cell grows, it becomes inefficient due to its decreasing surface-to-volume ratio. The solution to this problem is to divide.

Whatever the source of the message, the cell receives the signal, and a series of events within the cell allows it to proceed into interphase. Moving forward from this initiation point, every parameter required during each cell cycle phase must be met or the cycle cannot progress.

Regulation at Internal Checkpoints

It is essential that the daughter cells produced be exact duplicates of the parent cell. Mistakes in the duplication or distribution of the chromosomes lead to mutations that may be passed forward to every new cell produced from an abnormal cell. To prevent a compromised cell from continuing to divide, there are internal control mechanisms that operate at three main cell cycle checkpoints. A checkpoint is one of several points in the eukaryotic cell cycle at which the progression of a cell to the next stage in the cycle can be halted until conditions are favorable. These checkpoints occur near the end of G_1, at the G_2/M transition, and during metaphase (Figure 1).

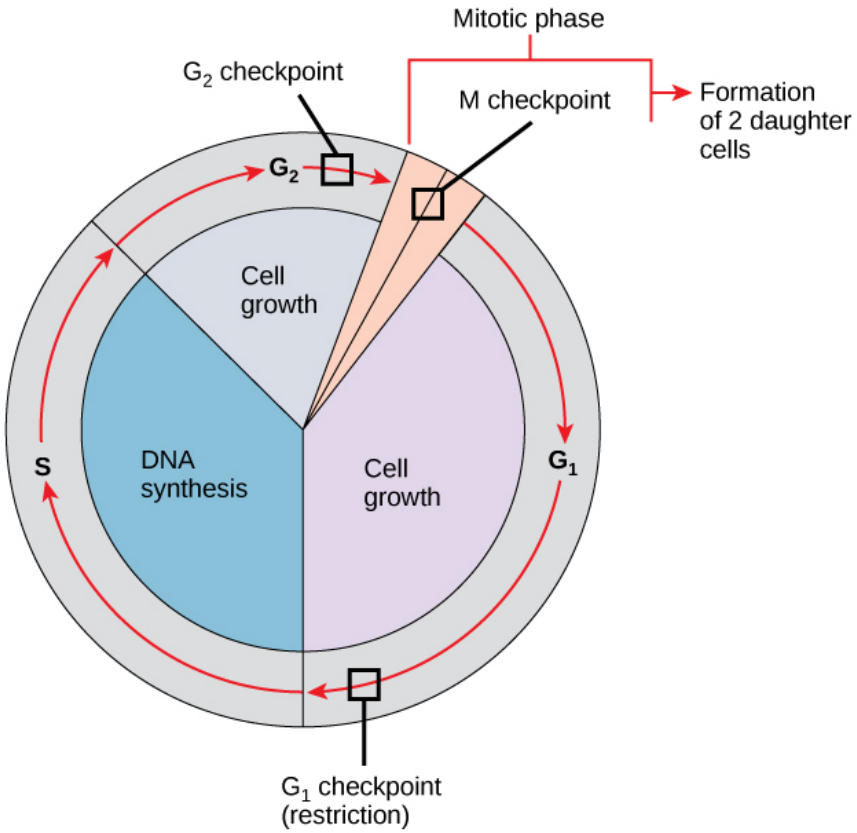


Figure 1 The cell cycle is controlled at three checkpoints. The integrity of the DNA is assessed at the G1 checkpoint. Proper chromosome duplication is assessed at the G2 checkpoint. Attachment of each kinetochore to a spindle fiber is assessed at the M checkpoint.

The G_1 Checkpoint

The G_1 checkpoint determines whether all conditions are favorable for cell division to proceed. The G_1 checkpoint, also called the restriction point (in yeast), is a point at which the cell irreversibly commits to the cell division process. External influences, such as growth factors, play a large role in carrying the cell past the G_1 checkpoint. In addition to adequate reserves and cell size, there is a check for genomic DNA damage at the G_1 checkpoint. A cell that does not meet all the requirements will not be allowed to progress into the S phase. The cell can halt the cycle and attempt to remedy the problematic condition, or the cell can advance into G_0 and await further signals when conditions improve.

The G_2 Checkpoint

The G_2 checkpoint bars entry into the mitotic phase if certain conditions are not met. As at the G_1 checkpoint, cell size and protein reserves are assessed. However, the most important role of the G_2 checkpoint is to ensure that

all of the chromosomes have been replicated and that the replicated DNA is not damaged. If the checkpoint mechanisms detect problems with the DNA, the cell cycle is halted, and the cell attempts to either complete DNA replication or repair the damaged DNA.

The M Checkpoint

The M checkpoint occurs near the end of the metaphase stage of karyokinesis. The M checkpoint is also known as the spindle checkpoint, because it determines whether all the sister chromatids are correctly attached to the spindle microtubules. Because the separation of the sister chromatids during anaphase is an irreversible step, the cycle will not proceed until the kinetochores of each pair of sister chromatids are firmly anchored to at least two spindle fibers arising from opposite poles of the cell.

Link to Learning

Watch what occurs at the G_1, G_2, and M checkpoints by visiting this website to see an animation of the cell cycle.

Regulator Molecules of the Cell Cycle

In addition to the internally controlled checkpoints, there are two groups of intracellular molecules that regulate the cell cycle. These regulatory molecules either promote progress of the cell to the next phase (positive regulation) or halt the cycle (negative regulation). Regulator molecules may act individually, or they can influence the activity or production of other regulatory proteins. Therefore, the failure of a single regulator may have almost no effect on the cell cycle, especially if more than one mechanism controls the same event. Conversely, the effect of a deficient or non-functioning regulator can be wide-ranging and possibly fatal to the cell if multiple processes are affected.

Positive Regulation of the Cell Cycle

Two groups of proteins, called cyclins and cyclin-dependent kinases (Cdks), are responsible for the progress of the cell through the various checkpoints. The levels of the four cyclin proteins fluctuate throughout the cell cycle in a predictable pattern (Figure 2). Increases in the concentration of cyclin proteins are triggered by both external and internal signals. After the cell moves to the next stage of the cell cycle, the cyclins that were active in the previous stage are degraded.

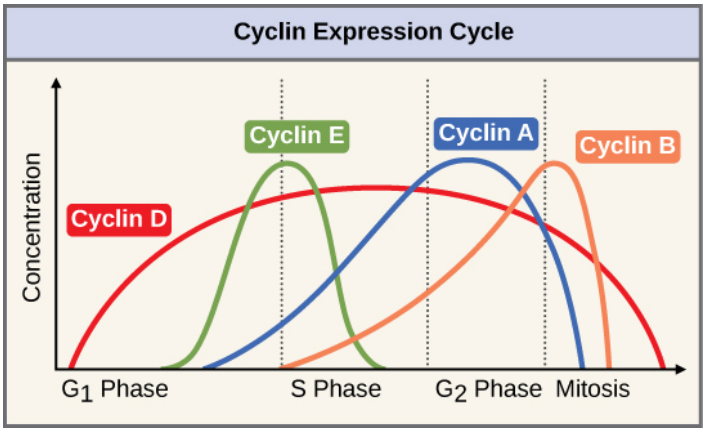


Figure 2. The concentrations of cyclin proteins change throughout the cell cycle. There is a direct correlation between cyclin accumulation and the three major cell cycle checkpoints. Also note the sharp decline of cyclin levels following each checkpoint (the transition between phases of the cell cycle), as cyclin is degraded by cytoplasmic enzymes. (credit: modification of work by "WikiMiMa"/Wikimedia Commons)

Cyclins regulate the cell cycle only when they are tightly bound to Cdks. To be fully active, the Cdk/cyclin complex must also be phosphorylated in specific locations. Like all kinases, Cdks are enzymes (kinases) that phosphorylate other proteins. Phosphorylation activates the protein by changing its shape. The proteins phosphorylated by Cdks are involved in advancing the cell to the next phase. (Figure 3). The levels of Cdk proteins are relatively stable throughout the cell cycle; however, the concentrations of cyclin fluctuate and determine when Cdk/cyclin complexes form. The different cyclins and Cdks bind at specific points in the cell cycle and thus regulate different checkpoints.

Since the cyclic fluctuations of cyclin levels are based on the timing of the cell cycle and not on specific events, regulation of the cell cycle usually occurs by either the Cdk molecules alone or the Cdk/cyclin complexes. Without a specific concentration of fully activated cyclin/Cdk complexes, the cell cycle cannot proceed through the checkpoints.

Although the cyclins are the main regulatory molecules that determine the forward momentum of the cell cycle, there are several other mechanisms that fine-tune the progress of the cycle with negative, rather than positive, effects. These mechanisms essentially block the progression of the cell cycle until problematic conditions are resolved. Molecules that prevent the full activation of Cdks are called Cdk inhibitors. Many of these inhibitor molecules directly or indirectly monitor a particular cell cycle event. The block placed on Cdks by inhibitor molecules will not be removed until the specific event that the inhibitor monitors is completed.

Negative Regulation of the Cell Cycle

The second group of cell cycle regulatory molecules are negative regulators. Negative regulators halt the cell cycle. Remember that in positive regulation, active molecules cause the cycle to progress.

The best understood negative regulatory molecules are retinoblastoma protein (Rb), p53, and p21. Retinoblastoma proteins are a group of tumor-suppressor proteins common in many cells. The 53 and 21 designations refer to the functional molecular masses of the proteins (p) in kilodaltons. Much of what is known about cell cycle regulation comes from research conducted with cells that have lost regulatory control. All three of these regulatory proteins were discovered to be damaged or non-functional in cells that had begun to replicate uncontrollably (became cancerous). In each case, the main cause of the unchecked progress through the cell cycle was a faulty copy of the regulatory protein.

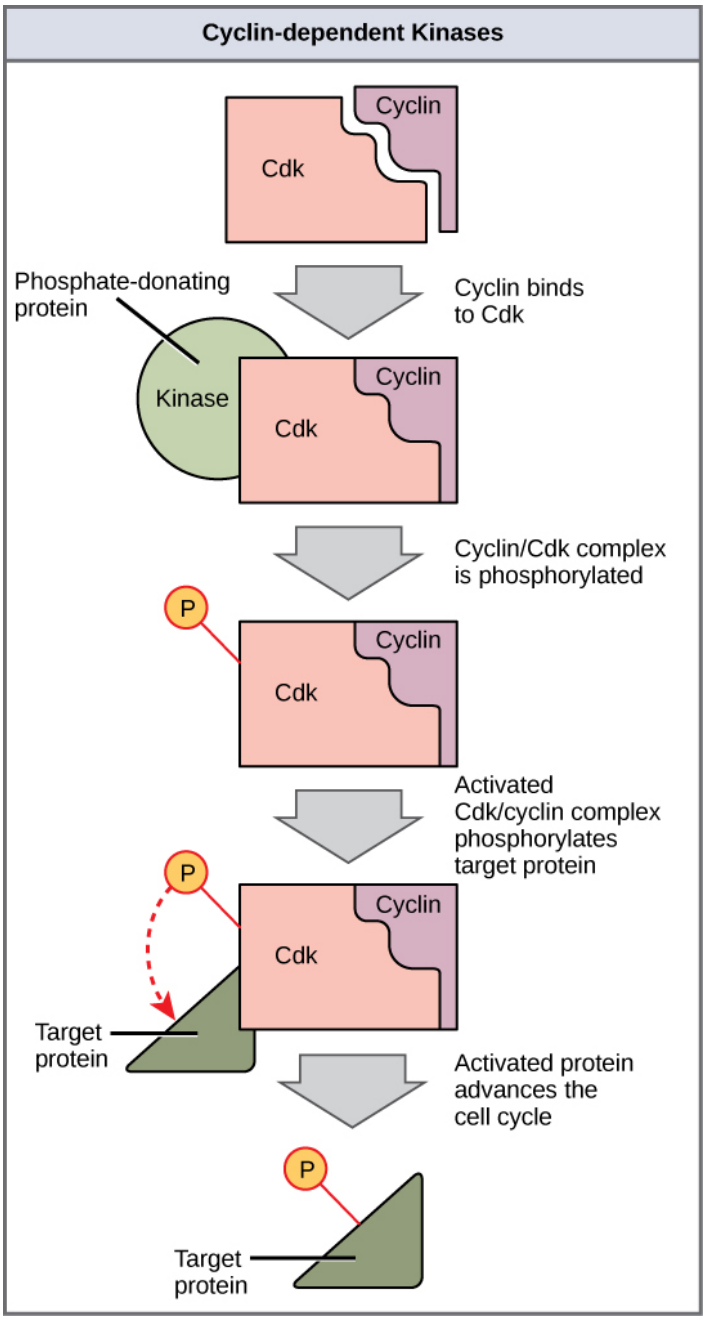


Figure 3 Cyclin-dependent kinases (Cdks) are protein kinases that, when fully activated, can phosphorylate and thus activate other proteins that advance the cell cycle past a checkpoint. To become fully activated, a Cdk must bind to a cyclin protein and then be phosphorylated by another kinase.

Rb, p53, and p21 act primarily at the G_1 checkpoint. p53 is a multi-functional protein that has a major impact on the commitment of a cell to division because it acts when there is damaged DNA in cells that are undergoing the preparatory processes during G_1. If damaged DNA is detected, p53 halts the cell cycle and recruits enzymes to repair the DNA. If the DNA cannot be repaired, p53 can trigger apoptosis, or cell suicide, to prevent the duplication of damaged chromosomes. As p53 levels rise, the production of p21 is triggered. p21 enforces the halt in the cycle dictated by p53 by binding to and inhibiting the activity of the Cdk/cyclin complexes. As a cell is exposed to more stress, higher levels of p53 and p21 accumulate, making it less likely that the cell will move into the S phase.

Rb exerts its regulatory influence on other positive regulator proteins. Chiefly, Rb monitors cell size. In the active, dephosphorylated state, Rb binds to proteins called transcription factors, most commonly, E2F (Figure 4). Transcription factors "turn on" specific genes, allowing the production of proteins encoded by that gene. When Rb is bound to E2F, production of proteins necessary for the G_1/S transition is blocked. As the cell increases in size, Rb is slowly phosphorylated until it becomes inactivated. Rb releases E2F, which can now turn on the gene that produces the transition protein, and this particular block is removed. For the cell to move past each of the checkpoints, all positive regulators must be "turned on," and all negative regulators must be "turned off."

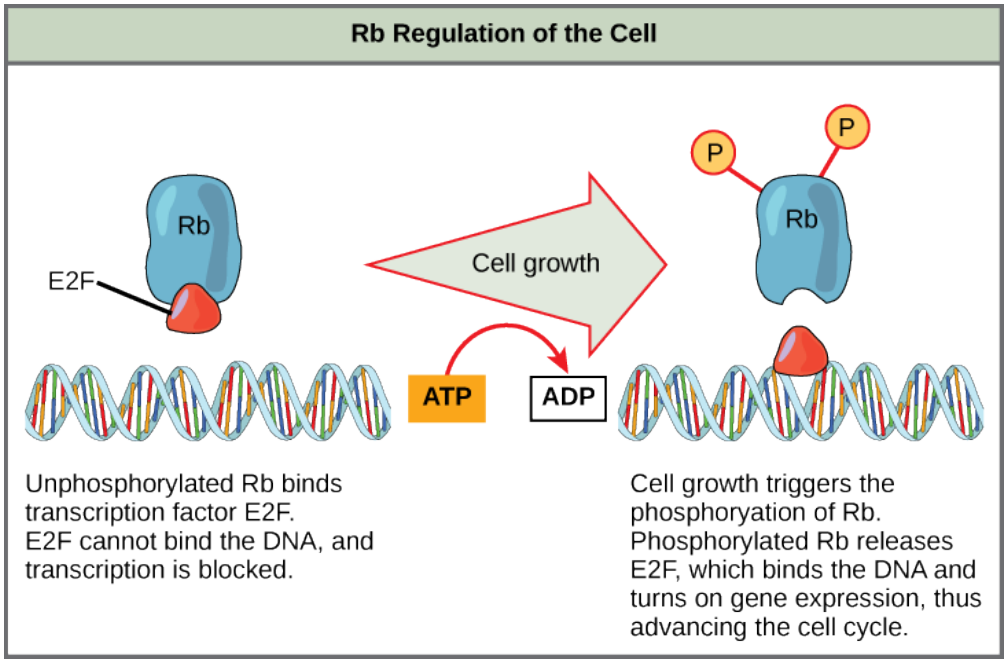


Figure 4. Rb halts the cell cycle and releases its hold in response to cell growth.

Rb and other proteins that negatively regulate the cell cycle are sometimes called tumor suppressors. Why do you think the name tumor suppressor might be appropriate for these proteins?

Section Summary

Each step of the cell cycle is monitored by internal controls called checkpoints. There are three major checkpoints in the cell cycle: one near the end of G_1, a second at the G_2/M transition, and the third during metaphase. Positive regulator molecules allow the cell cycle to advance to the next stage. Negative regulator molecules monitor cellular conditions and can halt the cycle until specific requirements are met.

Visit this page in your course online to check your understanding.

Additional Self Check Questions

1. Rb and other proteins that negatively regulate the cell cycle are sometimes called tumor suppressors. Why do you think the name tumor suppressor might be an appropriate for these proteins?

2. Describe the general conditions that must be met at each of the three main cell cycle checkpoints.

3. Explain the roles of the positive cell cycle regulators compared to the negative regulators.

4. What steps are necessary for Cdk to become fully active?

5. Rb is a negative regulator that blocks the cell cycle at the G_1 checkpoint until the cell achieves a requisite size. What molecular mechanism does Rb employ to halt the cell cycle?

Answers

1. Rb and other negative regulatory proteins control cell division and therefore prevent the formation of tumors. Mutations that prevent these proteins from carrying out their function can result in cancer.

2. The G_1 checkpoint monitors adequate cell growth, the state of the genomic DNA, adequate stores of energy, and materials for S phase. At the G_2 checkpoint, DNA is checked to ensure that all chromosomes were duplicated and that there are no mistakes in newly synthesized DNA. Additionally, cell size and energy reserves are evaluated. The M checkpoint confirms the correct attachment of the mitotic spindle fibers to the kinetochores.

3. Positive cell regulators such as cyclin and Cdk perform tasks that advance the cell cycle to the next stage. Negative regulators such as Rb, p53, and p21 block the progression of the cell cycle until certain events have occurred.

4. Cdk must bind to a cyclin, and it must be phosphorylated in the correct position to become fully active.

5. Rb is active when it is dephosphorylated. In this state, Rb binds to E2F, which is a transcription factor required for the transcription and eventual translation of molecules required for the G_1/S transition. E2F cannot transcribe certain genes when it is bound to Rb. As the cell increases in size, Rb becomes phosphorylated, inactivated, and releases E2F. E2F can then promote the transcription of the genes it controls, and the transition proteins will be produced.

Glossary

cell cycle checkpoint: mechanism that monitors the preparedness of a eukaryotic cell to advance through the various cell cycle stages

cyclin: one of a group of proteins that act in conjunction with cyclin-dependent kinases to help regulate the cell cycle by phosphorylating key proteins; the concentrations of cyclins fluctuate throughout the cell cycle

cyclin-dependent kinase: one of a group of protein kinases that helps to regulate the cell cycle when bound to cyclin; it functions to phosphorylate other proteins that are either activated or inactivated by phosphorylation

p21: cell cycle regulatory protein that inhibits the cell cycle; its levels are controlled by p53

p53: cell cycle regulatory protein that regulates cell growth and monitors DNA damage; it halts the progression of the cell cycle in cases of DNA damage and may induce apoptosis

retinoblastoma protein (Rb): regulatory molecule that exhibits negative effects on the cell cycle by interacting with a transcription factor (E2F)

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: Open Stax. **Located at:** http://cnx.org/contents/185cbf87-c72e-48f5-b51e-f14f21b5eabd@9.17:1/Biology. **License**: *CC BY: Attribution*

CANCER AND THE CELL CYCLE

Learning Objectives:

By the end of this section, you will be able to:

- Describe how cancer is caused by uncontrolled cell growth
- Understand how proto-oncogenes are normal cell genes that, when mutated, become oncogenes
- Describe how tumor suppressors function
- Explain how mutant tumor suppressors cause cancer

Cancer comprises many different diseases caused by a common mechanism: uncontrolled cell growth. Despite the redundancy and overlapping levels of cell cycle control, errors do occur. One of the critical processes monitored by the cell cycle checkpoint surveillance mechanism is the proper replication of DNA during the S phase. Even when all of the cell cycle controls are fully functional, a small percentage of replication errors (mutations) will be passed on to the daughter cells. If changes to the DNA nucleotide sequence occur within a coding portion of a gene and are not corrected, a gene mutation results. All cancers start when a gene mutation gives rise to a faulty protein that plays a key role in cell reproduction. The change in the cell that results from the malformed protein may be minor: perhaps a slight delay in the binding of Cdk to cyclin or an Rb protein that detaches from its target DNA while still phosphorylated. Even minor mistakes, however, may allow subsequent mistakes to occur more readily. Over and over, small uncorrected errors are passed from the parent cell to the daughter cells and amplified as each generation produces more non-functional proteins from uncorrected DNA damage. Eventually, the pace of the cell cycle speeds up as the effectiveness of the control and repair mechanisms decreases. Uncontrolled growth of the mutated cells outpaces the growth of normal cells in the area, and a tumor ("-oma") can result.

Proto-oncogenes

The genes that code for the positive cell cycle regulators are called proto-oncogenes. Proto-oncogenes are normal genes that, when mutated in certain ways, become oncogenes, genes that cause a cell to become cancerous. Consider what might happen to the cell cycle in a cell with a recently acquired oncogene. In most instances, the alteration of the DNA sequence will result in a less functional (or non-functional) protein. The result is detrimental to the cell and will likely prevent the cell from completing the cell cycle; however, the organism is not harmed because the mutation will not be carried forward. If a cell cannot reproduce, the mutation is not propagated and the damage is minimal. Occasionally, however, a gene mutation causes a change that increases the activity of a positive regulator. For example, a mutation that allows Cdk to be activated without being partnered with cyclin could push the cell cycle past a checkpoint before all of the required conditions are met. If the resulting daughter cells are too damaged to undergo further cell divisions, the mutation would not be propagated and no harm would come to the organism. However, if the atypical daughter cells are able to undergo further cell divisions, subsequent generations of cells will probably accumulate even more mutations, some possibly in additional genes that regulate the cell cycle.

The Cdk gene in the above example is only one of many genes that are considered proto-oncogenes. In addition to the cell cycle regulatory proteins, any protein that influences the cycle can be altered in such a way as to override cell cycle checkpoints. An oncogene is any gene that, when altered, leads to an increase in the rate of cell cycle progression.

Tumor Suppressor Genes

Like proto-oncogenes, many of the negative cell cycle regulatory proteins were discovered in cells that had become cancerous. Tumor suppressor genes are segments of DNA that code for negative regulator proteins, the

type of regulators that, when activated, can prevent the cell from undergoing uncontrolled division. The collective function of the best-understood tumor suppressor gene proteins, Rb, p53, and p21, is to put up a roadblock to cell cycle progression until certain events are completed. A cell that carries a mutated form of a negative regulator might not be able to halt the cell cycle if there is a problem. Tumor suppressors are similar to brakes in a vehicle: Malfunctioning brakes can contribute to a car crash.

Mutated p53 genes have been identified in more than one-half of all human tumor cells. This discovery is not surprising in light of the multiple roles that the p53 protein plays at the G_1 checkpoint. A cell with a faulty p53 may fail to detect errors present in the genomic DNA (Figure 1). Even if a partially functional p53 does identify the mutations, it may no longer be able to signal the necessary DNA repair enzymes. Either way, damaged DNA will remain uncorrected. At this point, a functional p53 will deem the cell unsalvageable and trigger programmed cell death (apoptosis). The damaged version of p53 found in cancer cells, however, cannot trigger apoptosis.

Art Connection

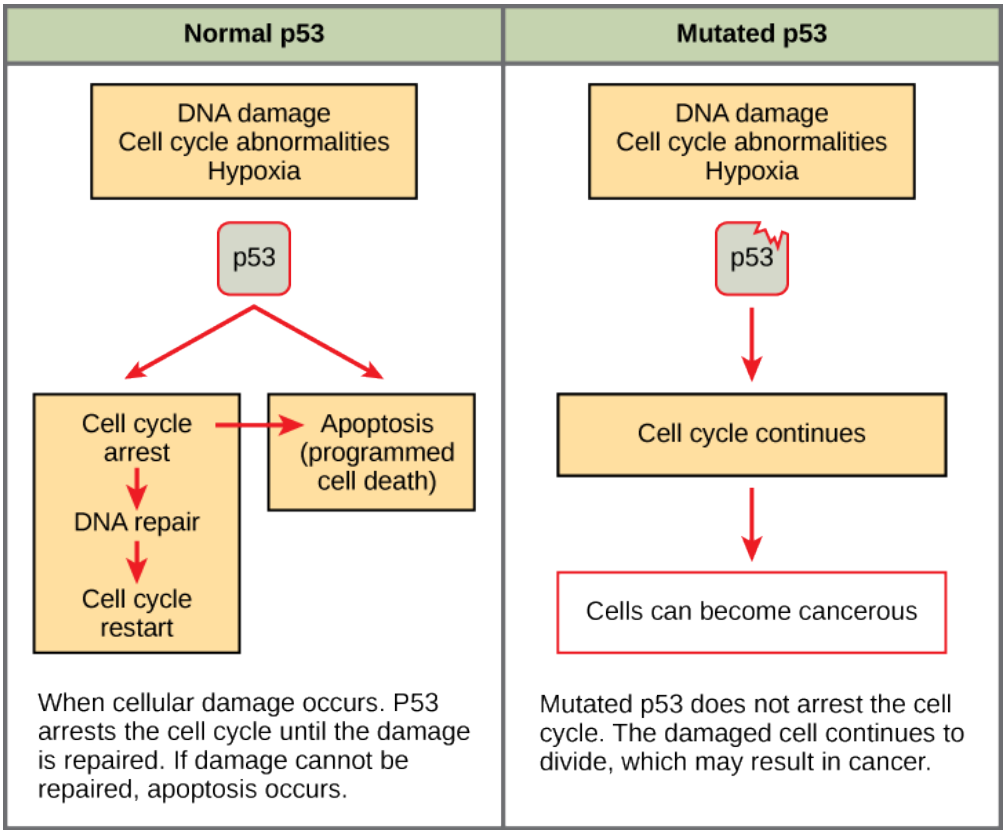


Figure 1 The role of normal p53 is to monitor DNA and the supply of oxygen (hypoxia is a condition of reduced oxygen supply). If damage is detected, p53 triggers repair mechanisms. If repairs are unsuccessful, p53 signals apoptosis. A cell with an abnormal p53 protein cannot repair damaged DNA and thus cannot signal apoptosis. Cells with abnormal p53 can become cancerous. (credit: modification of work by Thierry Soussi)

Human papillomavirus can cause cervical cancer. The virus encodes E6, a protein that binds p53. Based on this fact and what you know about p53, what effect do you think E6 binding has on p53 activity?

1. E6 activates p53
2. E6 inactivates p53
3. E6 mutates p53
4. E6 binding marks p53 for degradation

The loss of p53 function has other repercussions for the cell cycle. Mutated p53 might lose its ability to trigger p21 production. Without adequate levels of p21, there is no effective block on Cdk activation. Essentially, without a fully functional p53, the G_1 checkpoint is severely compromised and the cell proceeds directly from G_1 to S regardless of internal and external conditions. At the completion of this shortened cell cycle, two daughter cells are produced that have inherited the mutated p53 gene. Given the non-optimal conditions under which the parent cell reproduced, it is likely that the daughter cells will have acquired other mutations in addition to the faulty tumor suppressor gene. Cells such as these daughter cells quickly accumulate both oncogenes and non-functional tumor suppressor genes. Again, the result is tumor growth.

Link to Learning

Go to this website to watch an animation of how cancer results from errors in the cell cycle.

Section Summary

Cancer is the result of unchecked cell division caused by a breakdown of the mechanisms that regulate the cell cycle. The loss of control begins with a change in the DNA sequence of a gene that codes for one of the regulatory molecules. Faulty instructions lead to a protein that does not function as it should. Any disruption of the monitoring system can allow other mistakes to be passed on to the daughter cells. Each successive cell division will give rise to daughter cells with even more accumulated damage. Eventually, all checkpoints become nonfunctional, and rapidly reproducing cells crowd out normal cells, resulting in a tumor or leukemia (blood cancer).

Visit this page in your course online to check your understanding.

Additional Self Check Questions

1. Human papillomavirus can cause cervical cancer. The virus encodes E6, a protein that binds p53. Based on this fact and what you know about p53, what effect do you think E6 binding has on p53 activity?

    A. E6 activates p53

    B. E6 inactivates p53

    C. E6 mutates p53

    D. E6 binding marks p53 for degradation

2. Outline the steps that lead to a cell becoming cancerous.

3. Explain the difference between a proto-oncogene and a tumor suppressor gene.

4. List the regulatory mechanisms that might be lost in a cell producing faulty p53.

5. p53 can trigger apoptosis if certain cell cycle events fail. How does this regulatory outcome benefit a multicellular organism?

Answers

1. D. E6 binding marks p53 for degradation.

2. If one of the genes that produces regulator proteins becomes mutated, it produces a malformed, possibly non-functional, cell cycle regulator, increasing the chance that more mutations will be left unrepaired in the cell. Each subsequent generation of cells sustains more damage. The cell cycle can speed up as a result of the loss of functional checkpoint proteins. The cells can lose the ability to self-destruct and eventually become "immortalized."

3. A proto-oncogene is a segment of DNA that codes for one of the positive cell cycle regulators. If that gene becomes mutated so that it produces a hyperactivated protein product, it is considered an oncogene. A tumor suppressor gene is a segment of DNA that codes for one of the negative cell cycle regulators. If that gene becomes mutated so that the protein product becomes less active, the cell cycle will run unchecked. A single oncogene can initiate abnormal cell divisions; however, tumor suppressors lose their effectiveness only when both copies of the gene are damaged.

4. Regulatory mechanisms that might be lost include monitoring of the quality of the genomic DNA, recruiting of repair enzymes, and the triggering of apoptosis.

5. If a cell has damaged DNA, the likelihood of producing faulty proteins is higher. The daughter cells of such a damaged parent cell would also produce faulty proteins that might eventually become cancerous. If p53 recognizes this damage and triggers the cell to self-destruct, the damaged DNA is degraded and recycled. No further harm comes to the organism. Another healthy cell is triggered to divide instead.

Glossary

oncogene: mutated version of a normal gene involved in the positive regulation of the cell cycle

proto-oncogene: normal gene that when mutated becomes an oncogene

tumor suppressor gene: segment of DNA that codes for regulator proteins that prevent the cell from undergoing uncontrolled division

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: Open Stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License**: *CC BY: Attribution*

PROKARYOTIC CELL DIVISION

LEARNING OBJECTIVES

By the end of this section, you will be able to:

- Describe the process of binary fission in prokaryotes
- Explain how FtsZ and tubulin proteins are examples of homology

Prokaryotes, such as bacteria, propagate by binary fission. For unicellular organisms, cell division is the only method to produce new individuals. In both prokaryotic and eukaryotic cells, the outcome of cell reproduction is a pair of daughter cells that are genetically identical to the parent cell. In unicellular organisms, daughter cells are individuals.

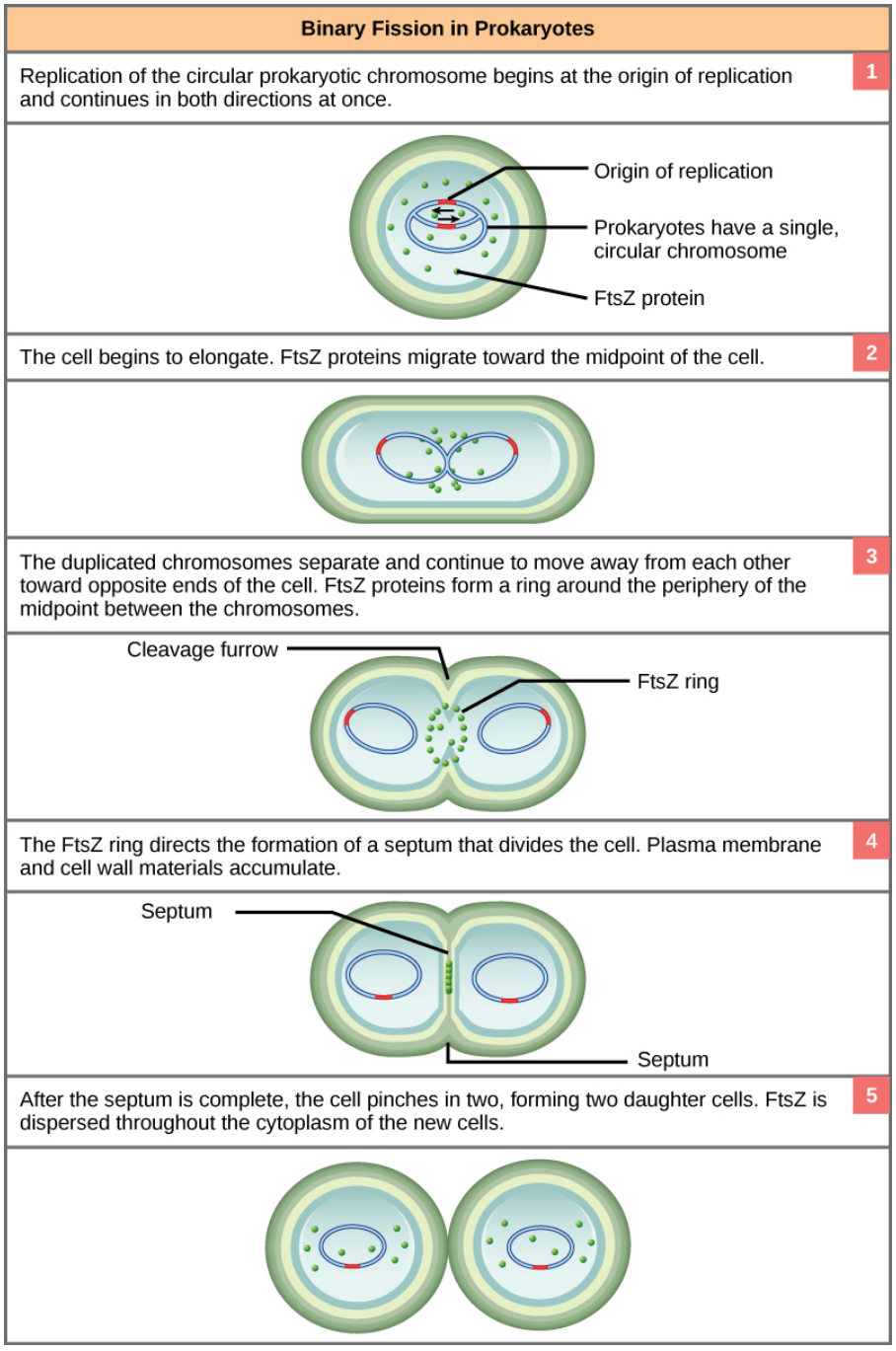
To achieve the outcome of cloned offspring, certain steps are essential. The genomic DNA must be replicated and then allocated into the daughter cells; the cytoplasmic contents must also be divided to give both new cells

the machinery to sustain life. In bacterial cells, the genome consists of a single, circular DNA chromosome; therefore, the process of cell division is simplified. Karyokinesis is unnecessary because there is no nucleus and thus no need to direct one copy of the multiple chromosomes into each daughter cell. This type of cell division is called binary (prokaryotic) fission.

Binary Fission

Due to the relative simplicity of the prokaryotes, the cell division process, called binary fission, is a less complicated and much more rapid process than cell division in eukaryotes. The single, circular DNA chromosome of bacteria is not enclosed in a nucleus, but instead occupies a specific location, the nucleoid, within the cell ([link]). Although the DNA of the nucleoid is associated with proteins that aid in packaging the molecule into a compact size, there are no histone proteins and thus no nucleosomes in prokaryotes. The packing proteins of bacteria are, however, related to the cohesin and condensin proteins involved in the chromosome compaction of eukaryotes.

The bacterial chromosome is attached to the plasma membrane at about the midpoint of the cell. The starting point of replication, the origin, is close to the binding site of the chromosome to the plasma membrane (Figure). Replication of the DNA is bidirectional, moving away from the origin on both strands of the loop simultaneously. As the new double strands are formed, each origin point moves away from the cell wall attachment toward the opposite ends of the cell. As the cell elongates, the growing membrane aids in the transport of the chromosomes. After the chromosomes have cleared the midpoint of the elongated cell, cytoplasmic separation begins. The formation of a ring composed of repeating units of a protein called FtsZ directs the partition between the nucleoids. Formation of the FtsZ ring triggers the accumulation of other proteins that work together to recruit new membrane and cell wall materials to the site. A septum is formed between the nucleoids, extending gradually from the periphery toward the center of the cell. When the new cell walls are in place, the daughter cells separate.



These images show the steps of binary fission in prokaryotes. (credit: modification of work by "Mcstrother"/Wikimedia Commons)

EVOLUTION CONNECTION

Mitotic Spindle ApparatusThe precise timing and formation of the mitotic spindle is critical to the success of eukaryotic cell division. Prokaryotic cells, on the other hand, do not undergo karyokinesis and therefore have no need for a mitotic spindle. However, the FtsZ protein that plays such a vital role in prokaryotic cytokinesis is structurally and functionally very similar to tubulin, the building block of the microtubules that make up the mitotic spindle fibers that are necessary for eukaryotes. FtsZ proteins can form filaments, rings, and other three-dimensional structures that resemble the way tubulin forms microtubules, centrioles, and various cytoskeletal components. In addition, both FtsZ and tubulin employ the same energy source, GTP (guanosine triphosphate), to rapidly assemble and disassemble complex structures.

FtsZ and tubulin are homologous structures derived from common evolutionary origins. In this example, FtsZ is the ancestor protein to tubulin (a modern protein). While both proteins are found in extant organisms, tubulin

function has evolved and diversified tremendously since evolving from its FtsZ prokaryotic origin. A survey of mitotic assembly components found in present-day unicellular eukaryotes reveals crucial intermediary steps to the complex membrane-enclosed genomes of multicellular eukaryotes (Table).

Cell Division Apparatus among Various Organisms			
	Structure of genetic material	Division of nuclear material	Separation of daughter cells
Prokaryotes	There is no nucleus. The single, circular chromosome exists in a region of cytoplasm called the nucleoid.	Occurs through binary fission. As the chromosome is replicated, the two copies move to opposite ends of the cell by an unknown mechanism.	FtsZ proteins assemble into a ring that pinches the cell in two.
Some protists	Linear chromosomes exist in the nucleus.	Chromosomes attach to the nuclear envelope, which remains intact. The mitotic spindle passes through the envelope and elongates the cell. No centrioles exist.	Microfilaments form a cleavage furrow that pinches the cell in two.
Other protists	Linear chromosomes exist in the nucleus.	A mitotic spindle forms from the centrioles and passes through the nuclear membrane, which remains intact. Chromosomes attach to the mitotic spindle, which separates the chromosomes and elongates the cell.	Microfilaments form a cleavage furrow that pinches the cell in two.
Animal cells	Linear chromosomes exist in the nucleus.	A mitotic spindle forms from the centrosomes. The nuclear envelope dissolves. Chromosomes attach to the mitotic spindle, which separates the chromosomes and elongates the cell.	Microfilaments form a cleavage furrow that pinches the cell in two.

Section Summary

In both prokaryotic and eukaryotic cell division, the genomic DNA is replicated and then each copy is allocated into a daughter cell. In addition, the cytoplasmic contents are divided evenly and distributed to the new cells. However, there are many differences between prokaryotic and eukaryotic cell division. Bacteria have a single, circular DNA chromosome but no nucleus. Therefore, mitosis is not necessary in bacterial cell division. Bacterial cytokinesis is directed by a ring composed of a protein called FtsZ. Ingrowth of membrane and cell wall material from the periphery of the cells results in the formation of a septum that eventually constructs the separate cell walls of the daughter cells.

Review Questions

Which eukaryotic cell cycle event is missing in binary fission?

1. cell growth
2. DNA duplication
3. karyokinesis
4. cytokinesis

FtsZ proteins direct the formation of a _____ that will eventually form the new cell walls of the daughter cells.

1. contractile ring
2. cell plate
3. cytoskeleton
4. septum

Free Response

Name the common components of eukaryotic cell division and binary fission.

Describe how the duplicated bacterial chromosomes are distributed into new daughter cells without the direction of the mitotic spindle.

Glossary

binary fission prokaryotic cell division process

FtsZ tubulin-like protein component of the prokaryotic cytoskeleton that is important in prokaryotic cytokinesis (name origin:**F**ilamenting **t**emperature-**s**ensitive mutant **Z**)

origin (also, ORI) region of the prokaryotic chromosome where replication begins (origin of replication)

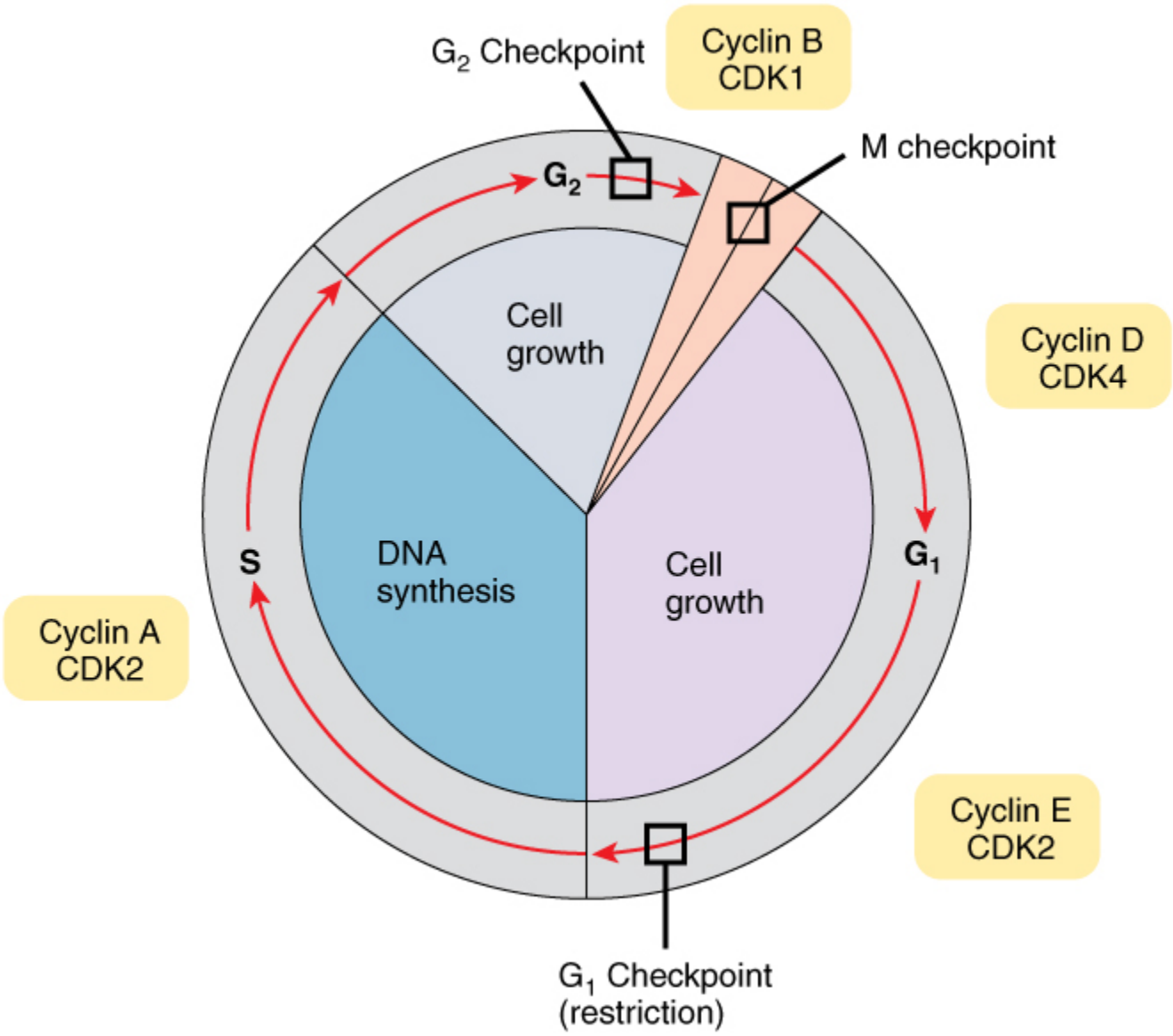
septum structure formed in a bacterial cell as a precursor to the separation of the cell into two daughter cells

Licensing & Attributions

CC licensed content, Shared previously

- Prokaryotic Cell Division. **Authored by:** OpenStax College. **Located at:** http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License:** *CC BY: Attribution*

FIGURE: CELL CYCLE WITH CYCLINS AND CHECKPOINTS



Licensing & Attributions

CC licensed content, Shared previously

- Image: Cell Cycle With Cyclins and Checkpoints. **Provided by:** Wikipedia. **Located at:** http://upload.wikimedia.org/wikipedia/commons/3/38/0332_Cell_Cycle_With_Cyclins_and_Checkpoints.jpg. **License:** *CC BY-SA: Attribution-ShareAlike*

STUDY GUIDE: MITOSIS

Study Questions

Objective: Illustrate the process of asexual reproduction in eukaryotic cells.

Use this page to check your understanding of the content.

Vocabulary

1. Chromosome
2. Chromatin
3. Haploid
4. Diploid
5. Homologous chromosome
6. Sister chromatid
7. Cytokinesis
8. Somatic cell
9. Gamete
10. Karyotype

Study Guide Questions

1. Generally compare and contrast mitosis and meiosis.
2. Carefully compare and contrast chromosomes and chromatin.
3. Explain the advantages/disadvantages of DNA in chromatin form, vs. chromosome form. Relate your response to the stages in the cell cycle when DNA is found in each form.
4. What are homologous chromosomes?
5. Define HAPLOID and DIPLOID.
6. How many chromosomes do humans have?
7. Where did your chromosomes come from?
8. Be able to label all stages in the cell cycle. Also be able to say exactly WHAT is happening during each stage.
9. What is the purpose of "checkpoints" during interphase?
10. What is a centriole?
11. Be able to describe what is happening in a cell during EACH stage in mitosis.
12. Be able to determine the number of chromosomes AND the amount of DNA in a cell during each stage of the cell cycle.
13. Given any diagram or picture of a cell in a phase of mitosis, be able to identify the phase.
14. Distinguish between cytokinesis in an animal cell and a plant cell.

Licensing & Attributions

CC licensed content, Shared previously

- Study Guide: Mitosis. **Authored by**: Wendy Riggs. **Provided by**: College of the Redwoods. **Project**: Kaleidoscope. **License**: *CC BY: Attribution*

INTRODUCTION TO MEIOSIS AND SEXUAL REPRODUCTION

The ability to reproduce in kind is a basic characteristic of all living things. In kind means that the offspring of any organism closely resemble their parent or parents. Hippopotamuses give birth to hippopotamus calves, Joshua trees produce seeds from which Joshua tree seedlings emerge, and adult flamingos lay eggs that hatch into flamingo chicks. In kind does not generally mean exactly the same. Whereas many unicellular organisms and a few multicellular organisms can produce genetically identical clones of themselves through cell division, many single-celled organisms and most multicellular organisms reproduce regularly using another method. Sexual reproduction is the production by parents of two haploid cells and the fusion of two haploid cells to form a single, unique diploid cell. In most plants and animals, through tens of rounds of mitotic cell division, this diploid cell will develop into an adult organism. Haploid cells that are part of the sexual reproductive cycle are produced by a type of cell division called meiosis. Sexual reproduction, specifically meiosis and fertilization, introduces variation into offspring that may account for the evolutionary success of sexual reproduction. The vast majority of eukaryotic organisms, both multicellular and unicellular, can or must employ some form of meiosis and fertilization to reproduce.

(a) (b) (c)

Figure 1. Each of us, like these other large multicellular organisms, begins life as a fertilized egg. After trillions of cell divisions, each of us develops into a complex, multicellular organism. (credit a: modification of work by Frank Wouters; credit b: modification of work by Ken Cole, USGS; credit c: modification of work by Martin Pettitt)

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: Open Stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License**: *CC BY: Attribution*

THE PROCESS OF MEIOSIS

Learning Objectives
By the end of this section, you will be able to:

- Describe the behavior of chromosomes during meiosis
- Describe cellular events during meiosis
- Explain the differences between meiosis and mitosis
- Explain the mechanisms within meiosis that generate genetic variation among the products of meiosis

Sexual reproduction requires fertilization, the union of two cells from two individual organisms. If those two cells each contain one set of chromosomes, then the resulting cell contains two sets of chromosomes. Haploid cells contain one set of chromosomes. Cells containing two sets of chromosomes are called diploid. The number of sets of chromosomes in a cell is called its ploidy level. If the reproductive cycle is to continue, then the diploid cell must somehow reduce its number of chromosome sets before fertilization can occur again, or there will be a continual doubling in the number of chromosome sets in every generation. So, in addition to fertilization, sexual reproduction includes a nuclear division that reduces the number of chromosome sets.

Most animals and plants are diploid, containing two sets of chromosomes. In each **somatic cell** of the organism (all cells of a multicellular organism except the gametes or reproductive cells), the nucleus contains two copies of each chromosome, called homologous chromosomes. Somatic cells are sometimes referred to as "body" cells. Homologous chromosomes are matched pairs containing the same genes in identical locations along their length. Diploid organisms inherit one copy of each homologous chromosome from each parent; all together, they are considered a full set of chromosomes. Haploid cells, containing a single copy of each homologous chromosome, are found only within structures that give rise to either gametes or spores. **Spores** are haploid cells that can produce a haploid organism or can fuse with another spore to form a diploid cell. All animals and most plants produce eggs and sperm, or gametes. Some plants and all fungi produce spores.

The nuclear division that forms haploid cells, which is called **meiosis**, is related to mitosis. As you have learned, mitosis is the part of a cell reproduction cycle that results in identical daughter nuclei that are also genetically identical to the original parent nucleus. In mitosis, both the parent and the daughter nuclei are at the same ploidy level—diploid for most plants and animals. Meiosis employs many of the same mechanisms as mitosis. However, the starting nucleus is always diploid and the nuclei that result at the end of a meiotic cell division are haploid. To achieve this reduction in chromosome number, meiosis consists of one round of chromosome duplication and two rounds of nuclear division. Because the events that occur during each of the division stages are analogous to the events of mitosis, the same stage names are assigned. However, because there are two rounds of division, the major process and the stages are designated with a "I" or a "II." Thus, **meiosis I** is the first round of meiotic division and consists of prophase I, prometaphase I, and so on. **Meiosis II**, in which the second round of meiotic division takes place, includes prophase II, prometaphase II, and so on.

Meiosis I

Meiosis is preceded by an interphase consisting of the G_1, S, and G_2 phases, which are nearly identical to the phases preceding mitosis. The G_1 phase, which is also called the first gap phase, is the first phase of the interphase and is focused on cell growth. The S phase is the second phase of interphase, during which the DNA of the chromosomes is replicated. Finally, the G_2 phase, also called the second gap phase, is the third and final phase of interphase; in this phase, the cell undergoes the final preparations for meiosis.

During DNA duplication in the S phase, each chromosome is replicated to produce two identical copies, called sister chromatids, that are held together at the centromere by **cohesin** proteins. Cohesin holds the chromatids together until anaphase II. The centrosomes, which are the structures that organize the microtubules of the meiotic spindle, also replicate. This prepares the cell to enter prophase I, the first meiotic phase.

Prophase I

Early in prophase I, before the chromosomes can be seen clearly microscopically, the homologous chromosomes are attached at their tips to the nuclear envelope by proteins. As the nuclear envelope begins to break down, the proteins associated with homologous chromosomes bring the pair close to each other. Recall that, in mitosis, homologous chromosomes do not pair together. In mitosis, homologous chromosomes line up end-to-end so that when they divide, each daughter cell receives a sister chromatid from both members of the homologous pair. The **synaptonemal complex**, a lattice of proteins between the homologous chromosomes, first forms at specific locations and then spreads to cover the entire length of the chromosomes. The tight pairing of the homologous chromosomes is called **synapsis**. In synapsis, the genes on the chromatids of the homologous chromosomes are aligned precisely with each other. The synaptonemal complex supports the exchange of chromosomal segments between non-sister homologous chromatids, a process called crossing over. Crossing over can be observed visually after the exchange as **chiasmata** (singular = chiasma) (Figure 1).

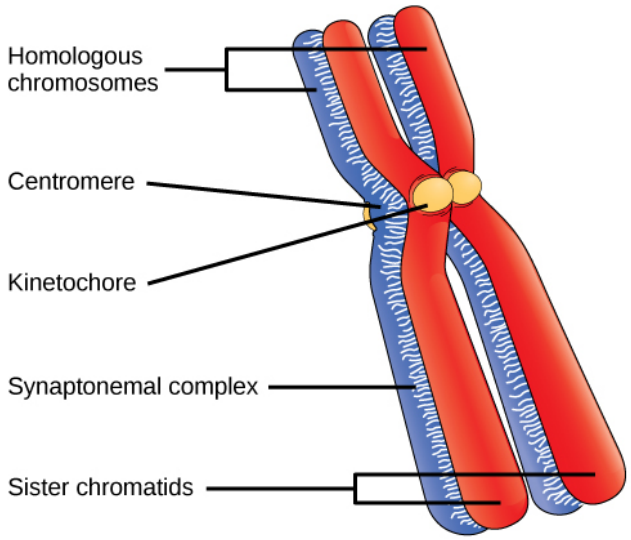


Figure 1. Early in prophase I, homologous chromosomes come together to form a synapse. The chromosomes are bound tightly together and in perfect alignment by a protein lattice called a synaptonemal complex and by cohesin proteins at the centromere.

In species such as humans, even though the X and Y sex chromosomes are not homologous (most of their genes differ), they have a small region of homology that allows the X and Y chromosomes to pair up during prophase I. A partial synaptonemal complex develops only between the regions of homology.

Located at intervals along the synaptonemal complex are large protein assemblies called **recombination nodules**. These assemblies mark the points of later chiasmata and mediate the multistep process of **crossover**—or genetic recombination—between the non-sister chromatids. Near the recombination nodule on each chromatid, the double-stranded DNA is cleaved, the cut ends are modified, and a new connection is made between the non-sister chromatids. As prophase I progresses, the synaptonemal complex begins to break down and the chromosomes begin to condense. When the synaptonemal complex is gone, the homologous chromosomes remain attached to each other at the centromere and at chiasmata. The chiasmata remain until anaphase I. The number of chiasmata varies according to the species and the length of the chromosome. There must be at least one chiasma per chromosome for proper separation of homologous chromosomes during meiosis I, but there may be as many as 25. Following crossover, the synaptonemal complex breaks down and the cohesin connection between homologous pairs is also removed. At the end of prophase I, the pairs are held together only at the chiasmata (Figure 2) and are called **tetrads** because the four sister chromatids of each pair of homologous chromosomes are now visible.

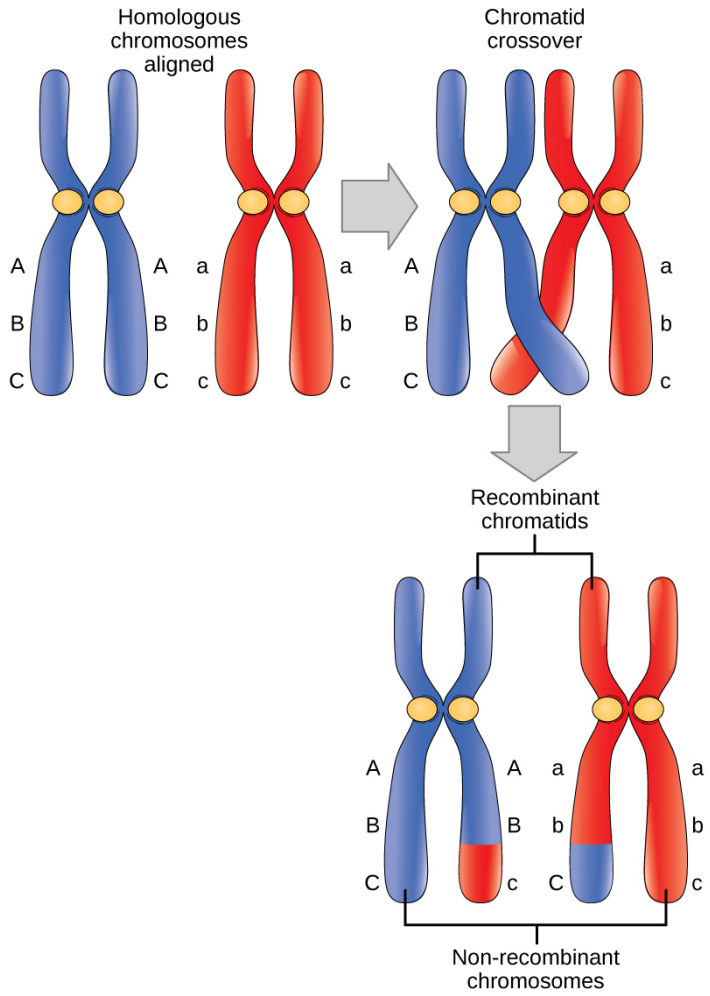


Figure 2. Crossover occurs between non-sister chromatids of homologous chromosomes. The result is an exchange of genetic material between homologous chromosomes.

The crossover events are the first source of genetic variation in the nuclei produced by meiosis. A single crossover event between homologous non-sister chromatids leads to a reciprocal exchange of equivalent DNA between a maternal chromosome and a paternal chromosome. Now, when that sister chromatid is moved into a gamete cell it will carry some DNA from one parent of the individual and some DNA from the other parent. The sister recombinant chromatid has a combination of maternal and paternal genes that did not exist before the crossover. Multiple crossovers in an arm of the chromosome have the same effect, exchanging segments of DNA to create recombinant chromosomes.

Prometaphase I

The key event in prometaphase I is the attachment of the spindle fiber microtubules to the kinetochore proteins at the centromeres. Kinetochore proteins are multiprotein complexes that bind the centromeres of a chromosome to the microtubules of the mitotic spindle. Microtubules grow from centrosomes placed at opposite poles of the cell. The microtubules move toward the middle of the cell and attach to one of the two fused homologous chromosomes. The microtubules attach at each chromosomes' kinetochores. With each member of the homologous pair attached to opposite poles of the cell, in the next phase, the microtubules can pull the homologous pair apart. A spindle fiber that has attached to a kinetochore is called a kinetochore microtubule. At the end of prometaphase I, each tetrad is attached to microtubules from both poles, with one homologous chromosome facing each pole. The homologous chromosomes are still held together at chiasmata. In addition, the nuclear membrane has broken down entirely.

Metaphase I

During metaphase I, the homologous chromosomes are arranged in the center of the cell with the kinetochores facing opposite poles. The homologous pairs orient themselves randomly at the equator. For example, if the two homologous members of chromosome 1 are labeled a and b, then the chromosomes could line up a-b, or b-a. This is important in determining the genes carried by a gamete, as each will only receive one of the two homologous chromosomes. Recall that homologous chromosomes are not identical. They contain slight differences in their genetic information, causing each gamete to have a unique genetic makeup.

This randomness is the physical basis for the creation of the second form of genetic variation in offspring. Consider that the homologous chromosomes of a sexually reproducing organism are originally inherited as two separate sets, one from each parent. Using humans as an example, one set of 23 chromosomes is present in the egg donated by the mother. The father provides the other set of 23 chromosomes in the sperm that fertilizes the egg. Every cell of the multicellular offspring has copies of the original two sets of homologous chromosomes. In prophase I of meiosis, the homologous chromosomes form the tetrads. In metaphase I, these pairs line up at the midway point between the two poles of the cell to form the metaphase plate. Because there is an equal chance that a microtubule fiber will encounter a maternally or paternally inherited chromosome, the arrangement of the tetrads at the metaphase plate is random. Any maternally inherited chromosome may face either pole. Any paternally inherited chromosome may also face either pole. The orientation of each tetrad is independent of the orientation of the other 22 tetrads.

This event—the random (or independent) assortment of homologous chromosomes at the metaphase plate—is the second mechanism that introduces variation into the gametes or spores. In each cell that undergoes meiosis, the arrangement of the tetrads is different. The number of variations is dependent on the number of chromosomes making up a set. There are two possibilities for orientation at the metaphase plate; the possible number of alignments therefore equals $2n$, where n is the number of chromosomes per set. Humans have 23 chromosome pairs, which results in over eight million (2^{23}) possible genetically-distinct gametes. This number does not include the variability that was previously created in the sister chromatids by crossover. Given these two mechanisms, it is highly unlikely that any two haploid cells resulting from meiosis will have the same genetic composition (Figure 3).

To summarize the genetic consequences of meiosis I, the maternal and paternal genes are recombined by crossover events that occur between each homologous pair during prophase I. In addition, the random assortment of tetrads on the metaphase plate produces a unique combination of maternal and paternal chromosomes that will make their way into the gametes.

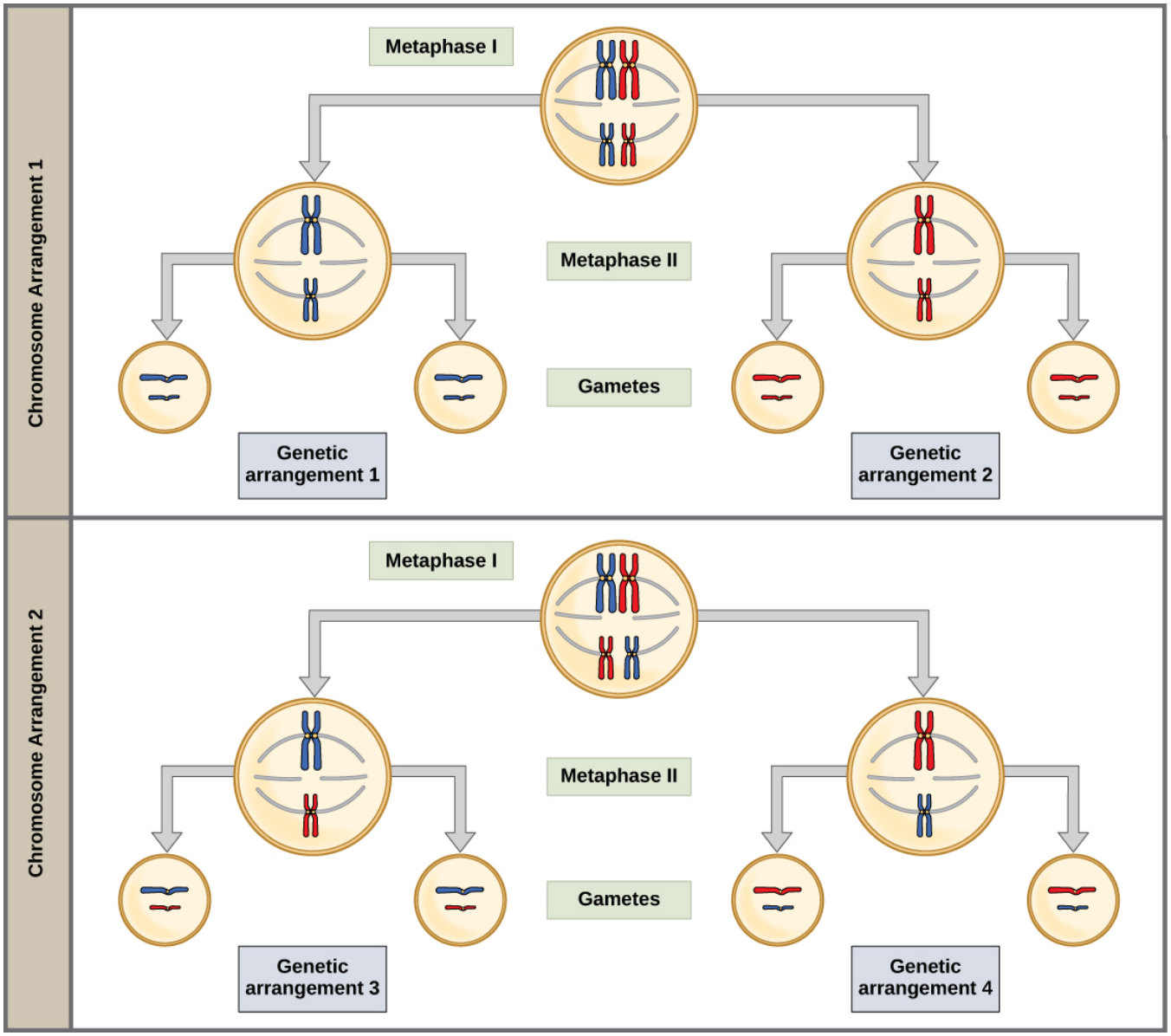


Figure 3. Random, independent assortment during metaphase I can be demonstrated by considering a cell with a set of two chromosomes (n = 2). In this case, there are two possible arrangements at the equatorial plane in metaphase I. The total possible number of different gametes is 2n, where n equals the number of chromosomes in a set. In this example, there are four possible genetic combinations for the gametes. With n = 23 in human cells, there are over 8 million possible combinations of paternal and maternal chromosomes.

Anaphase I

In anaphase I, the microtubules pull the linked chromosomes apart. The sister chromatids remain tightly bound together at the centromere. The chiasmata are broken in anaphase I as the microtubules attached to the fused kinetochores pull the homologous chromosomes apart (Figure 4).

Telophase I and Cytokinesis

In telophase, the separated chromosomes arrive at opposite poles. The remainder of the typical telophase events may or may not occur, depending on the species. In some organisms, the chromosomes decondense and nuclear envelopes form around the chromatids in telophase I. In other organisms, cytokinesis—the physical separation of

the cytoplasmic components into two daughter cells—occurs without reformation of the nuclei. In nearly all species of animals and some fungi, cytokinesis separates the cell contents via a cleavage furrow (constriction of the actin ring that leads to cytoplasmic division). In plants, a cell plate is formed during cell cytokinesis by Golgi vesicles fusing at the metaphase plate. This cell plate will ultimately lead to the formation of cell walls that separate the two daughter cells.

Two haploid cells are the end result of the first meiotic division. The cells are haploid because at each pole, there is just one of each pair of the homologous chromosomes. Therefore, only one full set of the chromosomes is present. This is why the cells are considered haploid—there is only one chromosome set, even though each homolog still consists of two sister chromatids. Recall that sister chromatids are merely duplicates of one of the two homologous chromosomes (except for changes that occurred during crossing over). In meiosis II, these two sister chromatids will separate, creating four haploid daughter cells.

Link to Learning

Review the process of meiosis, observing how chromosomes align and migrate, at Meiosis: An Interactive Animation.

Meiosis II

In some species, cells enter a brief interphase, or **interkinesis**, before entering meiosis II. Interkinesis lacks an S phase, so chromosomes are not duplicated. The two cells produced in meiosis I go through the events of meiosis II in synchrony. During meiosis II, the sister chromatids within the two daughter cells separate, forming four new haploid gametes. The mechanics of meiosis II is similar to mitosis, except that each dividing cell has only one set of homologous chromosomes. Therefore, each cell has half the number of sister chromatids to separate out as a diploid cell undergoing mitosis.

Prophase II

If the chromosomes decondensed in telophase I, they condense again. If nuclear envelopes were formed, they fragment into vesicles. The centrosomes that were duplicated during interkinesis move away from each other toward opposite poles, and new spindles are formed.

Prometaphase II

The nuclear envelopes are completely broken down, and the spindle is fully formed. Each sister chromatid forms an individual kinetochore that attaches to microtubules from opposite poles.

Metaphase II

The sister chromatids are maximally condensed and aligned at the equator of the cell.

Anaphase II

The sister chromatids are pulled apart by the kinetochore microtubules and move toward opposite poles. Non-kinetochore microtubules elongate the cell.

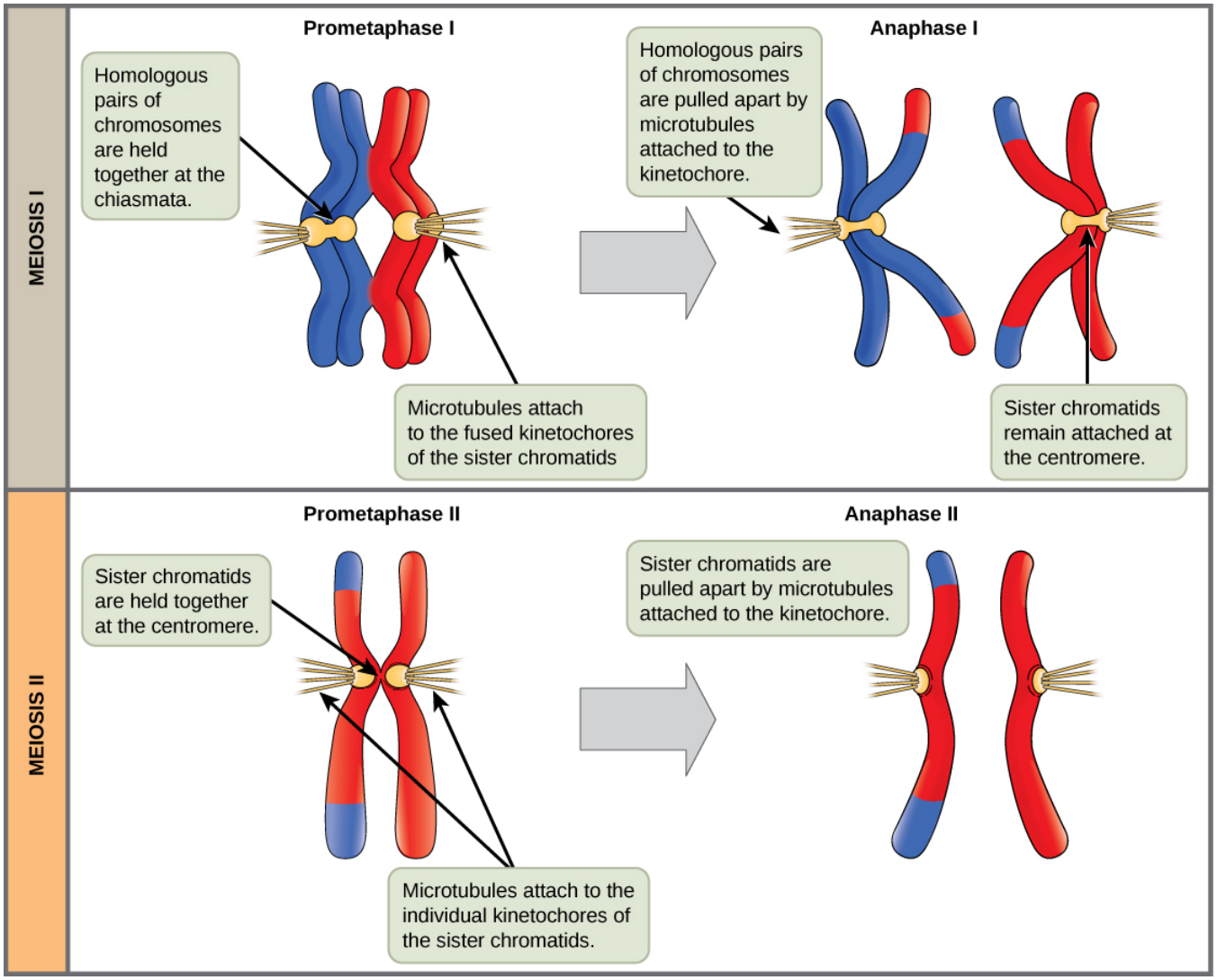


Figure 4. The process of chromosome alignment differs between meiosis I and meiosis II. In prometaphase I, microtubules attach to the fused kinetochores of homologous chromosomes, and the homologous chromosomes are arranged at the midpoint of the cell in metaphase I. In anaphase I, the homologous chromosomes are separated. In prometaphase II, microtubules attach to the kinetochores of sister chromatids, and the sister chromatids are arranged at the midpoint of the cells in metaphase II. In anaphase II, the sister chromatids are separated.

Telophase II and Cytokinesis

The chromosomes arrive at opposite poles and begin to decondense. Nuclear envelopes form around the chromosomes. Cytokinesis separates the two cells into four unique haploid cells. At this point, the newly formed nuclei are both haploid. The cells produced are genetically unique because of the random assortment of paternal and maternal homologs and because of the recombining of maternal and paternal segments of chromosomes (with their sets of genes) that occurs during crossover. The entire process of meiosis is outlined in Figure 5.

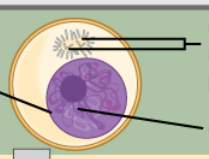
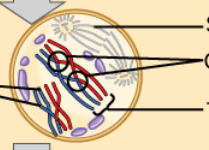
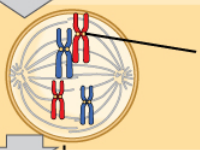
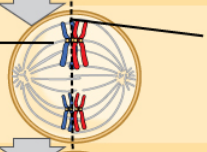
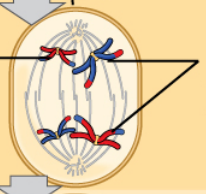
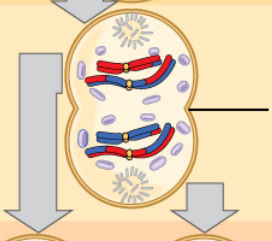
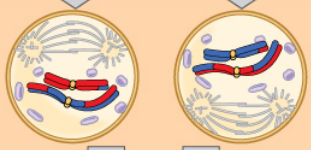
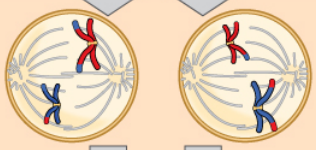
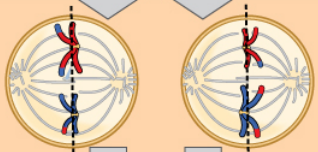
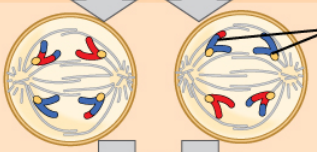
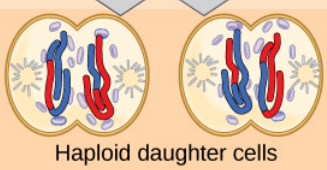
	Stage	Event	Outcome
INTERPHASE	S phase		Chromosomes are duplicated during interphase. The resulting sister chromatids are held together at the centromere. The centrosomes are also duplicated.
MEIOSIS I	Prophase I		Chromosomes condense, and the nuclear envelope fragments. Homologous chromosomes bind firmly together along their length, forming a tetrad. Chiasmata form between non-sister chromatids. Crossing over occurs at the chiasmata. Spindle fibers emerge from the centrosomes.
	Prometaphase I		Homologous chromosomes are attached to spindle microtubules at the fused kinetochore shared by the sister chromatids. Chromosomes continue to condense, and the nuclear envelope completely disappears.
	Metaphase I		Homologous chromosomes randomly assemble at the metaphase plate, where they have been maneuvered into place by the microtubules.
	Anaphase I		Spindle microtubules pull the homologous chromosomes apart. The sister chromatids are still attached at the centromere.
	Telophase I and Cytokinesis		Sister chromatids arrive at the poles of the cell and begin to decondense. A nuclear envelope forms around each nucleus and the cytoplasm is divided by a cleavage furrow. The result is two haploid cells. Each cell contains one duplicated copy of each homologous chromosome pair.
MEIOSIS II	Prophase II		Sister chromatids condense. A new spindle begins to form. The nuclear envelope starts to fragment.
	Prometaphase II		The nuclear envelope disappears, and the spindle fibers engage the individual kinetochores on the sister chromatids.
	Metaphase II		Sister chromatids line up at the metaphase plate.
	Anaphase II		Sister chromatids are pulled apart by the shortening of the kinetochore microtubules. Non-kinetochore microtubules lengthen the cell.
	Telophase II and Cytokinesis		Chromosomes arrive at the poles of the cell and decondense. Nuclear envelopes surround the four nuclei. Cleavage furrows divide the two cells into four haploid cells.

Figure 5. An animal cell with a diploid number of four (2n = 4) proceeds through the stages of meiosis to form four haploid daughter cells.

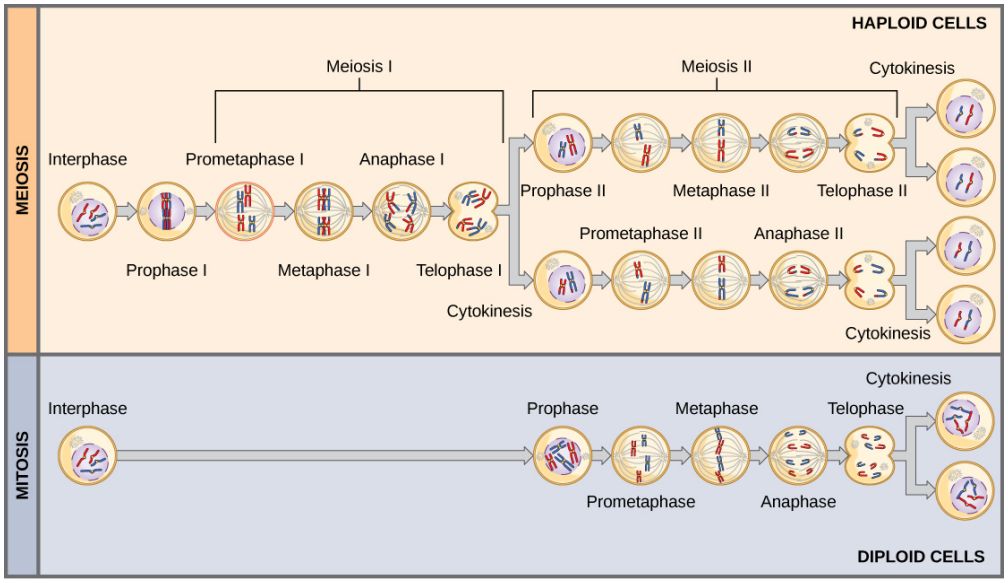
Comparing Meiosis and Mitosis

Mitosis and meiosis are both forms of division of the nucleus in eukaryotic cells. They share some similarities, but also exhibit distinct differences that lead to very different outcomes (Figure 6). Mitosis is a single nuclear division that results in two nuclei that are usually partitioned into two new cells. The nuclei resulting from a mitotic division are genetically identical to the original nucleus. They have the same number of sets of chromosomes, one set in the case of haploid cells and two sets in the case of diploid cells. In most plants and all animal species, it is typically diploid cells that undergo mitosis to form new diploid cells. In contrast, meiosis consists of two nuclear divisions resulting in four nuclei that are usually partitioned into four new cells. The nuclei resulting from meiosis are not genetically identical and they contain one chromosome set only. This is half the number of chromosome sets in the original cell, which is diploid.

The main differences between mitosis and meiosis occur in meiosis I, which is a very different nuclear division than mitosis. In meiosis I, the homologous chromosome pairs become associated with each other, are bound together with the synaptonemal complex, develop chiasmata and undergo crossover between sister chromatids, and line up along the metaphase plate in tetrads with kinetochore fibers from opposite spindle poles attached to each kinetochore of a homolog in a tetrad. All of these events occur only in meiosis I.

When the chiasmata resolve and the tetrad is broken up with the homologs moving to one pole or another, the ploidy level—the number of sets of chromosomes in each future nucleus—has been reduced from two to one. For this reason, meiosis I is referred to as a **reduction division**. There is no such reduction in ploidy level during mitosis.

Meiosis II is much more analogous to a mitotic division. In this case, the duplicated chromosomes (only one set of them) line up on the metaphase plate with divided kinetochores attached to kinetochore fibers from opposite poles. During anaphase II, as in mitotic anaphase, the kinetochores divide and one sister chromatid—now referred to as a chromosome—is pulled to one pole while the other sister chromatid is pulled to the other pole. If it were not for the fact that there had been crossover, the two products of each individual meiosis II division would be identical (like in mitosis). Instead, they are different because there has always been at least one crossover per chromosome. Meiosis II is not a reduction division because although there are fewer copies of the genome in the resulting cells, there is still one set of chromosomes, as there was at the end of meiosis I.



						OUTCOME
PROCESS	DNA synthesis	Synapsis of homologous chromosomes	Crossover	Homologous chromosomes line up at metaphase plate	Sister chromatids line up at metaphase plate	Number and genetic composition of daughter cells
MEIOSIS	Occurs in S phase of interphase	During prophase I	During prophase I	During metaphase I	During metaphase II	Four haploid cells at the end of meiosis II
MITOSIS	Occurs in S phase of interphase	Does not occur in mitosis	Does not occur in mitosis	Does not occur in mitosis	During metaphase	Two diploid cells at the end of mitosis

Figure 6. Meiosis and mitosis are both preceded by one round of DNA replication; however, meiosis includes two nuclear divisions. The four daughter cells resulting from meiosis are haploid and genetically distinct. The daughter cells resulting from mitosis are diploid and identical to the parent cell.

Evolution Connection

The Mystery of the Evolution of Meiosis

Some characteristics of organisms are so widespread and fundamental that it is sometimes difficult to remember that they evolved like other simpler traits. Meiosis is such an extraordinarily complex series of cellular events that biologists have had trouble hypothesizing and testing how it may have evolved. Although meiosis is inextricably entwined with sexual reproduction and its advantages and disadvantages, it is important to separate the questions of the evolution of meiosis and the evolution of sex, because early meiosis may have been advantageous for different reasons than it is now. Thinking outside the box and imagining what the early benefits from meiosis might have been is one approach to uncovering how it may have evolved.

Meiosis and mitosis share obvious cellular processes and it makes sense that meiosis evolved from mitosis. The difficulty lies in the clear differences between meiosis I and mitosis. ((Note: Adam S. Wilkins and Robin Holliday, “The Evolution of Meiosis from Mitosis,” Genetics 181 (2009): 3–12.))

summarized the unique events that needed to occur for the evolution of meiosis from mitosis. These steps are homologous chromosome pairing, crossover exchanges, sister chromatids remaining attached during anaphase, and suppression of DNA replication in interphase. They argue that the first step is the hardest and most important, and that understanding how it evolved would make the evolutionary process clearer. They suggest genetic experiments that might shed light on the evolution of synapsis.

There are other approaches to understanding the evolution of meiosis in progress. Different forms of meiosis exist in single-celled protists. Some appear to be simpler or more “primitive” forms of meiosis. Comparing the meiotic divisions of different protists may shed light on the evolution of meiosis. ((Note: Marilee A. Ramesh, Shehre-Banoo Malik and John M. Logsdon, Jr, “A Phylogenetic Inventory of Meiotic Genes: Evidence for Sex in Giardia and an Early Eukaryotic Origin of Meiosis,” Current Biology 15 (2005):185–91.))

compared the genes involved in meiosis in protists to understand when and where meiosis might have evolved. Although research is still ongoing, recent scholarship into meiosis in protists suggests that some aspects of meiosis may have evolved later than others. This kind of genetic comparison can tell us what aspects of meiosis are the oldest and what cellular processes they may have borrowed from in earlier cells.

Link to Learning

Click through the steps of this interactive animation to compare the meiotic process of cell division to that of mitosis: How Cells Divide.

Section Summary

Sexual reproduction requires that diploid organisms produce haploid cells that can fuse during fertilization to form diploid offspring. As with mitosis, DNA replication occurs prior to meiosis during the S-phase of the cell cycle. Meiosis is a series of events that arrange and separate chromosomes and chromatids into daughter cells. During the interphases of meiosis, each chromosome is duplicated. In meiosis, there are two rounds of nuclear division resulting in four nuclei and usually four daughter cells, each with half the number of chromosomes as the parent cell. The first separates homologs, and the second—like mitosis—separates chromatids into individual chromosomes. During meiosis, variation in the daughter nuclei is introduced because of crossover in prophase I and random alignment of tetrads at metaphase I. The cells that are produced by meiosis are genetically unique.

Meiosis and mitosis share similarities, but have distinct outcomes. Mitotic divisions are single nuclear divisions that produce daughter nuclei that are genetically identical and have the same number of chromosome sets as the original cell. Meiotic divisions include two nuclear divisions that produce four daughter nuclei that are genetically different and have one chromosome set instead of the two sets of chromosomes in the parent cell. The main differences between the processes occur in the first division of meiosis, in which homologous chromosomes are paired and exchange non-sister chromatid segments. The homologous chromosomes separate into different nuclei during meiosis I, causing a reduction of ploidy level in the first division. The second division of meiosis is more similar to a mitotic division, except that the daughter cells do not contain identical genomes because of crossover.

Visit this page in your course online to check your understanding.

Additional Self Check Questions

1. Describe the process that results i the formation of a tetrad.

2. Explain how the random alignment of homologous chromosomes during metaphase I contributes to the variation in gametes produced by meiosis.

3. What is the function of the fused kinetochore found on sister chromatids in prometaphase I?

4. In a comparison of the stages of meiosis to the stages of mitosis, which stages are unique to meiosis and which stages have the same events in both meiosis and mitosis?

Answers

1. During the meiotic interphase, each chromosome is duplicated. The sister chromatids that are formed during synthesis are held together at the centromere region by cohesin proteins. All chromosomes are attached to the nuclear envelope by their tips. As the cell enters prophase I, the nuclear envelope begins to fragment, and the proteins holding homologous chromosomes locate each other. The four sister chromatids align lengthwise, and a protein lattice called the synaptonemal complex is formed between them to bind them together. The synaptonemal complex facilitates crossover between non-sister chromatids, which is observed as chiasmata along the length of the chromosome. As prophase I progresses, the synaptonemal complex breaks down and the sister chromatids become free, except where they are attached by chiasmata. At this stage, the four chromatids are visible in each homologous pairing and are called a tetrad.
2. Random alignment leads to new combinations of traits. The chromosomes that were originally inherited by the gamete-producing individual came equally from the egg and the sperm. In metaphase I, the duplicated copies of these maternal and paternal homologous chromosomes line up across the center of the cell. The orientation of each tetrad is random. There is an equal chance that the maternally derived chromosomes will be facing either pole. The same is true of the paternally derived chromosomes. The alignment should occur differently in almost every meiosis. As the homologous chromosomes are pulled apart in anaphase I, any combination of maternal and paternal chromosomes will move toward each pole. The gametes formed from these two groups of chromosomes will have a mixture of traits from the individual's parents. Each gamete is unique.
3. In metaphase I, the homologous chromosomes line up at the metaphase plate. In anaphase I, the homologous chromosomes are pulled apart and move to opposite poles. Sister chromatids are not separated until meiosis II. The fused kinetochore formed during meiosis I ensures that each spindle microtubule that binds to the tetrad will attach to both sister chromatids.

4. All of the stages of meiosis I, except possibly telophase I, are unique because homologous chromosomes are separated, not sister chromatids. In some species, the chromosomes do not decondense and the nuclear envelopes do not form in telophase I. All of the stages of meiosis II have the same events as the stages of mitosis, with the possible exception of prophase II. In some species, the chromosomes are still condensed and there is no nuclear envelope. Other than this, all processes are the same.

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: Open Stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biologytt. **License**: *CC BY: Attribution*

SEXUAL REPRODUCTION

Learning Objectives

By the end of this section, you will be able to:

- Explain that meiosis and sexual reproduction are evolved traits
- Identify variation among offspring as a potential evolutionary advantage to sexual reproduction
- Describe the three different life-cycle types among sexual multicellular organisms and their commonalities

Sexual reproduction was an early evolutionary innovation after the appearance of eukaryotic cells. It appears to have been very successful because most eukaryotes are able to reproduce sexually, and in many animals, it is the only mode of reproduction. And yet, scientists recognize some real disadvantages to sexual reproduction. On the surface, creating offspring that are genetic clones of the parent appears to be a better system. If the parent organism is successfully occupying a habitat, offspring with the same traits would be similarly successful. There is also the obvious benefit to an organism that can produce offspring whenever circumstances are favorable by asexual budding, fragmentation, or asexual eggs. These methods of reproduction do not require another organism of the opposite sex. Indeed, some organisms that lead a solitary lifestyle have retained the ability to reproduce asexually. In addition, in asexual populations, every individual is capable of reproduction. In sexual populations, the males are not producing the offspring themselves, so in theory an asexual population could grow twice as fast.

However, multicellular organisms that exclusively depend on asexual reproduction are exceedingly rare. Why is sexuality (and meiosis) so common? This is one of the important unanswered questions in biology and has been the focus of much research beginning in the latter half of the twentieth century. There are several possible explanations, one of which is that the variation that sexual reproduction creates among offspring is very important to the survival and reproduction of the population. Thus, on average, a sexually reproducing population will leave more descendants than an otherwise similar asexually reproducing population. The only source of variation in asexual organisms is mutation. This is the ultimate source of variation in sexual organisms, but in addition, those different mutations are continually reshuffled from one generation to the next when different parents combine their unique genomes and the genes are mixed into different combinations by crossovers during prophase I and random assortment at metaphase I.

Evolution Connection

The Red Queen Hypothesis

It is not in dispute that sexual reproduction provides evolutionary advantages to organisms that employ this mechanism to produce offspring. But why, even in the face of fairly stable conditions, does sexual reproduction persist when it is more difficult and costly for individual organisms? Variation is the outcome of sexual reproduction, but why are ongoing variations necessary? Enter the Red Queen hypothesis, first proposed by Leigh Van Valen in 1973. ((Note: Leigh Van Valen, “A New Evolutionary Law,” Evolutionary Theory 1 (1973): 1–30)) The concept was named in reference to the Red Queen’s race in Lewis Carroll’s book, *Through the Looking-Glass*.

All species co-evolve with other organisms; for example predators evolve with their prey, and parasites evolve with their hosts. Each tiny advantage gained by favorable variation gives a species an edge over close competitors, predators, parasites, or even prey. The only method that will allow a co-evolving species to maintain its own share of the resources is to also continually improve its fitness. As one species gains an advantage, this increases selection on the other species; they must also develop an advantage or they will be outcompeted. No single species progresses too far ahead because genetic variation among the progeny of sexual reproduction provides all species with a mechanism to improve rapidly. Species that cannot keep up become extinct. The Red Queen’s catchphrase was, “It takes all the running you can do to stay in the same place.” This is an apt description of co-evolution between competing species.

Life Cycles of Sexually Reproducing Organisms

Fertilization and meiosis alternate in sexual **life cycles**. What happens between these two events depends on the organism. The process of meiosis reduces the chromosome number by half. Fertilization, the joining of two haploid gametes, restores the diploid condition. There are three main categories of life cycles in multicellular organisms: **diploid-dominant**, in which the multicellular diploid stage is the most obvious life stage, such as with most animals including humans; **haploid-dominant**, in which the multicellular haploid stage is the most obvious life stage, such as with all fungi and some algae; and **alternation of generations**, in which the two stages are apparent to different degrees depending on the group, as with plants and some algae.

Diploid-Dominant Life Cycle

Nearly all animals employ a diploid-dominant life-cycle strategy in which the only haploid cells produced by the organism are the gametes. Early in the development of the embryo, specialized diploid cells, called **germ cells**, are produced within the gonads, such as the testes and ovaries. Germ cells are capable of mitosis to perpetuate the cell line and meiosis to produce gametes. Once the haploid gametes are formed, they lose the ability to divide again. There is no multicellular haploid life stage. Fertilization occurs with the fusion of two gametes, usually from different individuals, restoring the diploid state (Figure 1).

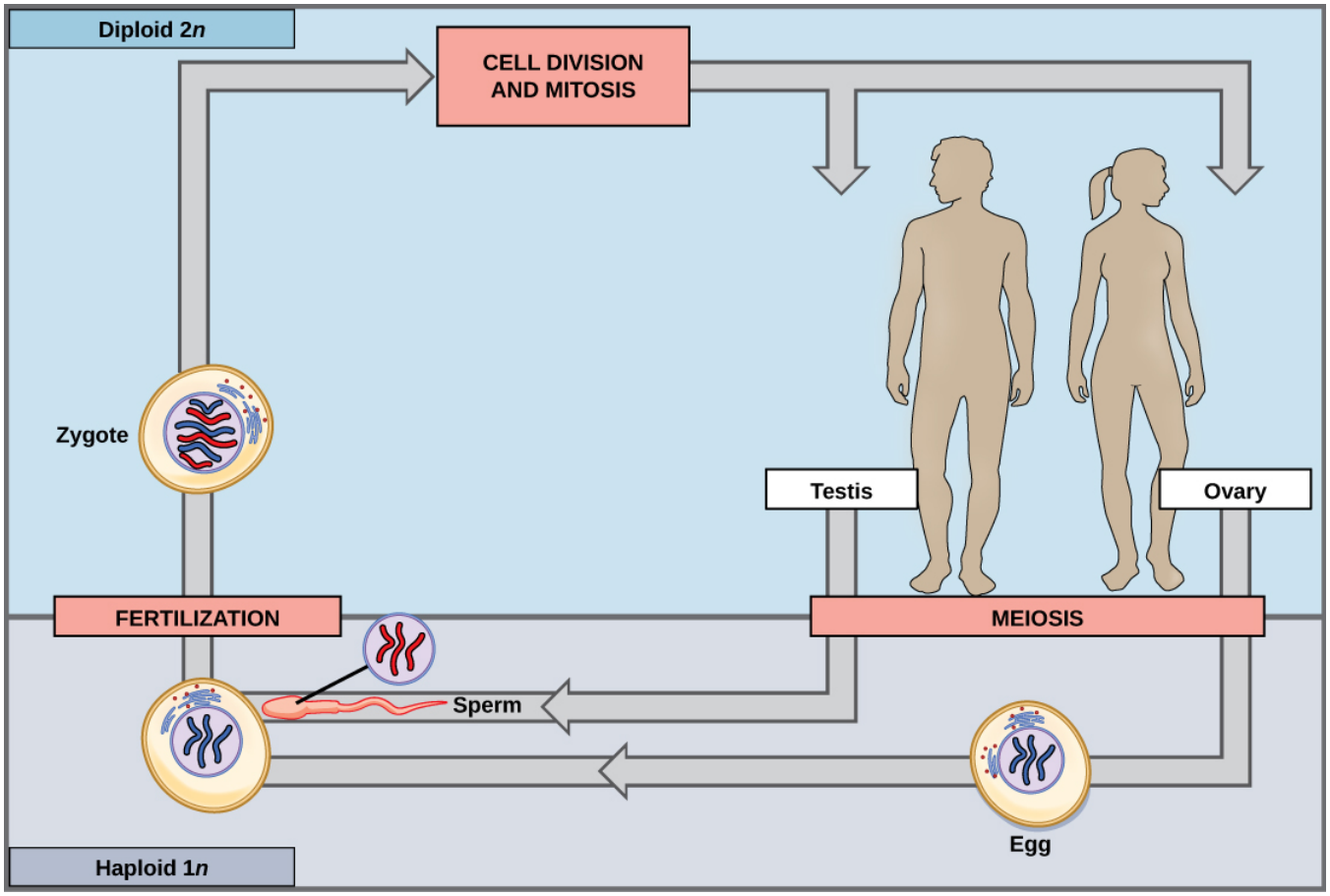


Figure 1. In animals, sexually reproducing adults form haploid gametes from diploid germ cells. Fusion of the gametes gives rise to a fertilized egg cell, or zygote. The zygote will undergo multiple rounds of mitosis to produce a multicellular offspring. The germ cells are generated early in the development of the zygote.

Haploid-Dominant Life Cycle

Most fungi and algae employ a life-cycle type in which the "body" of the organism—the ecologically important part of the life cycle—is haploid. The haploid cells that make up the tissues of the dominant multicellular stage are formed by mitosis. During sexual reproduction, specialized haploid cells from two individuals, designated the (+) and (–) mating types, join to form a diploid zygote. The zygote immediately undergoes meiosis to form four haploid cells called spores. Although haploid like the "parents," these spores contain a new genetic combination from two parents. The spores can remain dormant for various time periods. Eventually, when conditions are conducive, the spores form multicellular haploid structures by many rounds of mitosis (Figure 2).

Art Connection

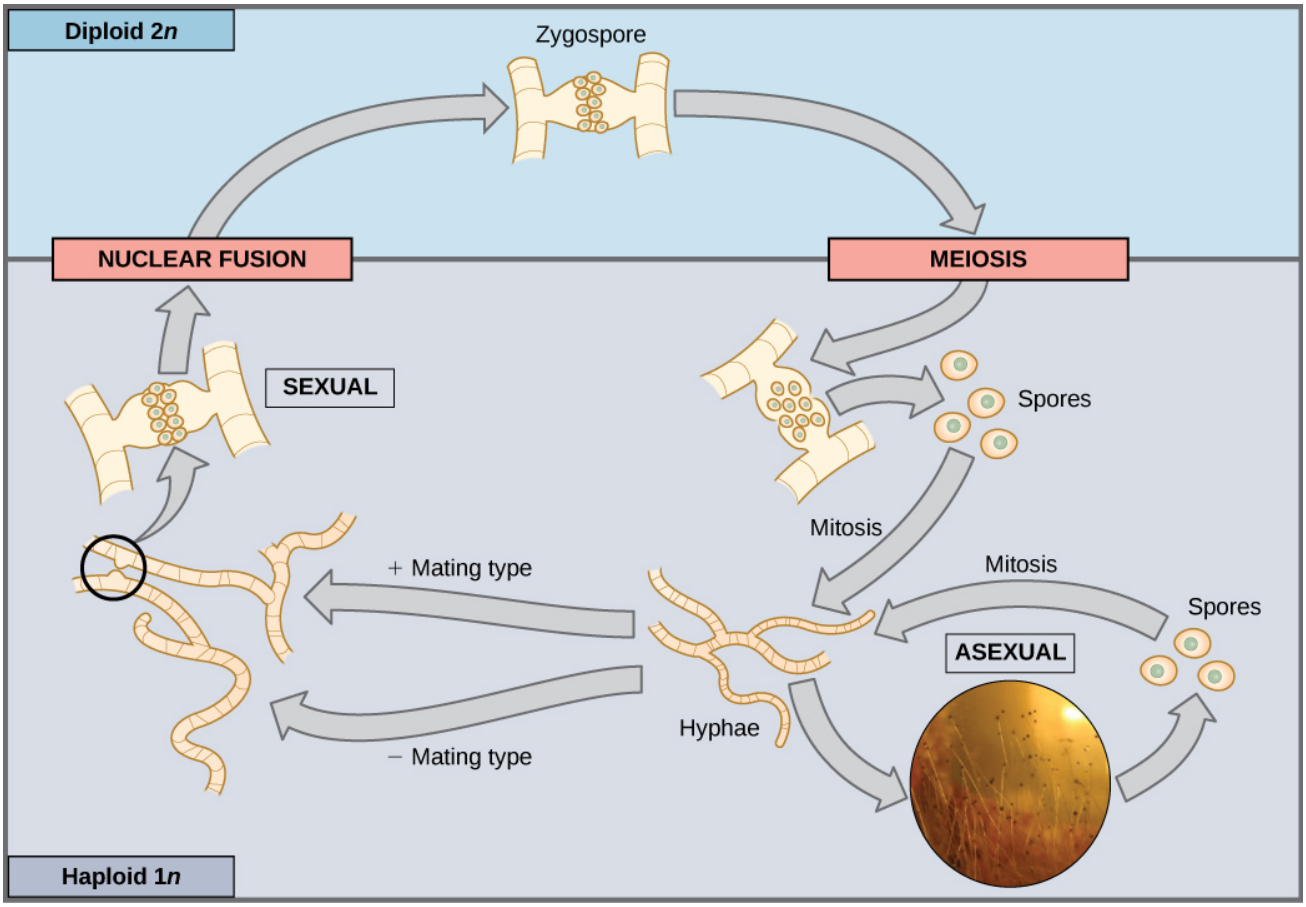


Figure 2 Fungi, such as black bread mold (Rhizopus nigricans), have haploid-dominant life cycles. The haploid multicellular stage produces specialized haploid cells by mitosis that fuse to form a diploid zygote. The zygote undergoes meiosis to produce haploid spores. Each spore gives rise to a multicellular haploid organism by mitosis. (credit "zygomycota" micrograph: modification of work by "Fanaberka"/Wikimedia Commons)

If a mutation occurs so that a fungus is no longer able to produce a minus mating type, will it still be able to reproduce?

Alternation of Generations

The third life-cycle type, employed by some algae and all plants, is a blend of the haploid-dominant and diploid-dominant extremes. Species with alternation of generations have both haploid and diploid multicellular organisms as part of their life cycle. The haploid multicellular plants are called **gametophytes**, because they produce gametes from specialized cells. Meiosis is not directly involved in the production of gametes in this case, because the organism that produces the gametes is already a haploid. Fertilization between the gametes forms a diploid zygote. The zygote will undergo many rounds of mitosis and give rise to a diploid multicellular plant called a **sporophyte**. Specialized cells of the sporophyte will undergo meiosis and produce haploid spores. The spores will subsequently develop into the gametophytes (Figure 3).

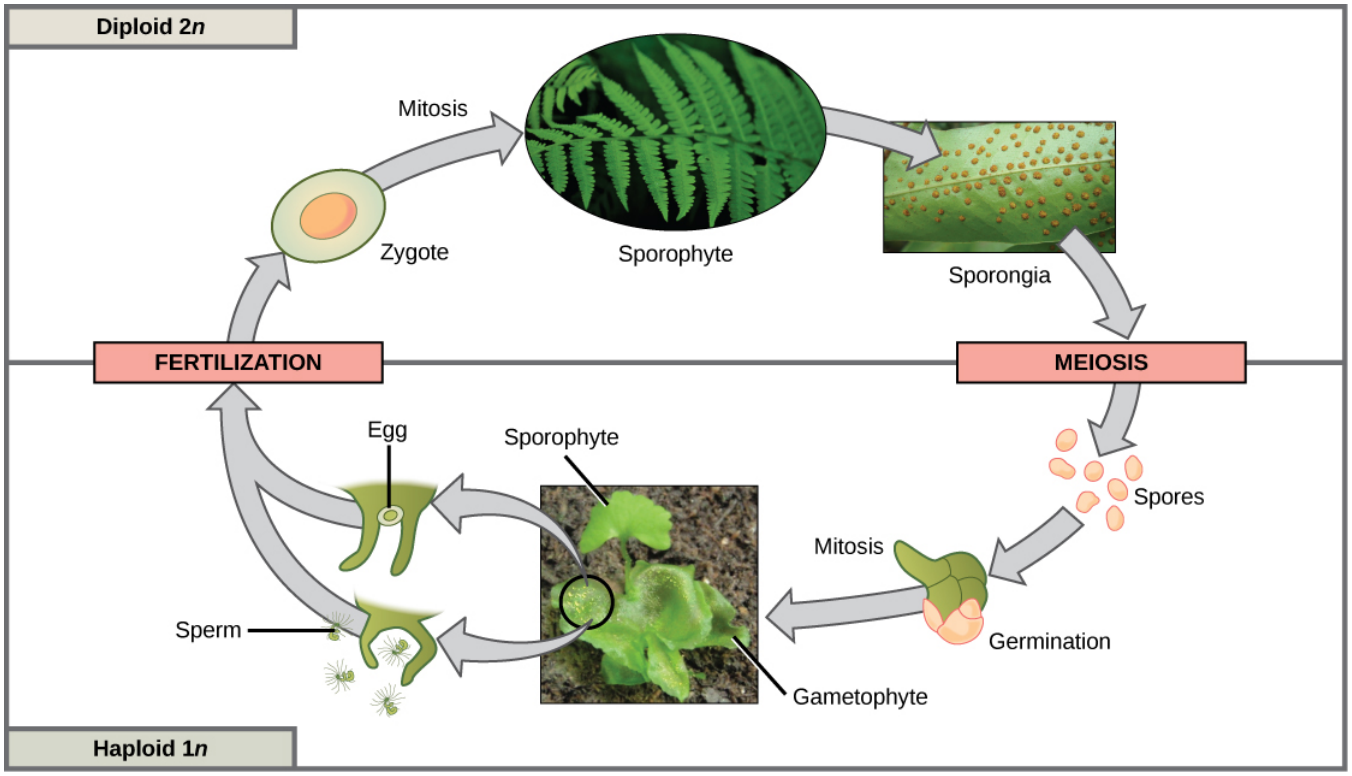


Figure 3 Plants have a life cycle that alternates between a multicellular haploid organism and a multicellular diploid organism. In some plants, such as ferns, both the haploid and diploid plant stages are free-living. The diploid plant is called a sporophyte because it produces haploid spores by meiosis. The spores develop into multicellular, haploid plants called gametophytes because they produce gametes. The gametes of two individuals will fuse to form a diploid zygote that becomes the sporophyte. (credit "fern": modification of work by Cory Zanker; credit "sporangia": modification of work by "Obsidian Soul"/Wikimedia Commons; credit "gametophyte and sporophyte": modification of work by "Vlmastra"/Wikimedia Commons)

Although all plants utilize some version of the alternation of generations, the relative size of the sporophyte and the gametophyte and the relationship between them vary greatly. In plants such as moss, the gametophyte organism is the free-living plant, and the sporophyte is physically dependent on the gametophyte. In other plants, such as ferns, both the gametophyte and sporophyte plants are free-living; however, the sporophyte is much larger. In seed plants, such as magnolia trees and daisies, the gametophyte is composed of only a few cells and, in the case of the female gametophyte, is completely retained within the sporophyte.

Sexual reproduction takes many forms in multicellular organisms. However, at some point in each type of life cycle, meiosis produces haploid cells that will fuse with the haploid cell of another organism. The mechanisms of variation—crossover, random assortment of homologous chromosomes, and random fertilization—are present in all versions of sexual reproduction. The fact that nearly every multicellular organism on Earth employs sexual reproduction is strong evidence for the benefits of producing offspring with unique gene combinations, though there are other possible benefits as well.

Section Summary

Nearly all eukaryotes undergo sexual reproduction. The variation introduced into the reproductive cells by meiosis appears to be one of the advantages of sexual reproduction that has made it so successful. Meiosis and fertilization alternate in sexual life cycles. The process of meiosis produces unique reproductive cells called gametes, which have half the number of chromosomes as the parent cell. Fertilization, the fusion of haploid gametes from two individuals, restores the diploid condition. Thus, sexually reproducing organisms alternate between haploid and diploid stages. However, the ways in which reproductive cells are produced and the timing between meiosis and fertilization vary greatly. There are three main categories of life cycles: diploid-dominant, demonstrated by most animals; haploid-dominant, demonstrated by all fungi and some algae; and the alternation of generations, demonstrated by plants and some algae.

Visit this page in your course online to check your understanding.

Additional Self Check Questions

1. If a mutation occurs so that a fungus is no longer able to produce a minus mating type, will it still be able to reproduce?

2. List and briefly describe the three processes that lead to variation in offspring with the same parents.

3. Compare the three main types of life cycles in multicellular organisms and give an example of an organism that employs each.

Answers

1. Yes, it will be able to reproduce asexually.
2. Crossover occurs in prophase I between non-sister homologous chromosomes. Segments of DNA are exchanged between maternally derived and paternally derived chromosomes, and new gene combinations are formed. b. Random alignment during metaphase I leads to gametes that have a mixture of maternal and paternal chromosomes. c. Fertilization is random, in that any two gametes can fuse.
3. In the haploid-dominant life cycle, the multicellular stage is haploid. The diploid stage is a spore that undergoes meiosis to produce cells that will divide mitotically to produce new multicellular organisms. Fungi have a haploid-dominant life cycle. b. In the diploid-dominant life cycle, the most visible or largest multicellular stage is diploid. The haploid stage is usually reduced to a single cell type, such as a gamete or spore. Animals, such as humans, have a diploid-dominant life cycle. c. In the alternation of generations life cycle, there are both haploid and diploid multicellular stages, although the haploid stage may be completely retained by the diploid stage. Plants have a life cycle with alternation of generations.

Glossary

alternation of generations: life-cycle type in which the diploid and haploid stages alternate

diploid-dominant: life-cycle type in which the multicellular diploid stage is prevalent

haploid-dominant: life-cycle type in which the multicellular haploid stage is prevalent

gametophyte: a multicellular haploid life-cycle stage that produces gametes

germ cells: specialized cell line that produces gametes, such as eggs or sperm

life cycle: the sequence of events in the development of an organism and the production of cells that produce offspring

sporophyte: a multicellular diploid life-cycle stage that produces haploid spores by meiosis

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: Open Stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License**: *CC BY: Attribution*

STUDY GUIDE: MEIOSIS

Study Questions

Objective: Explain how eukaryotic organisms sexually reproduce.

Use this page to check your understanding of the content.

Study Guide Questions

1. Carefully compare and contrast mitosis and meiosis.
2. What is the very unique thing that happens between homologous chromosomes during prophase I? Why is this important?
3. Be able to determine the number of chromosomes AND the amount of DNA in a cell during meiosis. Know when the "reduction division" occurs.
4. Given any diagram or picture of a cell in any phase of meiosis or mitosis, be able to identify the phase.
5. Where does meiosis occur?
6. Be able to provide a rationale for WHY meiosis is necessary.
7. Clearly explain how meiosis results in genetic diversity.
8. What is independent assortment? When does it occur? How does it increase genetic diversity?
9. Be able to draw a picture of independent assortment.
10. What is crossing over? How does it increase genetic diversity?
11. Be able to draw a picture of crossing over.

Licensing & Attributions

CC licensed content, Shared previously

- Study Guide: Meiosis. **Authored by**: Wendy Riggs. **Provided by**: College of the Redwoods. **Project**: Kaleidoscope. **License**: *CC BY: Attribution*

FIG: MEIOSIS

http://cnx.org/resources/bc371c7c2a8cd37190ddb5e9fba46730

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: OpenStax. **Located at**: http://cnx.org/resources/bc371c7c2a8cd37190ddb5e9fba46730. **License**: *CC BY: Attribution*

GENETICS AND INHERITANCE

INTRODUCTION (MENDEL AND HEREDITY)

Genetics is the study of heredity. Johann Gregor Mendel set the framework for genetics long before chromosomes or genes had been identified, at a time when meiosis was not well understood. Mendel selected a simple biological system and conducted methodical, quantitative analyses using large sample sizes. Because of Mendel's work, the fundamental principles of heredity were revealed. We now know that genes, carried on chromosomes, are the basic functional units of heredity with the capability to be replicated, expressed, or mutated. Today, the postulates put forth by Mendel form the basis of classical, or Mendelian, genetics. Not all genes are transmitted from parents to offspring according to Mendelian genetics, but Mendel's experiments serve as an excellent starting point for thinking about inheritance.

Figure 1. Experimenting with thousands of garden peas, Mendel uncovered the fundamentals of genetics. (credit: modification of work by Jerry Kirkhart)

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: Open Stax. **Located at**: http://cnx.org/contents/185cbf87-c72e-48f5-b51e-f14f21b5eabd@9.17:1/Biology. **License**: *CC BY: Attribution*

MENDEL'S EXPERIMENTS AND THE LAWS OF PROBABILITY

Learning Objective

By the end of this section you will be able to:,

- Describe the scientific reasons for the success of Mendel's experimental work
- Describe the expected outcomes of monohybrid crosses involving dominant and recessive alleles
- Apply the sum and product rules to calculate probabilities

Johann Gregor Mendel (1822–1884) (Figure 1) was a lifelong learner, teacher, scientist, and man of faith. As a young adult, he joined the Augustinian Abbey of St. Thomas in Brno in what is now the Czech Republic. Supported by the monastery, he taught physics, botany, and natural science courses at the secondary and university levels. In 1856, he began a decade-long research pursuit involving inheritance patterns in honeybees and plants, ultimately settling on pea plants as his primary model system (a system with convenient characteristics used to study a specific biological phenomenon to be applied to other systems). In 1865, Mendel presented the results of his experiments with nearly 30,000 pea plants to the local Natural History Society. He demonstrated that traits are transmitted faithfully from parents to offspring independently of other traits and in dominant and recessive patterns. In 1866, he published his work, *Experiments in Plant Hybridization*, ((Note: Johann Gregor Mendel, Versuche über Pflanzenhybriden Verhandlungen des naturforschenden Vereines in Brünn, Bd. IV für das Jahr, 1865 Abhandlungen, 3–47. (for English translation see http://www.mendelweb.org/Mendel.plain.html))) in the proceedings of the Natural History Society of Brünn.

Figure 1. Johann Gregor Mendel is considered the father of genetics.

Mendel's work went virtually unnoticed by the scientific community that believed, incorrectly, that the process of inheritance involved a blending of parental traits that produced an intermediate physical appearance in offspring; this hypothetical process appeared to be correct because of what we know now as continuous variation. Continuous variation results from the action of many genes to determine a characteristic like human height. Offspring appear to be a "blend" of their parents' traits when we look at characteristics that exhibit continuous variation. The blending theory of inheritance asserted that the original parental traits were lost or absorbed by the blending in the offspring, but we now know that this is not the case. Mendel was the first researcher to see it. Instead of continuous characteristics, Mendel worked with traits that were inherited in distinct classes (specifically, violet versus white flowers); this is referred to as discontinuous variation. Mendel's choice of these kinds of traits allowed him to see experimentally that the traits were not blended in the offspring, nor were they absorbed, but rather that they kept their distinctness and could be passed on. In 1868, Mendel became abbot of the monastery and exchanged his scientific pursuits for his pastoral duties. He was not recognized for his extraordinary scientific contributions during his lifetime. In fact, it was not until 1900 that his work was rediscovered, reproduced, and revitalized by scientists on the brink of discovering the chromosomal basis of heredity.

Mendel's Model System

Mendel's seminal work was accomplished using the garden pea, *Pisum sativum*, to study inheritance. This species naturally self-fertilizes, such that pollen encounters ova within individual flowers. The flower petals remain sealed tightly until after pollination, preventing pollination from other plants. The result is highly inbred, or "true-breeding," pea plants. These are plants that always produce offspring that look like the parent. By experimenting with true-breeding pea plants, Mendel avoided the appearance of unexpected traits in offspring that might occur if the plants were not true breeding. The garden pea also grows to maturity within one season, meaning that several generations could be evaluated over a relatively short time. Finally, large quantities of garden peas could be cultivated simultaneously, allowing Mendel to conclude that his results did not come about simply by chance.

Mendelian Crosses

Mendel performed hybridizations, which involve mating two true-breeding individuals that have different traits. In the pea, which is naturally self-pollinating, this is done by manually transferring pollen from the anther of a mature pea plant of one variety to the stigma of a separate mature pea plant of the second variety. In plants, pollen carries the male gametes (sperm) to the stigma, a sticky organ that traps pollen and allows the sperm to move down the pistil to the female gametes (ova) below. To prevent the pea plant that was receiving pollen from self-fertilizing and confounding his results, Mendel painstakingly removed all of the anthers from the plant's flowers before they had a chance to mature.

Plants used in first-generation crosses were called P_0, or parental generation one, plants (Figure). Mendel collected the seeds belonging to the P_0 plants that resulted from each cross and grew them the following season. These offspring were called the F_1, or the first filial (*filial* = offspring, daughter or son), generation. Once Mendel examined the characteristics in the F_1 generation of plants, he allowed them to self-fertilize naturally. He then collected and grew the seeds from the F_1 plants to produce the F_2, or second filial, generation. Mendel's experiments extended beyond the F_2 generation to the F_3 and F_4 generations, and so on, but it was the ratio of characteristics in the P_0–F_1–F_2 generations that were the most intriguing and became the basis for Mendel's postulates.

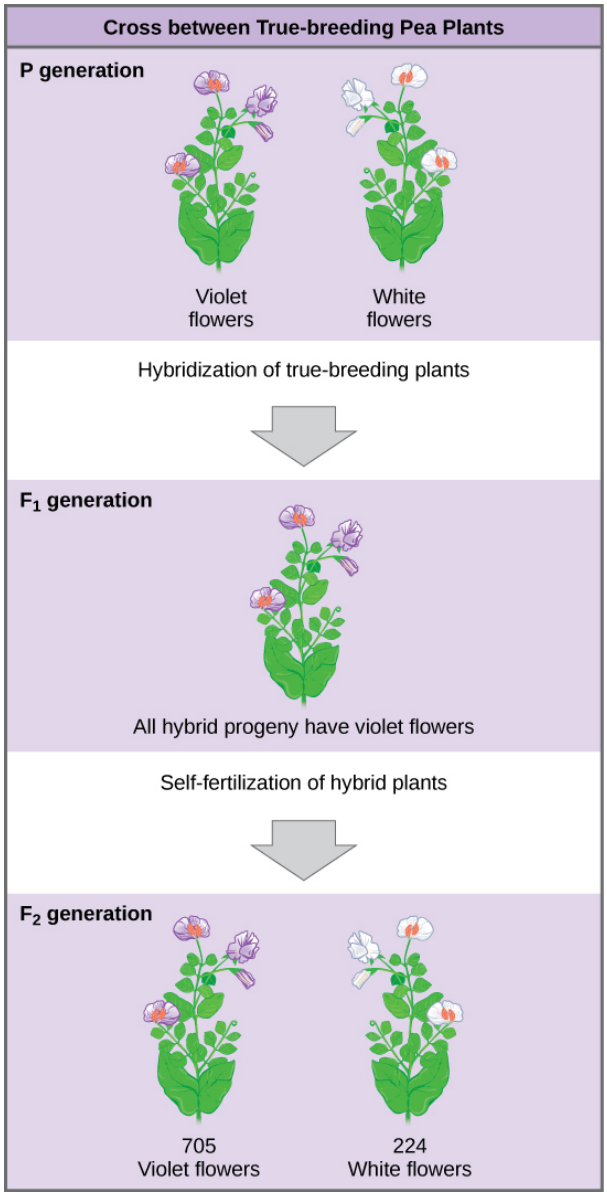


Figure 2. In one of his experiments on inheritance patterns, Mendel crossed plants that were true-breeding for violet flower color with plants true-breeding for white flower color (the P generation). The resulting hybrids in the F1 generation all had violet flowers. In the F2 generation, approximately three quarters of the plants had violet flowers, and one quarter had white flowers.

Garden Pea Characteristics Revealed the Basics of Heredity

In his 1865 publication, Mendel reported the results of his crosses involving seven different characteristics, each with two contrasting traits. A trait is defined as a variation in the physical appearance of a heritable characteristic. The characteristics included plant height, seed texture, seed color, flower color, pea pod size, pea pod color, and flower position. For the characteristic of flower color, for example, the two contrasting traits were white versus violet. To fully examine each characteristic, Mendel generated large numbers of F_1 and F_2 plants, reporting results from 19,959 F_2 plants alone. His findings were consistent.

What results did Mendel find in his crosses for flower color? First, Mendel confirmed that he had plants that bred true for white or violet flower color. Regardless of how many generations Mendel examined, all self-crossed offspring of parents with white flowers had white flowers, and all self-crossed offspring of parents with violet flowers had violet flowers. In addition, Mendel confirmed that, other than flower color, the pea plants were physically identical.

Once these validations were complete, Mendel applied the pollen from a plant with violet flowers to the stigma of a plant with white flowers. After gathering and sowing the seeds that resulted from this cross, Mendel found that 100 percent of the F_1 hybrid generation had violet flowers. Conventional wisdom at that time would have predicted the hybrid flowers to be pale violet or for hybrid plants to have equal numbers of white and violet flowers. In other words, the contrasting parental traits were expected to blend in the offspring. Instead, Mendel's results demonstrated that the white flower trait in the F_1 generation had completely disappeared.

Importantly, Mendel did not stop his experimentation there. He allowed the F_1 plants to self-fertilize and found that, of F_2-generation plants, 705 had violet flowers and 224 had white flowers. This was a ratio of 3.15 violet flowers per one white flower, or approximately 3:1. When Mendel transferred pollen from a plant with violet flowers to the stigma of a plant with white flowers and vice versa, he obtained about the same ratio regardless of which parent, male or female, contributed which trait. This is called a reciprocal cross—a paired cross in which the respective traits of the male and female in one cross become the respective traits of the female and male in the other cross. For the other six characteristics Mendel examined, the F_1 and F_2 generations behaved in the same way as they had for flower color. One of the two traits would disappear completely from the F_1 generation only to reappear in the F_2 generation at a ratio of approximately 3:1 (Table 1).

Table 1. The Results of Mendel's Garden Pea Hybridizations

Characteristic	Contrasting P_0 Traits	F_1 Offspring Traits	F_2 Offspring Traits	F_2 Trait Ratios
Flower color	Violet vs. white	100 percent violet	• 705 violet • 224 white	3.15:1
Flower position	Axial vs. terminal	100 percent axial	• 651 axial • 207 terminal	3.14:1
Plant height	Tall vs. dwarf	100 percent tall	• 787 tall • 277 dwarf	2.84:1
Seed texture	Round vs. wrinkled	100 percent round	• 5,474 round • 1,850 wrinkled	2.96:1
Seed color	Yellow vs. green	100 percent yellow	• 6,022 yellow • 2,001 green	3.01:1
Pea pod texture	Inflated vs. constricted	100 percent inflated	• 882 inflated • 299 constricted	2.95:1
Pea pod color	Green vs. yellow	100 percent green	• 428 green • 152 yellow	2.82:1

Upon compiling his results for many thousands of plants, Mendel concluded that the characteristics could be divided into expressed and latent traits. He called these, respectively, dominant and recessive traits. Dominant traits are those that are inherited unchanged in a hybridization. Recessive traits become latent, or disappear, in the offspring of a hybridization. The recessive trait does, however, reappear in the progeny of the hybrid offspring. An example of a dominant trait is the violet-flower trait. For this same characteristic (flower color), white-colored flowers are a recessive trait. The fact that the recessive trait reappeared in the F_2 generation meant that the traits remained separate (not blended) in the plants of the F_1 generation. Mendel also proposed that plants possessed two copies of the trait for the flower-color characteristic, and that each parent transmitted one of its two copies to its offspring, where they came together. Moreover, the physical observation of a dominant trait could mean that the genetic composition of the organism included two dominant versions of the characteristic or that it included one dominant and one recessive version. Conversely, the observation of a recessive trait meant that the organism lacked any dominant versions of this characteristic.

So why did Mendel repeatedly obtain 3:1 ratios in his crosses? To understand how Mendel deduced the basic mechanisms of inheritance that lead to such ratios, we must first review the laws of probability.

Probability Basics

Probabilities are mathematical measures of likelihood. The empirical probability of an event is calculated by dividing the number of times the event occurs by the total number of opportunities for the event to occur. It is also possible to calculate theoretical probabilities by dividing the number of times that an event is expected to occur by the number of times that it could occur. Empirical probabilities come from observations, like those of Mendel. Theoretical probabilities come from knowing how the events are produced and assuming that the probabilities of individual outcomes are equal. A probability of one for some event indicates that it is guaranteed to occur, whereas a probability of zero indicates that it is guaranteed not to occur. An example of a genetic event is a round seed produced by a pea plant. In his experiment, Mendel demonstrated that the probability of the event “round seed” occurring was one in the F_1 offspring of true-breeding parents, one of which has round seeds and one of which has wrinkled seeds. When the F_1 plants were subsequently self-crossed, the probability of any given F_2 offspring having round seeds was now three out of four. In other words, in a large population of F_2 offspring chosen at random, 75 percent were expected to have round seeds, whereas 25 percent were expected to have wrinkled seeds. Using large numbers of crosses, Mendel was able to calculate probabilities and use these to predict the outcomes of other crosses.

The Product Rule and Sum Rule

Mendel demonstrated that the pea-plant characteristics he studied were transmitted as discrete units from parent to offspring. As will be discussed, Mendel also determined that different characteristics, like seed color and seed texture, were transmitted independently of one another and could be considered in separate probability analyses. For instance, performing a cross between a plant with green, wrinkled seeds and a plant with yellow, round seeds still produced offspring that had a 3:1 ratio of green:yellow seeds (ignoring seed texture) and a 3:1 ratio of round:wrinkled seeds (ignoring seed color). The characteristics of color and texture did not influence each other.

The product rule of probability can be applied to this phenomenon of the independent transmission of characteristics. The product rule states that the probability of two independent events occurring together can be calculated by multiplying the individual probabilities of each event occurring alone. To demonstrate the product rule, imagine that you are rolling a six-sided die (D) and flipping a penny (P) at the same time. The die may roll any number from 1–6 ($D_{\#}$), whereas the penny may turn up heads (P_H) or tails (P_T). The outcome of rolling the die has no effect on the outcome of flipping the penny and vice versa. There are 12 possible outcomes of this action (Table 2), and each event is expected to occur with equal probability.

Table 2. Twelve Equally Likely Outcomes of Rolling a Die and Flipping a Penny

Rolling Die	Flipping Penny
D_1	P_H

Table 2. Twelve Equally Likely Outcomes of Rolling a Die and Flipping a Penny	
Rolling Die	Flipping Penny
D_1	P_T
D_2	P_H
D_2	P_T
D_3	P_H
D_3	P_T
D_4	P_H
D_4	P_T
D_5	P_H
D_5	P_T
D_6	P_H
D_6	P_T

Of the 12 possible outcomes, the die has a 2/12 (or 1/6) probability of rolling a two, and the penny has a 6/12 (or 1/2) probability of coming up heads. By the product rule, the probability that you will obtain the combined outcome 2 and heads is: $(D_2) \times (P_H) = (1/6) \times (1/2)$ or 1/12 (Table). Notice the word “and” in the description of the probability. The “and” is a signal to apply the product rule. For example, consider how the product rule is applied to the dihybrid cross: the probability of having both dominant traits in the F_2 progeny is the product of the probabilities of having the dominant trait for each characteristic, as shown here:
34/ × 34/ = 916/

On the other hand, the sum rule of probability is applied when considering two mutually exclusive outcomes that can come about by more than one pathway. The sum rule states that the probability of the occurrence of one event or the other event, of two mutually exclusive events, is the sum of their individual probabilities. Notice the word “or” in the description of the probability. The “or” indicates that you should apply the sum rule. In this case, let’s imagine you are flipping a penny (P) and a quarter (Q). What is the probability of one coin coming up heads and one coin coming up tails? This outcome can be achieved by two cases: the penny may be heads (P_H) and the quarter may be tails (Q_T), or the quarter may be heads (Q_H) and the penny may be tails (P_T). Either case fulfills the outcome. By the sum rule, we calculate the probability of obtaining one head and one tail as $[(P_H) \times (Q_T)] + [(Q_H) \times (P_T)] = [(1/2) \times (1/2)] + [(1/2) \times (1/2)] = 1/2$ (Table). You should also notice that we used the product rule to calculate the probability of P_H and Q_T, and also the probability of P_T and Q_H, before we summed them. Again, the sum rule can be applied to show the probability of having just one dominant trait in the F_2 generation of a dihybrid cross:
316/ + 34/ = 1516/

Table 3. The Product Rule and Sum Rule	
Product Rule	Sum Rule
For independent events A and B, the probability (P) of them both occurring (A *and* B) is ($P_A \times P_B$)	For mutually exclusive events A and B, the probability (P) that at least one occurs (A *or* B) is ($P_A + P_B$)

To use probability laws in practice, it is necessary to work with large sample sizes because small sample sizes are prone to deviations caused by chance. The large quantities of pea plants that Mendel examined allowed him calculate the probabilities of the traits appearing in his F_2 generation. As you will learn, this discovery meant that when parental traits were known, the offspring's traits could be predicted accurately even before fertilization.

Section Summary

Working with garden pea plants, Mendel found that crosses between parents that differed by one trait produced F_1 offspring that all expressed the traits of one parent. Observable traits are referred to as dominant, and non-expressed traits are described as recessive. When the offspring in Mendel's experiment were self-crossed, the F_2 offspring exhibited the dominant trait or the recessive trait in a 3:1 ratio, confirming that the recessive trait had been transmitted faithfully from the original P_0 parent. Reciprocal crosses generated identical F_1 and F_2 offspring ratios. By examining sample sizes, Mendel showed that his crosses behaved reproducibly according to the laws of probability, and that the traits were inherited as independent events.

Two rules in probability can be used to find the expected proportions of offspring of different traits from different crosses. To find the probability of two or more independent events occurring together, apply the product rule and multiply the probabilities of the individual events. The use of the word "and" suggests the appropriate application of the product rule. To find the probability of two or more events occurring in combination, apply the sum rule and add their individual probabilities together. The use of the word "or" suggests the appropriate application of the sum rule.

Visit this page in your course online to check your understanding.

Additional Self Check Questions

1. Describe one of the reasons why the garden pea was an excellent choice of model system for studying inheritance.

2. How would you perform a reciprocal cross for the characteristic of stem height in the garden pea?

Answers

1. The garden pea is sessile and has flowers that close tightly during self-pollination. These features help to prevent accidental or unintentional fertilizations that could have diminished the accuracy of Mendel's data.

2. Two sets of P_0 parents would be used. In the first cross, pollen would be transferred from a true-breeding tall plant to the stigma of a true-breeding dwarf plant. In the second cross, pollen would be transferred from a true-breeding dwarf plant to the stigma of a true-breeding tall plant. For each cross, F_1 and F_2 offspring would be analyzed to determine if offspring traits were affected according to which parent donated each trait.

Glossary

blending theory of inheritance: hypothetical inheritance pattern in which parental traits are blended together in the offspring to produce an intermediate physical appearance

continuous variation: inheritance pattern in which a character shows a range of trait values with small gradations rather than large gaps between them

discontinuous variation: inheritance pattern in which traits are distinct and are transmitted independently of one another

dominant: trait which confers the same physical appearance whether an individual has two copies of the trait or one copy of the dominant trait and one copy of the recessive trait

F_1: first filial generation in a cross; the offspring of the parental generation

F_2: second filial generation produced when F_1 individuals are self-crossed or fertilized with each other

hybridization: process of mating two individuals that differ with the goal of achieving a certain characteristic in their offspring

model system: species or biological system used to study a specific biological phenomenon to be applied to other different species

P_0: parental generation in a cross

product rule: probability of two independent events occurring simultaneously can be calculated by multiplying the individual probabilities of each event occurring alone

recessive: trait that appears "latent" or non-expressed when the individual also carries a dominant trait for that same characteristic; when present as two identical copies, the recessive trait is expressed

reciprocal cross: paired cross in which the respective traits of the male and female in one cross become the respective traits of the female and male in the other cross

sum rule: probability of the occurrence of at least one of two mutually exclusive events is the sum of their individual probabilities

trait: variation in the physical appearance of a heritable characteristic

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: Open Stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License**: *CC BY: Attribution*

TRAITS AND CHARACTERISTICS

Learning Objectives

By the end of this section, you will be able to:

- Explain the relationship between genotypes and phenotypes in dominant and recessive gene systems
- Develop a Punnett square to calculate the expected proportions of genotypes and phenotypes in a monohybrid cross
- Explain the purpose and methods of a test cross
- Identify non-Mendelian inheritance patterns such as incomplete dominance, codominance, recessive lethals, multiple alleles, and sex linkage

The seven characteristics that Mendel evaluated in his pea plants were each expressed as one of two versions, or traits. The physical expression of characteristics is accomplished through the expression of genes carried on chromosomes. The genetic makeup of peas consists of two similar or homologous copies of each chromosome, one from each parent. Each pair of homologous chromosomes has the same linear order of genes. In other

words, peas are diploid organisms in that they have two copies of each chromosome. The same is true for many other plants and for virtually all animals. Diploid organisms utilize meiosis to produce haploid gametes, which contain one copy of each homologous chromosome that unite at fertilization to create a diploid zygote.

For cases in which a single gene controls a single characteristic, a diploid organism has two genetic copies that may or may not encode the same version of that characteristic. Gene variants that arise by mutation and exist at the same relative locations on homologous chromosomes are called **alleles**. Mendel examined the inheritance of genes with just two allele forms, but it is common to encounter more than two alleles for any given gene in a natural population.

Phenotypes and Genotypes

Two alleles for a given gene in a diploid organism are expressed and interact to produce physical characteristics. The observable traits expressed by an organism are referred to as its **phenotype**. An organism's underlying genetic makeup, consisting of both physically visible and non-expressed alleles, is called its **genotype**. Mendel's hybridization experiments demonstrate the difference between phenotype and genotype. When true-breeding plants in which one parent had yellow pods and one had green pods were cross-fertilized, all of the F_1 hybrid offspring had yellow pods. That is, the hybrid offspring were phenotypically identical to the true-breeding parent with yellow pods. However, we know that the allele donated by the parent with green pods was not simply lost because it reappeared in some of the F_2 offspring. Therefore, the F_1 plants must have been genotypically different from the parent with yellow pods.

The P_1 plants that Mendel used in his experiments were each homozygous for the trait he was studying. Diploid organisms that are **homozygous** at a given gene, or locus, have two identical alleles for that gene on their homologous chromosomes. Mendel's parental pea plants always bred true because both of the gametes produced carried the same trait. When P_1 plants with contrasting traits were cross-fertilized, all of the offspring were **heterozygous** for the contrasting trait, meaning that their genotype reflected that they had different alleles for the gene being examined.

Dominant and Recessive Alleles

Our discussion of homozygous and heterozygous organisms brings us to why the F_1 heterozygous offspring were identical to one of the parents, rather than expressing both alleles. In all seven pea-plant characteristics, one of the two contrasting alleles was dominant, and the other was recessive. Mendel called the dominant allele the expressed unit factor; the recessive allele was referred to as the latent unit factor. We now know that these so-called unit factors are actually genes on homologous chromosome pairs. For a gene that is expressed in a dominant and recessive pattern, homozygous dominant and heterozygous organisms will look identical (that is, they will have different genotypes but the same phenotype). The recessive allele will only be observed in homozygous recessive individuals (Table 1).

Table 1. Human Inheritance in Dominant and Recessive Patterns

Dominant Traits	Recessive Traits
Achondroplasia	Albinism
Brachydactyly	Cystic fibrosis
Huntington's disease	Duchenne muscular dystrophy
Marfan syndrome	Galactosemia
Neurofibromatosis	Phenylketonuria
Widow's peak	Sickle-cell anemia

Table 1. Human Inheritance in Dominant and Recessive Patterns	
Dominant Traits	Recessive Traits
Wooly hair	Tay-Sachs disease

Several conventions exist for referring to genes and alleles. For the purposes of this chapter, we will abbreviate genes using the first letter of the gene's corresponding dominant trait. For example, violet is the dominant trait for a pea plant's flower color, so the flower-color gene would be abbreviated as *V* (note that it is customary to italicize gene designations). Furthermore, we will use uppercase and lowercase letters to represent dominant and recessive alleles, respectively. Therefore, we would refer to the genotype of a homozygous dominant pea plant with violet flowers as *VV*, a homozygous recessive pea plant with white flowers as *vv*, and a heterozygous pea plant with violet flowers as *Vv*.

The Punnett Square Approach for a Monohybrid Cross

When fertilization occurs between two true-breeding parents that differ in only one characteristic, the process is called a **monohybrid** cross, and the resulting offspring are monohybrids. Mendel performed seven monohybrid crosses involving contrasting traits for each characteristic. On the basis of his results in F_1 and F_2 generations, Mendel postulated that each parent in the monohybrid cross contributed one of two paired unit factors to each offspring, and every possible combination of unit factors was equally likely.

To demonstrate a monohybrid cross, consider the case of true-breeding pea plants with yellow versus green pea seeds. The dominant seed color is yellow; therefore, the parental genotypes were *YY* for the plants with yellow seeds and *yy* for the plants with green seeds, respectively. A **Punnett square**, devised by the British geneticist Reginald Punnett, can be drawn that applies the rules of probability to predict the possible outcomes of a genetic cross or mating and their expected frequencies. To prepare a Punnett square, all possible combinations of the parental alleles are listed along the top (for one parent) and side (for the other parent) of a grid, representing their meiotic segregation into haploid gametes. Then the combinations of egg and sperm are made in the boxes in the table to show which alleles are combining. Each box then represents the diploid genotype of a zygote, or fertilized egg, that could result from this mating. Because each possibility is equally likely, genotypic ratios can be determined from a Punnett square. If the pattern of inheritance (dominant or recessive) is known, the phenotypic ratios can be inferred as well. For a monohybrid cross of two true-breeding parents, each parent contributes one type of allele. In this case, only one genotype is possible. All offspring are *Yy* and have yellow seeds (Figure 1).

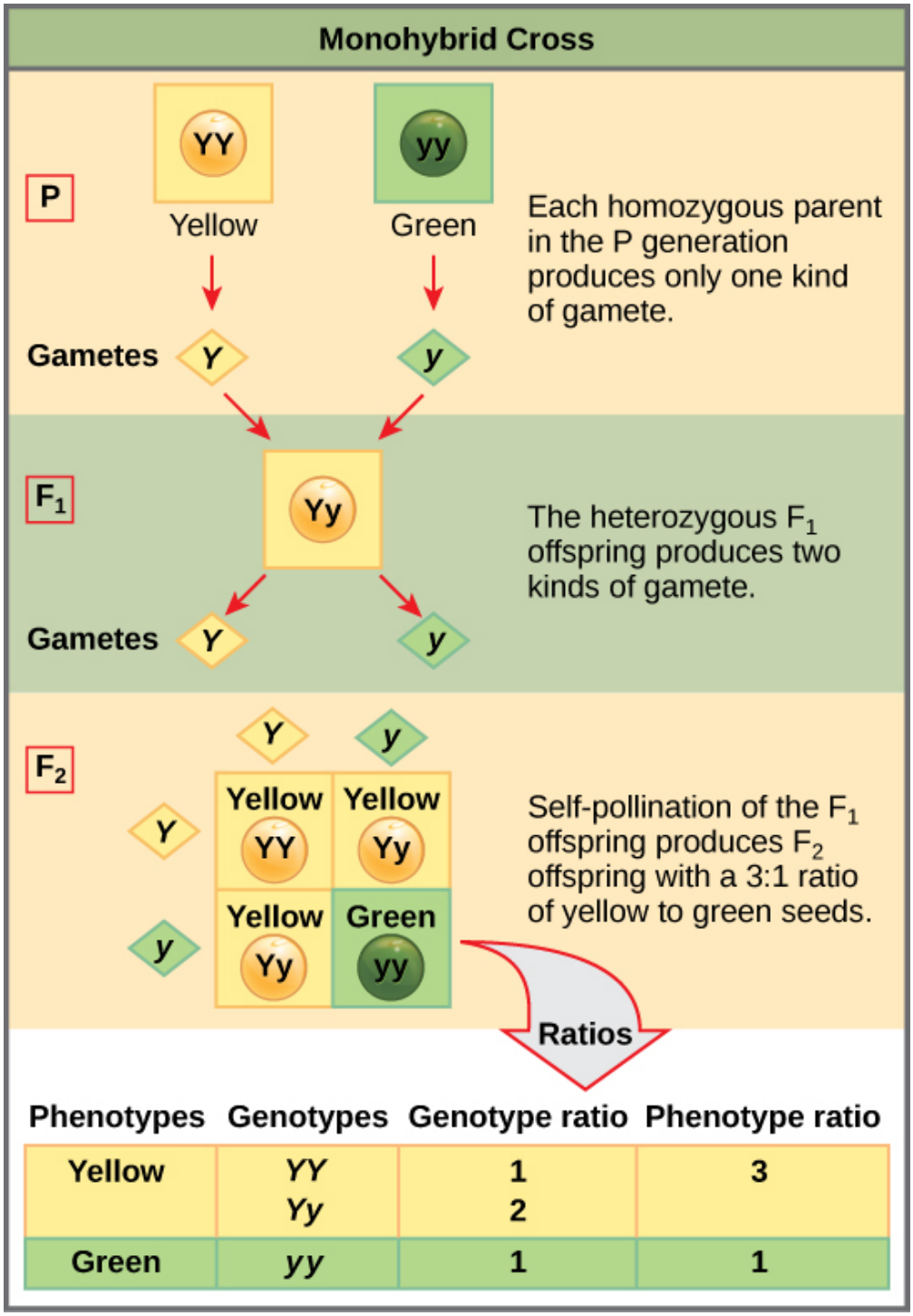


Figure 1. In the P generation, pea plants that are true-breeding for the dominant yellow phenotype are crossed with plants with the recessive green phenotype. This cross produces F1 heterozygotes with a yellow phenotype. Punnett square analysis can be used to predict the genotypes of the F2 generation.

A self-cross of one of the *Yy* heterozygous offspring can be represented in a 2 × 2 Punnett square because each parent can donate one of two different alleles. Therefore, the offspring can potentially have one of four allele combinations: *YY*, *Yy*, *yY*, or *yy* (Figure 1). Notice that there are two ways to obtain the *Yy* genotype: a *Y* from the egg and a *y* from the sperm, or a *y* from the egg and a *Y* from the sperm. Both of these possibilities must be counted. Recall that Mendel's pea-plant characteristics behaved in the same way in reciprocal crosses. Therefore, the two possible heterozygous combinations produce offspring that are genotypically and phenotypically identical despite their dominant and recessive alleles deriving from different parents. They are

grouped together. Because fertilization is a random event, we expect each combination to be equally likely and for the offspring to exhibit a ratio of *YY*:*Yy*:*yy* genotypes of 1:2:1 (Figure 1). Furthermore, because the *YY* and *Yy* offspring have yellow seeds and are phenotypically identical, applying the sum rule of probability, we expect the offspring to exhibit a phenotypic ratio of 3 yellow:1 green. Indeed, working with large sample sizes, Mendel observed approximately this ratio in every F_2 generation resulting from crosses for individual traits.

Mendel validated these results by performing an F_3 cross in which he self-crossed the dominant- and recessive-expressing F_2 plants. When he self-crossed the plants expressing green seeds, all of the offspring had green seeds, confirming that all green seeds had homozygous genotypes of *yy*. When he self-crossed the F_2 plants expressing yellow seeds, he found that one-third of the plants bred true, and two-thirds of the plants segregated at a 3:1 ratio of yellow:green seeds. In this case, the true-breeding plants had homozygous (*YY*) genotypes, whereas the segregating plants corresponded to the heterozygous (*Yy*) genotype. When these plants self-fertilized, the outcome was just like the F_1 self-fertilizing cross.

The Test Cross Distinguishes the Dominant Phenotype

Beyond predicting the offspring of a cross between known homozygous or heterozygous parents, Mendel also developed a way to determine whether an organism that expressed a dominant trait was a heterozygote or a homozygote. Called the **test cross**, this technique is still used by plant and animal breeders. In a test cross, the dominant-expressing organism is crossed with an organism that is homozygous recessive for the same characteristic. If the dominant-expressing organism is a homozygote, then all F_1 offspring will be heterozygotes expressing the dominant trait (Figure 2). Alternatively, if the dominant expressing organism is a heterozygote, the F_1 offspring will exhibit a 1:1 ratio of heterozygotes and recessive homozygotes (Figure 2). The test cross further validates Mendel's postulate that pairs of unit factors segregate equally.

Art Connection

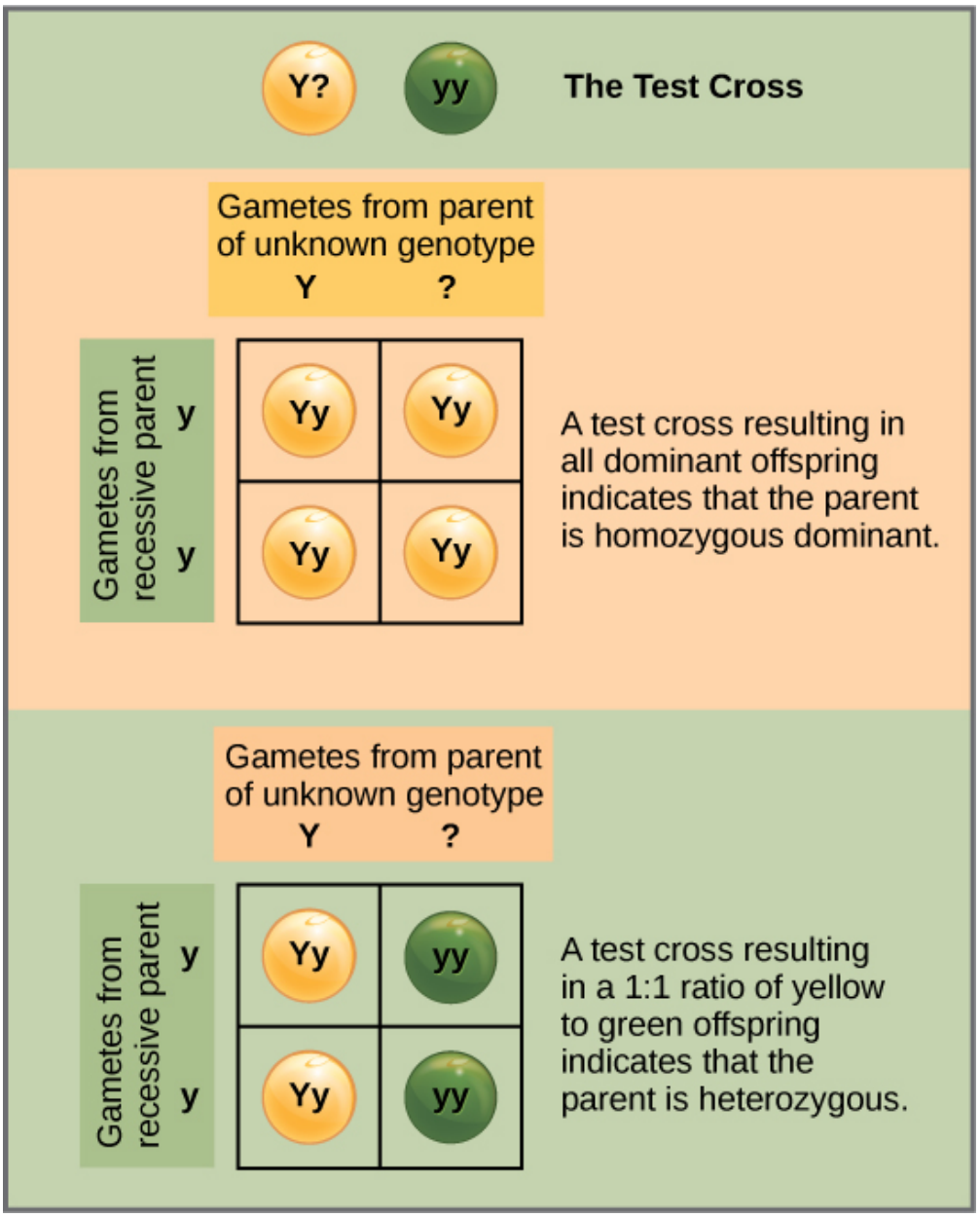


Figure 2. A test cross can be performed to determine whether an organism expressing a dominant trait is a homozygote or a heterozygote.

In pea plants, round peas (*R*) are dominant to wrinkled peas (*r*). You do a test cross between a pea plant with wrinkled peas (genotype *rr*) and a plant of unknown genotype that has round peas. You end up with three plants, all which have round peas. From this data, can you tell if the round pea parent plant is homozygous dominant or heterozygous? If the round pea parent plant is heterozygous, what is the probability that a random sample of 3 progeny peas will all be round?

Many human diseases are genetically inherited. A healthy person in a family in which some members suffer from a recessive genetic disorder may want to know if he or she has the disease-causing gene and what risk exists of passing the disorder on to his or her offspring. Of course, doing a test cross in humans is unethical and impractical. Instead, geneticists use **pedigree analysis** to study the inheritance pattern of human genetic diseases (Figure 3).

Art Connection

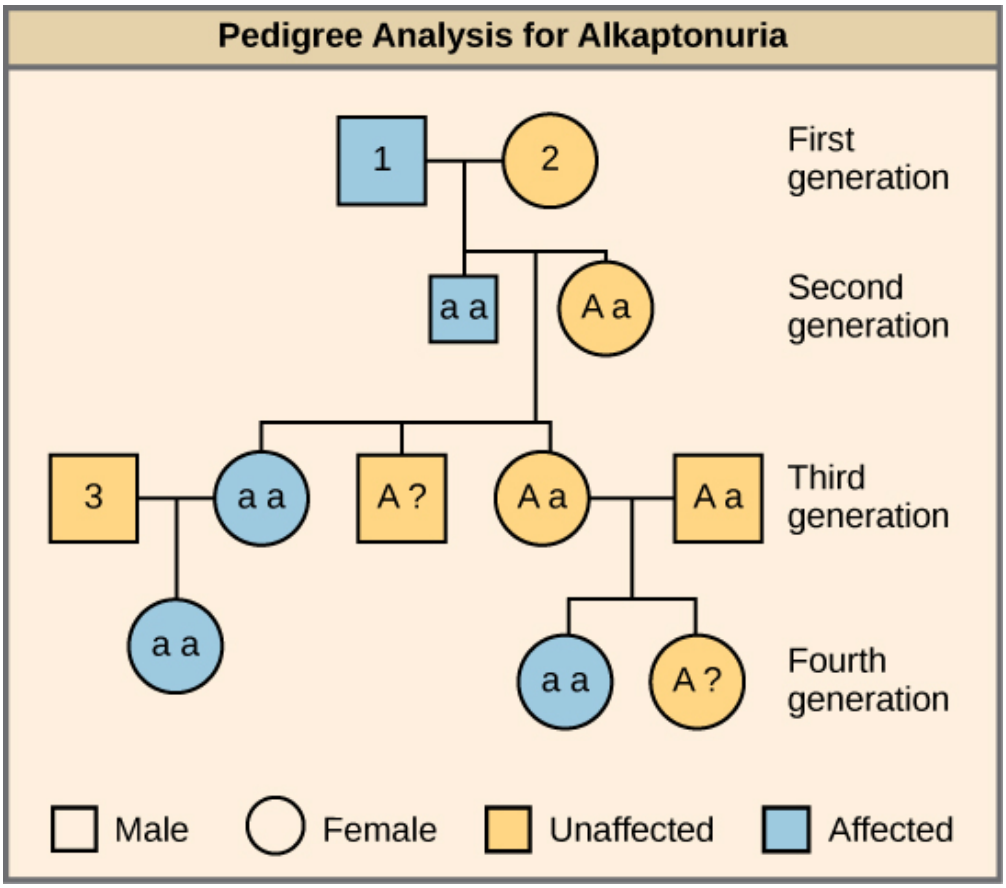


Figure 3. In this pedigree analysis for alkaptonuria, individuals with the disorder are indicated in blue and have the genotype aa. Unaffected individuals are indicated in yellow and have the genotype AA or Aa.

Alkaptonuria is a recessive genetic disorder in which two amino acids, phenylalanine and tyrosine, are not properly metabolized. Affected individuals may have darkened skin and brown urine, and may suffer joint damage and other complications. Looking at Figure 3, we can see that it is often possible to determine a person's genotype from the genotype of their offspring. For example, if neither parent has the disorder but their child does, they must be heterozygous. Two individuals on the pedigree have an unaffected phenotype but unknown genotype. Because they do not have the disorder, they must have at least one normal allele, so their genotype gets the "*A?*" designation.

What are the genotypes of the individuals labeled 1, 2 and 3?

Alternatives to Dominance and Recessiveness

Mendel's experiments with pea plants suggested that: (1) two "units" or alleles exist for every gene; (2) alleles maintain their integrity in each generation (no blending); and (3) in the presence of the dominant allele, the recessive allele is hidden and makes no contribution to the phenotype. Therefore, recessive alleles can be "carried" and not expressed by individuals. Such heterozygous individuals are sometimes referred to as "carriers." Further genetic studies in other plants and animals have shown that much more complexity exists, but that the fundamental principles of Mendelian genetics still hold true. In the sections to follow, we consider some of the extensions of Mendelism. If Mendel had chosen an experimental system that exhibited these genetic complexities, it's possible that he would not have understood what his results meant.

Incomplete Dominance

Mendel's results, that traits are inherited as dominant and recessive pairs, contradicted the view at that time that offspring exhibited a blend of their parents' traits. However, the heterozygote phenotype occasionally does appear to be intermediate between the two parents. For example, in the snapdragon, *Antirrhinum majus* (Figure), a cross between a homozygous parent with white flowers (C^WC^W) and a homozygous parent with red flowers (C^RC^R) will produce offspring with pink flowers (C^RC^W). (Note that different genotypic abbreviations are used for Mendelian extensions to distinguish these patterns from simple dominance and recessiveness.) This pattern of inheritance is described as **incomplete dominance**, denoting the expression of two contrasting alleles such that the individual displays an intermediate phenotype. The allele for red flowers is incompletely dominant over the allele for white flowers. However, the results of a heterozygote self-cross can still be predicted, just as with Mendelian dominant and recessive crosses. In this case, the genotypic ratio would be 1 C^RC^R:2 C^RC^W:1 C^WC^W, and the phenotypic ratio would be 1:2:1 for red:pink:white.

Figure 4. These pink flowers of a heterozygote snapdragon result from incomplete dominance. (credit: "storebukkebruse"/Flickr)

Codominance

A variation on incomplete dominance is **codominance**, in which both alleles for the same characteristic are simultaneously expressed in the heterozygote. An example of codominance is the MN blood groups of humans. The M and N alleles are expressed in the form of an M or N antigen present on the surface of red blood cells. Homozygotes (L^ML^M and L^NL^N) express either the M or the N allele, and heterozygotes (L^ML^N) express both alleles equally. In a self-cross between heterozygotes expressing a codominant trait, the three possible offspring genotypes are phenotypically distinct. However, the 1:2:1 genotypic ratio characteristic of a Mendelian monohybrid cross still applies.

Multiple Alleles

Mendel implied that only two alleles, one dominant and one recessive, could exist for a given gene. We now know that this is an oversimplification. Although individual humans (and all diploid organisms) can only have two alleles for a given gene, multiple alleles may exist at the population level such that many combinations of two alleles are observed. Note that when many alleles exist for the same gene, the convention is to denote the most common phenotype or genotype among wild animals as the **wild type** (often abbreviated "+"); this is considered the standard or norm. All other phenotypes or genotypes are considered **variants** of this standard, meaning that they deviate from the wild type. The variant may be recessive or dominant to the wild-type allele.

An example of multiple alleles is coat color in rabbits (Figure 5). Here, four alleles exist for the *c* gene. The wild-type version, C^+C^+, is expressed as brown fur. The chinchilla phenotype, $c^{ch}c^{ch}$, is expressed as black-tipped white fur. The Himalayan phenotype, c^hc^h, has black fur on the extremities and white fur elsewhere. Finally, the albino, or "colorless" phenotype, *cc*, is expressed as white fur. In cases of multiple alleles, dominance hierarchies can exist. In this case, the wild-type allele is dominant over all the others, chinchilla is incompletely dominant over Himalayan and albino, and Himalayan is dominant over albino. This hierarchy, or allelic series, was revealed by observing the phenotypes of each possible heterozygote offspring.

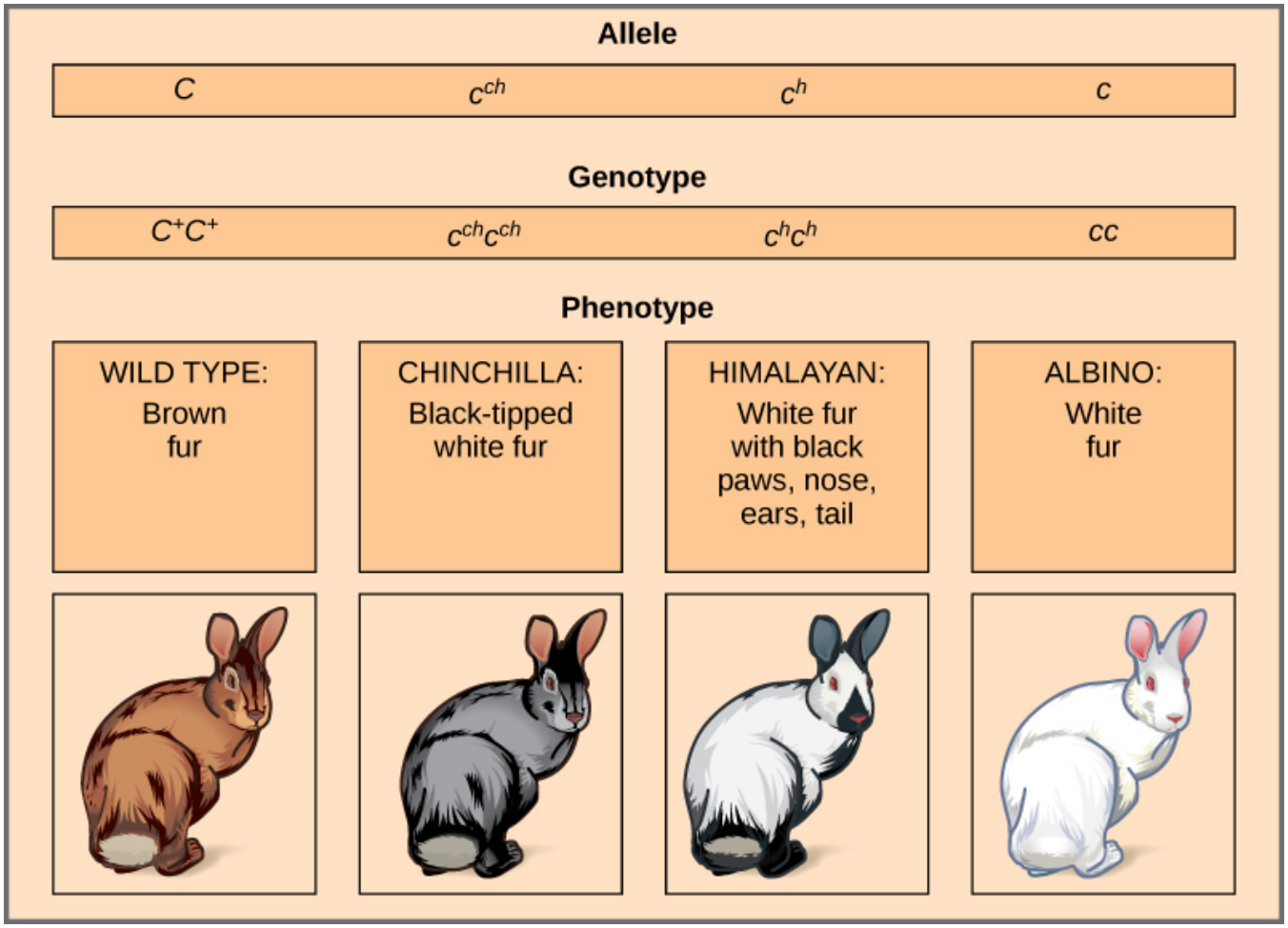


Figure 5. Four different alleles exist for the rabbit coat color (C) gene.

The complete dominance of a wild-type phenotype over all other mutants often occurs as an effect of "dosage" of a specific gene product, such that the wild-type allele supplies the correct amount of gene product whereas the mutant alleles cannot. For the allelic series in rabbits, the wild-type allele may supply a given dosage of fur pigment, whereas the mutants supply a lesser dosage or none at all. Interestingly, the Himalayan phenotype is the result of an allele that produces a temperature-sensitive gene product that only produces pigment in the cooler extremities of the rabbit's body.

Alternatively, one mutant allele can be dominant over all other phenotypes, including the wild type. This may occur when the mutant allele somehow interferes with the genetic message so that even a heterozygote with one wild-type allele copy expresses the mutant phenotype. One way in which the mutant allele can interfere is by enhancing the function of the wild-type gene product or changing its distribution in the body. One example of this is the *Antennapedia* mutation in *Drosophila* (Figure 6). In this case, the mutant allele expands the distribution of the gene product, and as a result, the *Antennapedia* heterozygote develops legs on its head where its antennae should be.

Figure 6. As seen in comparing the wild-type Drosophila (left) and the Antennapedia mutant (right), the Antennapedia mutant has legs on its head in place of antennae.

Evolution Connection

Multiple Alleles Confer Drug Resistance in the Malaria Parasite

Malaria is a parasitic disease in humans that is transmitted by infected female mosquitoes, including *Anopheles gambiae* (Figure 7a), and is characterized by cyclic high fevers, chills, flu-like symptoms, and severe anemia. *Plasmodium falciparum* and *P. vivax* are the most common causative agents of malaria, and *P. falciparum* is the most deadly (Figure 7b). When promptly and correctly treated, *P. falciparum* malaria has a mortality rate of 0.1 percent. However, in some parts of the world, the parasite has evolved resistance to commonly used malaria treatments, so the most effective malarial treatments can vary by geographic region.

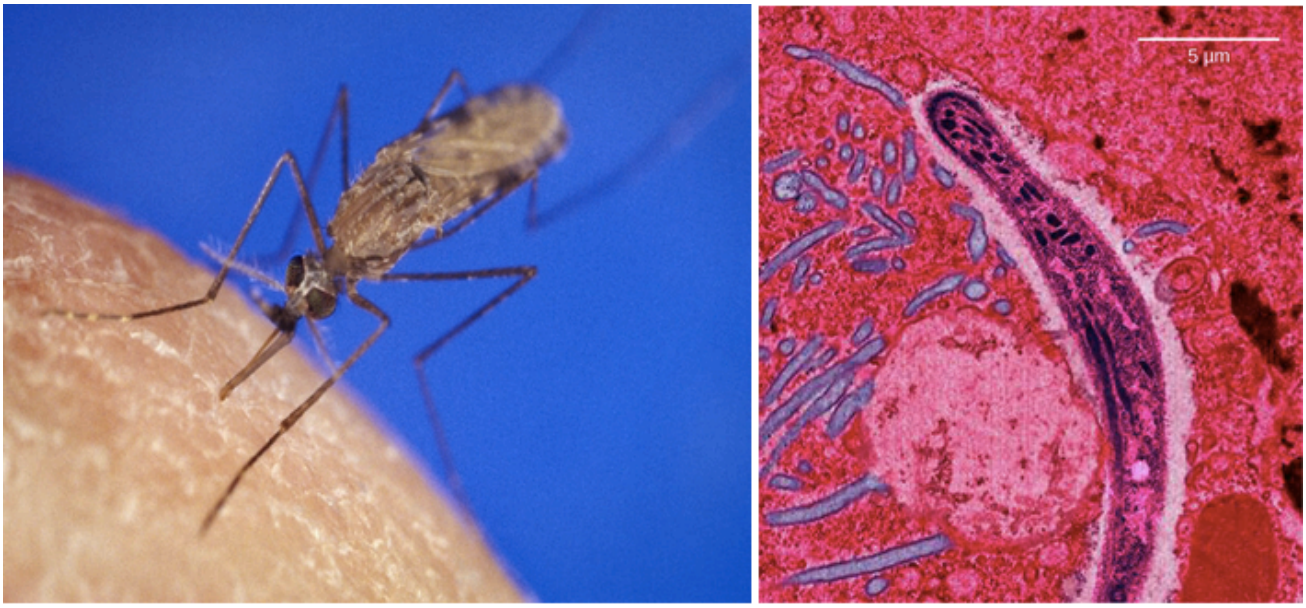


Figure 7. The (a) Anopheles gambiae, or African malaria mosquito, acts as a vector in the transmission to humans of the malaria-causing parasite (b) Plasmodium falciparum, here visualized using false-color transmission electron microscopy. (credit a: James D. Gathany; credit b: Ute Frevert; false color by Margaret Shear; scale-bar data from Matt Russell)

In Southeast Asia, Africa, and South America, *P. falciparum* has developed resistance to the anti-malarial drugs chloroquine, mefloquine, and sulfadoxine-pyrimethamine. *P. falciparum*, which is haploid during the life stage in which it is infectious to humans, has evolved multiple drug-resistant mutant alleles of the *dhps* gene. Varying degrees of sulfadoxine resistance are associated with each of these alleles. Being haploid, *P. falciparum* needs only one drug-resistant allele to express this trait.

In Southeast Asia, different sulfadoxine-resistant alleles of the *dhps* gene are localized to different geographic regions. This is a common evolutionary phenomenon that occurs because drug-resistant mutants arise in a population and interbreed with other *P. falciparum* isolates in close proximity. Sulfadoxine-resistant parasites cause considerable human hardship in regions where this drug is widely used as an over-the-counter malaria remedy. As is common with pathogens that multiply to large numbers within an infection cycle, *P. falciparum* evolves relatively rapidly (over a decade or so) in response to the selective pressure of commonly used anti-malarial drugs. For this reason, scientists must constantly work to develop new drugs or drug combinations to combat the worldwide malaria burden. ((Note: Sumiti Vinayak, et al., “Origin and Evolution of Sulfadoxine Resistant Plasmodium falciparum,” Public Library of Science Pathogens 6, no. 3 (2010): e1000830, doi:10.1371/journal.ppat.1000830.))

X-Linked Traits

In humans, as well as in many other animals and some plants, the sex of the individual is determined by sex chromosomes. The sex chromosomes are one pair of non-homologous chromosomes. Until now, we have only

considered inheritance patterns among non-sex chromosomes, or **autosomes**. In addition to 22 homologous pairs of autosomes, human females have a homologous pair of X chromosomes, whereas human males have an XY chromosome pair. Although the Y chromosome contains a small region of similarity to the X chromosome so that they can pair during meiosis, the Y chromosome is much shorter and contains many fewer genes. When a gene being examined is present on the X chromosome, but not on the Y chromosome, it is said to be **X-linked**.

Eye color in *Drosophila* was one of the first X-linked traits to be identified. Thomas Hunt Morgan mapped this trait to the X chromosome in 1910. Like humans, *Drosophila* males have an XY chromosome pair, and females are XX. In flies, the wild-type eye color is red (X^W) and it is dominant to white eye color (X^w) (Figure 8). Because of the location of the eye-color gene, reciprocal crosses do not produce the same offspring ratios. Males are said to be **hemizygous**, because they have only one allele for any X-linked characteristic. Hemizygosity makes the descriptions of dominance and recessiveness irrelevant for XY males. *Drosophila* males lack a second allele copy on the Y chromosome; that is, their genotype can only be X^WY or X^wY. In contrast, females have two allele copies of this gene and can be X^WX^W, X^WX^w, or X^wX^w.

Figure 8. In Drosophila, the gene for eye color is located on the X chromosome. Clockwise from top left are brown, cinnabar, sepia, vermilion, white, and red. Red eye color is wild-type and is dominant to white eye color.

In an X-linked cross, the genotypes of F_1 and F_2 offspring depend on whether the recessive trait was expressed by the male or the female in the P_1 generation. With regard to *Drosophila* eye color, when the P_1 male expresses the white-eye phenotype and the female is homozygous red-eyed, all members of the F_1 generation exhibit red eyes (Figure). The F_1 females are heterozygous (X^WX^w), and the males are all X^WY, having received their X chromosome from the homozygous dominant P_1 female and their Y chromosome from the P_1 male. A subsequent cross between the X^WX^w female and the X^WY male would produce only red-eyed females (with X^WX^W or X^WX^w genotypes) and both red- and white-eyed males (with X^WY or X^wY genotypes). Now, consider a cross between a homozygous white-eyed female and a male with red eyes. The F_1 generation would exhibit only heterozygous red-eyed females (X^WX^w) and only white-eyed males (X^wY). Half of the F_2 females would be red-eyed (X^WX^w) and half would be white-eyed (X^wX^w). Similarly, half of the F_2 males would be red-eyed (X^WY) and half would be white-eyed (X^wY).

Art Connection

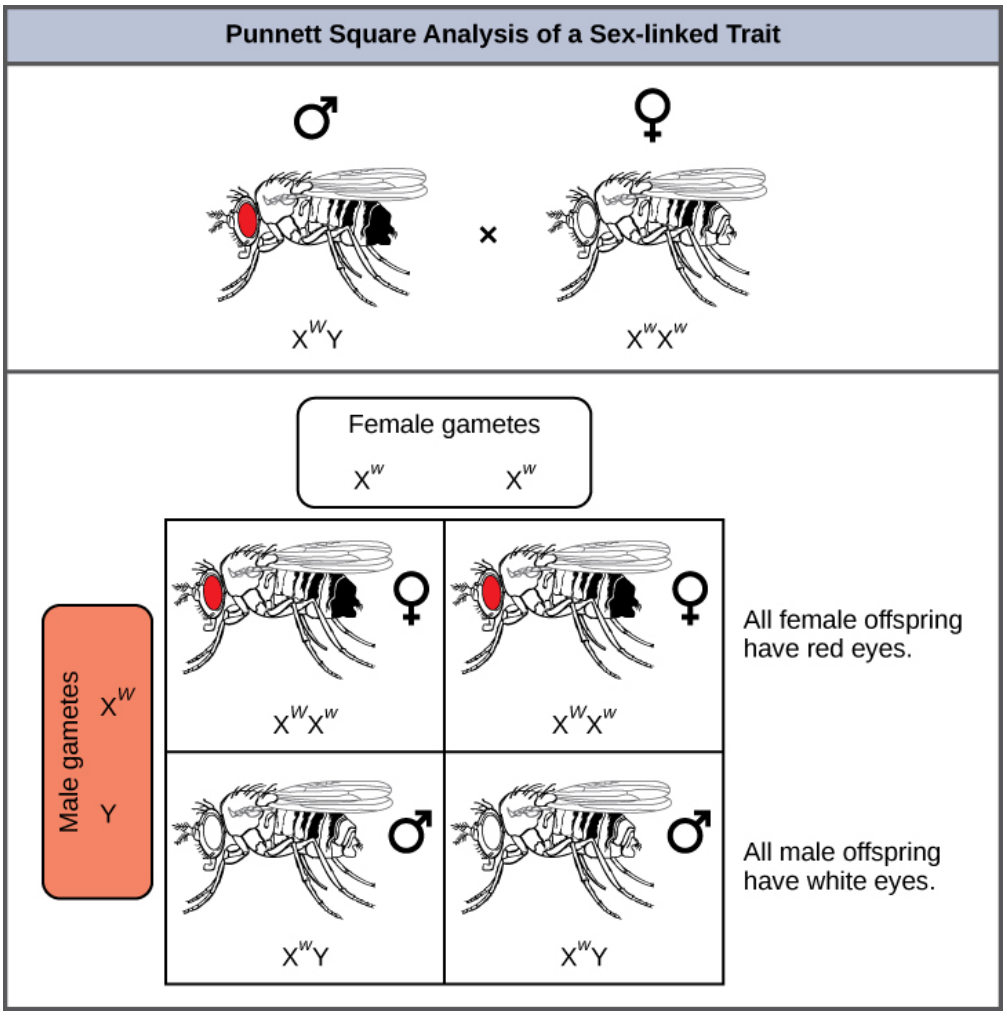


Figure 9. Punnett square analysis is used to determine the ratio of offspring from a cross between a red-eyed male fruit fly and a white-eyed female fruit fly.

What ratio of offspring would result from a cross between a white-eyed male and a female that is heterozygous for red eye color?

Discoveries in fruit fly genetics can be applied to human genetics. When a female parent is homozygous for a recessive X-linked trait, she will pass the trait on to 100 percent of her offspring. Her male offspring are, therefore, destined to express the trait, as they will inherit their father's Y chromosome. In humans, the alleles for certain conditions (some forms of color blindness, hemophilia, and muscular dystrophy) are X-linked. Females who are heterozygous for these diseases are said to be carriers and may not exhibit any phenotypic effects. These females will pass the disease to half of their sons and will pass carrier status to half of their daughters; therefore, recessive X-linked traits appear more frequently in males than females.

In some groups of organisms with sex chromosomes, the gender with the non-homologous sex chromosomes is the female rather than the male. This is the case for all birds. In this case, sex-linked traits will be more likely to appear in the female, in which they are hemizygous.

Human Sex-linked Disorders

Sex-linkage studies in Morgan's laboratory provided the fundamentals for understanding X-linked recessive disorders in humans, which include red-green color blindness, and Types A and B hemophilia. Because human males need to inherit only one recessive mutant X allele to be affected, X-linked disorders are disproportionately observed in males. Females must inherit recessive X-linked alleles from both of their parents in order to express the trait. When they inherit one recessive X-linked mutant allele and one dominant X-linked wild-type allele, they are carriers of the trait and are typically unaffected. Carrier females can manifest mild forms of the trait due to the inactivation of the dominant allele located on one of the X chromosomes. However, female carriers can contribute the trait to their sons, resulting in the son exhibiting the trait, or they can contribute the recessive allele to their daughters, resulting in the daughters being carriers of the trait (Figure 10). Although some Y-linked recessive disorders exist, typically they are associated with infertility in males and are therefore not transmitted to subsequent generations.

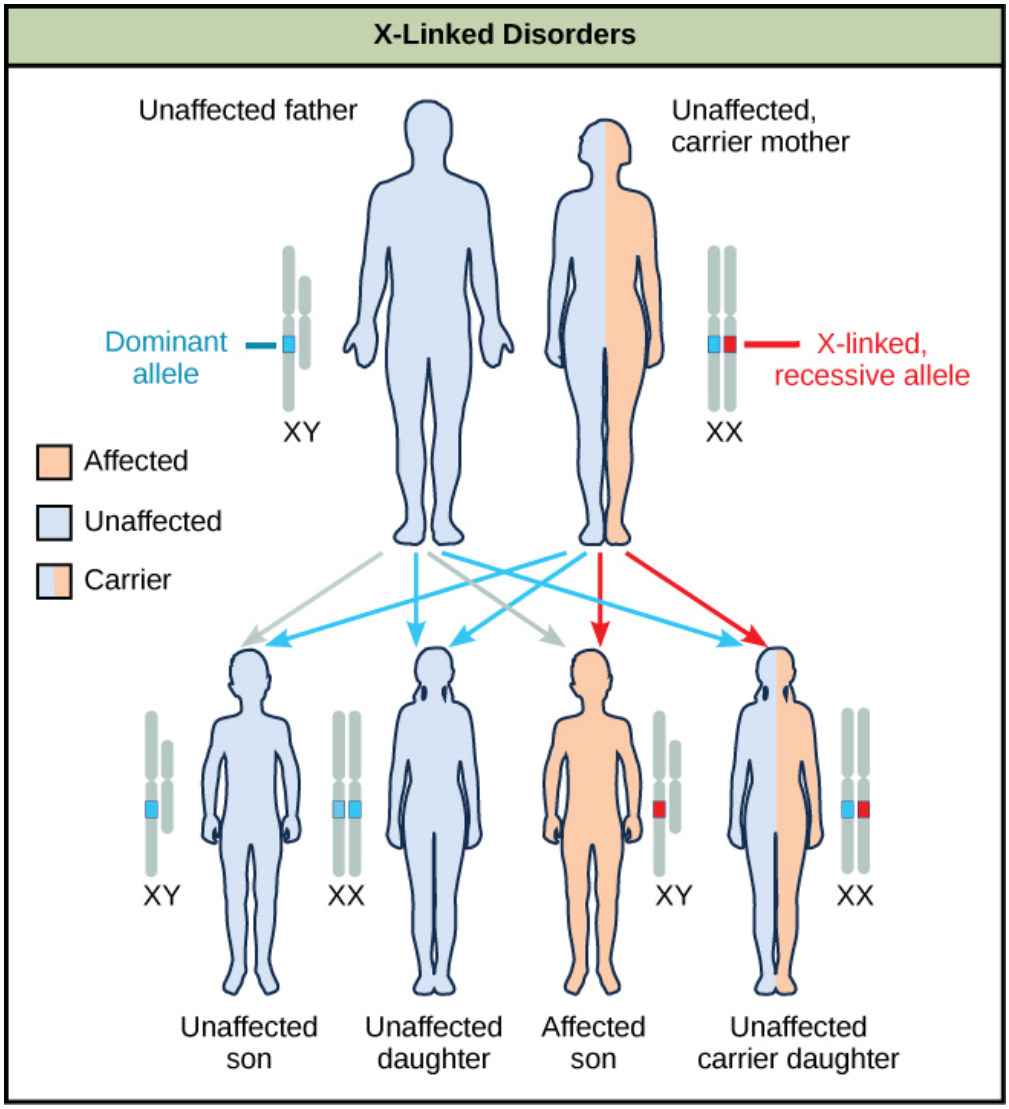


Figure 10. The son of a woman who is a carrier of a recessive X-linked disorder will have a 50 percent chance of being affected. A daughter will not be affected, but she will have a 50 percent chance of being a carrier like her mother.

Link to Learning

Watch this video to learn more about sex-linked traits.
Watch this video online: https://youtu.be/-ROhfKyxgCo

Lethality

A large proportion of genes in an individual's genome are essential for survival. Occasionally, a nonfunctional allele for an essential gene can arise by mutation and be transmitted in a population as long as individuals with this allele also have a wild-type, functional copy. The wild-type allele functions at a capacity sufficient to sustain life and is therefore considered to be dominant over the nonfunctional allele. However, consider two heterozygous parents that have a genotype of wild-type/nonfunctional mutant for a hypothetical essential gene. In one quarter of their offspring, we would expect to observe individuals that are homozygous recessive for the nonfunctional allele. Because the gene is essential, these individuals might fail to develop past fertilization, die *in utero*, or die later in life, depending on what life stage requires this gene. An inheritance pattern in which an allele is only lethal in the homozygous form and in which the heterozygote may be normal or have some altered non-lethal phenotype is referred to as **recessive lethal**.

For crosses between heterozygous individuals with a recessive lethal allele that causes death before birth when homozygous, only wild-type homozygotes and heterozygotes would be observed. The genotypic ratio would therefore be 2:1. In other instances, the recessive lethal allele might also exhibit a dominant (but not lethal) phenotype in the heterozygote. For instance, the recessive lethal *Curly* allele in *Drosophila* affects wing shape in the heterozygote form but is lethal in the homozygote.

A single copy of the wild-type allele is not always sufficient for normal functioning or even survival. The **dominant lethal** inheritance pattern is one in which an allele is lethal both in the homozygote and the heterozygote; this allele can only be transmitted if the lethality phenotype occurs after reproductive age. Individuals with mutations that result in dominant lethal alleles fail to survive even in the heterozygote form. Dominant lethal alleles are very rare because, as you might expect, the allele only lasts one generation and is not transmitted. However, just as the recessive lethal allele might not immediately manifest the phenotype of death, dominant lethal alleles also might not be expressed until adulthood. Once the individual reaches reproductive age, the allele may be unknowingly passed on, resulting in a delayed death in both generations. An example of this in humans is Huntington's disease, in which the nervous system gradually wastes away (Figure 11). People who are heterozygous for the dominant Huntington allele (*Hh*) will inevitably develop the fatal disease. However, the onset of Huntington's disease may not occur until age 40, at which point the afflicted persons may have already passed the allele to 50 percent of their offspring.

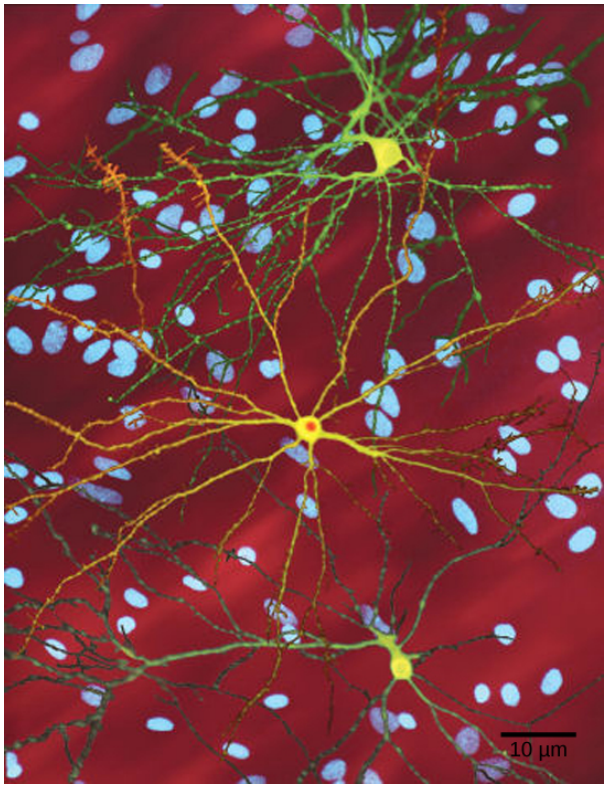


Figure 11. The neuron in the center of this micrograph (yellow) has nuclear inclusions characteristic of Huntington's disease (orange area in the center of the neuron). Huntington's disease occurs when an abnormal dominant allele for the Huntington gene is present. (credit:

Section Summary

When true-breeding or homozygous individuals that differ for a certain trait are crossed, all of the offspring will be heterozygotes for that trait. If the traits are inherited as dominant and recessive, the F_1 offspring will all exhibit the

same phenotype as the parent homozygous for the dominant trait. If these heterozygous offspring are self-crossed, the resulting F_2 offspring will be equally likely to inherit gametes carrying the dominant or recessive trait, giving rise to offspring of which one quarter are homozygous dominant, half are heterozygous, and one quarter are homozygous recessive. Because homozygous dominant and heterozygous individuals are phenotypically identical, the observed traits in the F_2 offspring will exhibit a ratio of three dominant to one recessive.

Dr. Steven Finkbeiner, Gladstone Institute of Neurological Disease, The Taube-Koret Center for Huntington's Disease Research, and the University of California San Francisco/ Wikimedia)

Alleles do not always behave in dominant and recessive patterns. Incomplete dominance describes situations in which the heterozygote exhibits a phenotype that is intermediate between the homozygous phenotypes. Codominance describes the simultaneous expression of both of the alleles in the heterozygote. Although diploid organisms can only have two alleles for any given gene, it is common for more than two alleles of a gene to exist in a population. In humans, as in many animals and some plants, females have two X chromosomes and males have one X and one Y chromosome. Genes that are present on the X but not the Y chromosome are said to be X-linked, such that males only inherit one allele for the gene, and females inherit two. Finally, some alleles can be lethal. Recessive lethal alleles are only lethal in homozygotes, but dominant lethal alleles are fatal in heterozygotes as well.

Additional Self Check Questions

1. In pea plants, round peas (*R*) are dominant to wrinkled peas (*r*). You do a test cross between a pea plant with wrinkled peas (genotype *rr*) and a plant of unknown genotype that has round peas. You end up with three plants, all which have round peas. From this data, can you tell if the round pea parent plant is homozygous dominant or heterozygous? If the round pea parent plant is heterozygous, what is the probability that a random sample of 3 progeny peas will all be round?

2. What are the genotypes of the individuals labeled 1, 2 and 3?

3. What ratio of offspring would result from a cross between a white-eyed male and a female that is heterozygous for red eye color?

4. The gene for flower position in pea plants exists as axial or terminal alleles. Given that axial is dominant to terminal, list all of the possible F_1 and F_2 genotypes and phenotypes from a cross involving parents that are homozygous for each trait. Express genotypes with conventional genetic abbreviations.

5. Use a Punnett square to predict the offspring in a cross between a dwarf pea plant (homozygous recessive) and a tall pea plant (heterozygous). What is the phenotypic ratio of the offspring?

6. Can a human male be a carrier of red-green color blindness?

Answers

1. You cannot be sure if the plant is homozygous or heterozygous as the data set is too small: by random chance, all three plants might have acquired only the dominant gene even if the recessive one is present. If the round pea parent is heterozygous, there is a one-eighth probability that a random sample of three progeny peas will all be round.

2. Individual 1 has the genotype *aa*. Individual 2 has the genotype *Aa*. Individual 3 has the genotype *Aa*.

3. Half of the female offspring would be heterozygous (X^WX^w) with red eyes, and half would be homozygous recessive (X^wX^w) with white eyes. Half of the male offspring would be hemizygous dominant (X^WY) withe red yes, and half would be hemizygous recessive (X^wY) with white eyes.

4. Because axial is dominant, the gene would be designated as *A*. F_1 would be all heterozygous *Aa* with axial phenotype. F_2 would have possible genotypes of *AA*, *Aa*, and *aa*; these would correspond to axial, axial, and terminal phenotypes, respectively.

5. The Punnett square would be 2 × 2 and will have *T* and *T* along the top, and *T* and *t* along the left side. Clockwise from the top left, the genotypes listed within the boxes will be *Tt*, *Tt*, *tt*, and *tt*. The phenotypic ratio will be 1 tall:1 dwarf.

6. No, males can only express color blindness. They cannot carry it because an individual needs two X chromosomes to be a carrier.

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: Open Stax. **Project**: http://cnx.org/contents/185cbf87-c72e-48f5-b51e-f14f21b5eabd@9.17:1/Biology. **License**: *CC BY: Attribution*

All rights reserved content

- Biology. **Authored by**: OpenStax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License**: *CC BY: Attribution*

LAWS OF INHERITANCE

Learning Objectives

By the end of this section, you will be able to:

- Explain Mendel's law of segregation and independent assortment in terms of genetics and the events of meiosis
- Use the forked-line method and the probability rules to calculate the probability of genotypes and phenotypes from multiple gene crosses
- Explain the effect of linkage and recombination on gamete genotypes
- Explain the phenotypic outcomes of epistatic effects between genes

Mendel generalized the results of his pea-plant experiments into four postulates, some of which are sometimes called "laws," that describe the basis of dominant and recessive inheritance in diploid organisms. As you have learned, more complex extensions of Mendelism exist that do not exhibit the same F_2 phenotypic ratios (3:1). Nevertheless, these laws summarize the basics of classical genetics.

Pairs of Unit Factors, or Genes

Mendel proposed first that paired unit factors of heredity were transmitted faithfully from generation to generation by the dissociation and reassociation of paired factors during gametogenesis and fertilization, respectively. After he crossed peas with contrasting traits and found that the recessive trait resurfaced in the F_2 generation, Mendel deduced that hereditary factors must be inherited as discrete units. This finding contradicted the belief at that time that parental traits were blended in the offspring.

Alleles Can Be Dominant or Recessive

Mendel's **law of dominance** states that in a heterozygote, one trait will conceal the presence of another trait for the same characteristic. Rather than both alleles contributing to a phenotype, the dominant allele will be expressed exclusively. The recessive allele will remain "latent" but will be transmitted to offspring by the same manner in which the dominant allele is transmitted. The recessive trait will only be expressed by offspring that have two copies of this allele (Figure 1), and these offspring will breed true when self-crossed.

Since Mendel's experiments with pea plants, other researchers have found that the law of dominance does not always hold true. Instead, several different patterns of inheritance have been found to exist.

Figure 1. The child in the photo expresses albinism, a recessive trait.

Equal Segregation of Alleles

Observing that true-breeding pea plants with contrasting traits gave rise to F_1 generations that all expressed the dominant trait and F_2 generations that expressed the dominant and recessive traits in a 3:1 ratio, Mendel proposed the law of segregation. This law states that paired unit factors (genes) must segregate equally into gametes such that offspring have an equal likelihood of inheriting either factor. For the F_2 generation of a monohybrid cross, the following three possible combinations of genotypes could result: homozygous dominant, heterozygous, or homozygous recessive. Because heterozygotes could arise from two different pathways (receiving one dominant and one recessive allele from either parent), and because heterozygotes and homozygous dominant individuals are phenotypically identical, the law supports Mendel's observed 3:1 phenotypic ratio. The equal segregation of alleles is the reason we can apply the Punnett square to accurately predict the offspring of parents with known genotypes. The physical basis of Mendel's law of segregation is the first division of meiosis, in which the homologous chromosomes with their different versions of each gene are segregated into daughter nuclei. The role of the meiotic segregation of chromosomes in sexual reproduction was not understood by the scientific community during Mendel's lifetime.

Independent Assortment

Mendel's **law of independent assortment** states that genes do not influence each other with regard to the sorting of alleles into gametes, and every possible combination of alleles for every gene is equally likely to occur. The independent assortment of genes can be illustrated by the **dihybrid** cross, a cross between two true-breeding parents that express different traits for two characteristics. Consider the characteristics of seed color and seed texture for two pea plants, one that has green, wrinkled seeds (*yyrr*) and another that has yellow, round seeds (*YYRR*). Because each parent is homozygous, the law of segregation indicates that the gametes for the green/wrinkled plant all are *yr*, and the gametes for the yellow/round plant are all *YR*. Therefore, the F_1 generation of offspring all are *YyRr* (Figure 2).

Art Connection

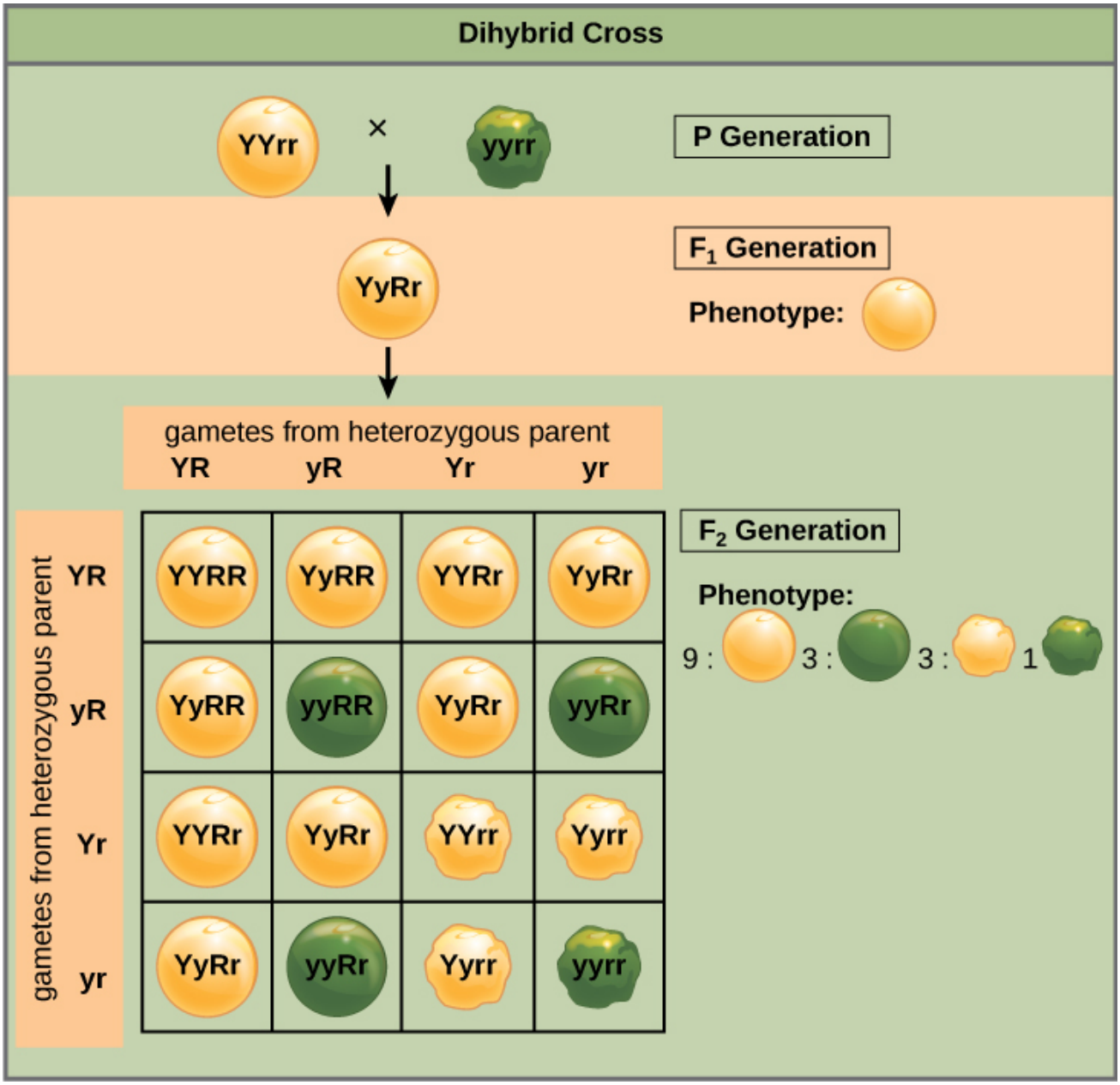


Figure 2. This dihybrid cross of pea plants involves the genes for seed color and texture.

In pea plants, purple flowers (P) are dominant to white flowers (p) and yellow peas (Y) are dominant to green peas (y). What are the possible genotypes and phenotypes for a cross between PpYY and ppYy pea plants? How many squares do you need to do a Punnett square analysis of this cross?

For the F_2 generation, the law of segregation requires that each gamete receive either an *R* allele or an *r* allele along with either a *Y* allele or a *y* allele. The law of independent assortment states that a gamete into which an *r* allele sorted would be equally likely to contain either a *Y* allele or a *y* allele. Thus, there are four equally likely gametes that can be formed when the *YyRr* heterozygote is self-crossed, as follows: *YR*, *Yr*, *yR*, and *yr*. Arranging these gametes along the top and left of a 4 × 4 Punnett square (Figure) gives us 16 equally likely genotypic combinations. From these genotypes, we infer a phenotypic ratio of 9 round/yellow:3 round/green:3 wrinkled/yellow:1 wrinkled/green (Figure 2). These are the offspring ratios we would expect, assuming we performed the crosses with a large enough sample size.

Because of independent assortment and dominance, the 9:3:3:1 dihybrid phenotypic ratio can be collapsed into two 3:1 ratios, characteristic of any monohybrid cross that follows a dominant and recessive pattern. Ignoring seed color and considering only seed texture in the above dihybrid cross, we would expect that three quarters of the F_2 generation offspring would be round, and one quarter would be wrinkled. Similarly, isolating only seed color, we would assume that three quarters of the F_2 offspring would be yellow and one quarter would be green. The sorting of alleles for texture and color are independent events, so we can apply the product rule. Therefore, the proportion of round and yellow F_2 offspring is expected to be (3/4) × (3/4) = 9/16, and the proportion of wrinkled and green offspring is expected to be (1/4) × (1/4) = 1/16. These proportions are identical to those obtained using a Punnett square. Round, green and wrinkled, yellow offspring can also be calculated using the product rule, as each of these genotypes includes one dominant and one recessive phenotype. Therefore, the proportion of each is calculated as (3/4) × (1/4) = 3/16.

The law of independent assortment also indicates that a cross between yellow, wrinkled (*YYrr*) and green, round (*yyRR*) parents would yield the same F_1 and F_2 offspring as in the *YYRR* x *yyrr* cross.

The physical basis for the law of independent assortment also lies in meiosis I, in which the different homologous pairs line up in random orientations. Each gamete can contain any combination of paternal and maternal chromosomes (and therefore the genes on them) because the orientation of tetrads on the metaphase plane is random.

Forked-Line Method

When more than two genes are being considered, the Punnett-square method becomes unwieldy. For instance, examining a cross involving four genes would require a 16 × 16 grid containing 256 boxes. It would be extremely cumbersome to manually enter each genotype. For more complex crosses, the forked-line and probability methods are preferred.

To prepare a forked-line diagram for a cross between F_1 heterozygotes resulting from a cross between *AABBCC* and *aabbcc* parents, we first create rows equal to the number of genes being considered, and then segregate the alleles in each row on forked lines according to the probabilities for individual monohybrid crosses (Figure 3). We then multiply the values along each forked path to obtain the F_2 offspring probabilities. Note that this process is a diagrammatic version of the product rule. The values along each forked pathway can be multiplied because each gene assorts independently. For a trihybrid cross, the F_2 phenotypic ratio is 27:9:9:3:3:3:1.

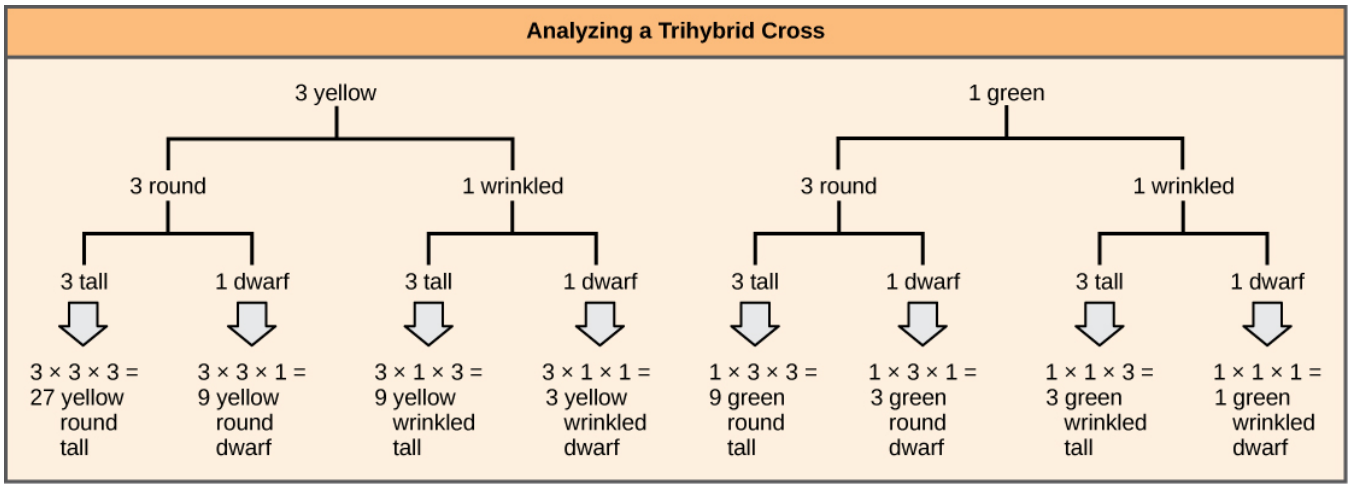


Figure 3. The forked-line method can be used to analyze a trihybrid cross. Here, the probability for color in the F2 generation occupies the top row (3 yellow:1 green). The probability for shape occupies the second row (3 round:1 wrinked), and the probability for height occupies the third row (3 tall:1 dwarf). The probability for each possible combination of traits is calculated by multiplying the probability for each individual trait. Thus, the probability of F2 offspring having yellow, round, and tall traits is 3 × 3 × 3, or 27.

Probability Method

While the forked-line method is a diagrammatic approach to keeping track of probabilities in a cross, the probability method gives the proportions of offspring expected to exhibit each phenotype (or genotype) without the added visual assistance. Both methods make use of the product rule and consider the alleles for each gene separately. Earlier, we examined the phenotypic proportions for a trihybrid cross using the forked-line method; now we will use the probability method to examine the genotypic proportions for a cross with even more genes.

For a trihybrid cross, writing out the forked-line method is tedious, albeit not as tedious as using the Punnett-square method. To fully demonstrate the power of the probability method, however, we can consider specific genetic calculations. For instance, for a tetrahybrid cross between individuals that are heterozygotes for all four genes, and in which all four genes are sorting independently and in a dominant and recessive pattern, what proportion of the offspring will be expected to be homozygous recessive for all four alleles? Rather than writing out every possible genotype, we can use the probability method. We know that for each gene, the fraction of homozygous recessive offspring will be 1/4. Therefore, multiplying this fraction for each of the four genes, (1/4) × (1/4) × (1/4) × (1/4), we determine that 1/256 of the offspring will be quadruply homozygous recessive.

For the same tetrahybrid cross, what is the expected proportion of offspring that have the dominant phenotype at all four loci? We can answer this question using phenotypic proportions, but let's do it the hard way—using genotypic proportions. The question asks for the proportion of offspring that are 1) homozygous dominant at *A* or heterozygous at *A,* and 2) homozygous at *B* or heterozygous at *B*, and so on. Noting the "or" and "and" in each circumstance makes clear where to apply the sum and product rules. The probability of a homozygous dominant at *A* is 1/4 and the probability of a heterozygote at *A* is 1/2. The probability of the homozygote or the heterozygote is 1/4 + 1/2 = 3/4 using the sum rule. The same probability can be obtained in the same way for each of the other genes, so that the probability of a dominant phenotype at *A* and *B* and *C* and *D* is, using the product rule, equal to 3/4 × 3/4 × 3/4 × 3/4, or 27/64. If you are ever unsure about how to combine probabilities, returning to the forked-line method should make it clear.

Rules for Multihybrid Fertilization

Predicting the genotypes and phenotypes of offspring from given crosses is the best way to test your knowledge of Mendelian genetics. Given a multihybrid cross that obeys independent assortment and follows a dominant and recessive pattern, several generalized rules exist; you can use these rules to check your results as you work through genetics calculations (Table 1). To apply these rules, first you must determine *n*, the number of heterozygous gene pairs (the number of genes segregating two alleles each). For example, a cross between *AaBb* and *AaBb* heterozygotes has an *n* of 2. In contrast, a cross between *AABb* and *AABb* has an *n* of 1 because *A* is not heterozygous.

Table 1. General Rules for Multihybrid Crosses

General Rule	Number of Heterozygous Gene Pairs
Number of different F_1 gametes	2^n
Number of different F_2 genotypes	3^n
Given dominant and recessive inheritance, the number of different F_2 phenotypes	2^n

Linked Genes Violate the Law of Independent Assortment

Although all of Mendel's pea characteristics behaved according to the law of independent assortment, we now know that some allele combinations are not inherited independently of each other. Genes that are located on separate non-homologous chromosomes will always sort independently. However, each chromosome contains

hundreds or thousands of genes, organized linearly on chromosomes like beads on a string. The segregation of alleles into gametes can be influenced by linkage, in which genes that are located physically close to each other on the same chromosome are more likely to be inherited as a pair. However, because of the process of recombination, or "crossover," it is possible for two genes on the same chromosome to behave independently, or as if they are not linked. To understand this, let's consider the biological basis of gene linkage and recombination.

Homologous chromosomes possess the same genes in the same linear order. The alleles may differ on homologous chromosome pairs, but the genes to which they correspond do not. In preparation for the first division of meiosis, homologous chromosomes replicate and synapse. Like genes on the homologs align with each other. At this stage, segments of homologous chromosomes exchange linear segments of genetic material (Figure 4). This process is called recombination, or crossover, and it is a common genetic process. Because the genes are aligned during recombination, the gene order is not altered. Instead, the result of recombination is that maternal and paternal alleles are combined onto the same chromosome. Across a given chromosome, several recombination events may occur, causing extensive shuffling of alleles.

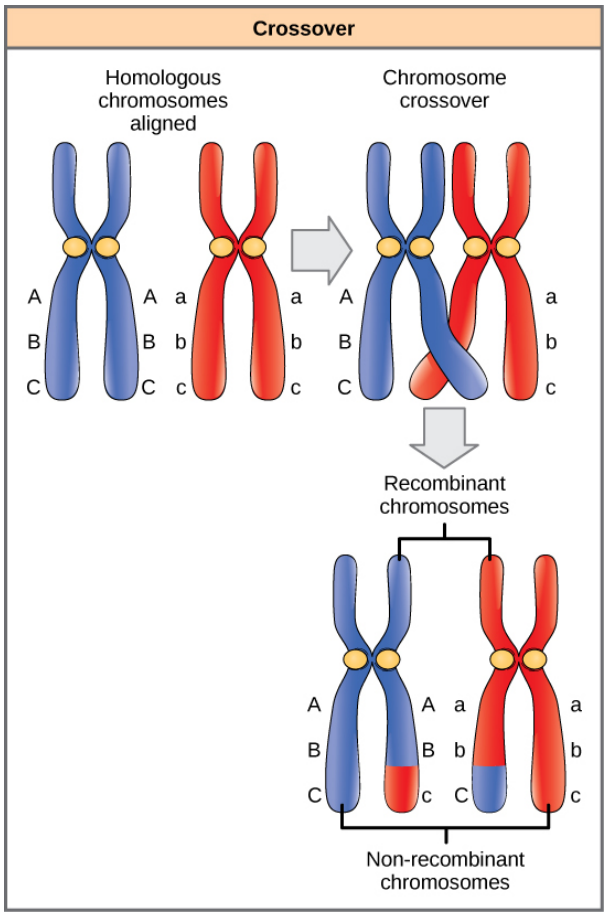


Figure 4. The process of crossover, or recombination, occurs when two homologous chromosomes align during meiosis and exchange a segment of genetic material. Here, the alleles for gene C were exchanged. The result is two recombinant and two non-recombinant chromosomes.

When two genes are located in close proximity on the same chromosome, they are considered linked, and their alleles tend to be transmitted through meiosis together. To exemplify this, imagine a dihybrid cross involving flower color and plant height in which the genes are next to each other on the chromosome. If one homologous chromosome has alleles for tall plants and red flowers, and the other chromosome has genes for short plants and yellow flowers, then when the gametes are formed, the tall and red alleles will go together into a gamete and the short and yellow alleles will go into other gametes. These are called the parental genotypes because they have been inherited intact from the parents of the individual producing gametes. But unlike if the genes were on different chromosomes, there will be no gametes with tall and yellow alleles and no gametes with short and red alleles. If you create the Punnett square with these gametes, you will see that the classical Mendelian prediction of a 9:3:3:1 outcome of a dihybrid cross would not apply. As the distance between two genes increases, the probability of one or more crossovers between them increases, and the genes behave more like they are on separate chromosomes. Geneticists have used the proportion of recombinant gametes (the ones not like the parents) as a measure of how far apart genes are on a chromosome. Using this information, they have constructed elaborate maps of genes on chromosomes for well-studied organisms, including humans.

Mendel's seminal publication makes no mention of linkage, and many researchers have questioned whether he encountered linkage but chose not to publish those crosses out of concern that they would invalidate his independent assortment postulate. The garden pea has seven chromosomes, and some have suggested that his choice of seven characteristics was not a coincidence. However, even if the genes he examined were not located on separate chromosomes, it is possible that he simply did not observe linkage because of the extensive shuffling effects of recombination.

Scientific Method Connection

Testing the Hypothesis of Independent Assortment

To better appreciate the amount of labor and ingenuity that went into Mendel's experiments, proceed through one of Mendel's dihybrid crosses.

Question: What will be the offspring of a dihybrid cross?

Background: Consider that pea plants mature in one growing season, and you have access to a large garden in which you can cultivate thousands of pea plants. There are several true-breeding plants with the following pairs of traits: tall plants with inflated pods, and dwarf plants with constricted pods. Before the plants have matured, you remove the pollen-producing organs from the tall/inflated plants in your crosses to prevent self-fertilization. Upon plant maturation, the plants are manually crossed by transferring pollen from the dwarf/constricted plants to the stigmata of the tall/inflated plants.

Hypothesis: Both trait pairs will sort independently according to Mendelian laws. When the true-breeding parents are crossed, all of the F_1 offspring are tall and have inflated pods, which indicates that the tall and inflated traits are dominant over the dwarf and constricted traits, respectively. A self-cross of the F_1 heterozygotes results in 2,000 F_2 progeny.

Test the hypothesis: Because each trait pair sorts independently, the ratios of tall:dwarf and inflated:constricted are each expected to be 3:1. The tall/dwarf trait pair is called *T/t*, and the inflated/constricted trait pair is designated *I/i*. Each member of the F_1 generation therefore has a genotype of *TtIi*. Construct a grid analogous to Figure 5, in which you cross two *TtIi* individuals. Each individual can donate four combinations of two traits: *TI*, *Ti*, *tI*, or *ti*, meaning that there are 16 possibilities of offspring genotypes. Because the *T* and *I* alleles are dominant, any individual having one or two of those alleles will express the tall or inflated phenotypes, respectively, regardless if they also have a *t* or *i* allele. Only individuals that are *tt* or *ii* will express the dwarf and constricted alleles, respectively. As shown in Figure 5, you predict that you will observe the following offspring proportions: tall/inflated:tall/constricted:dwarf/inflated:dwarf/constricted in a 9:3:3:1 ratio. Notice from the grid that when considering the tall/dwarf and inflated/constricted trait pairs in isolation, they are each inherited in 3:1 ratios.

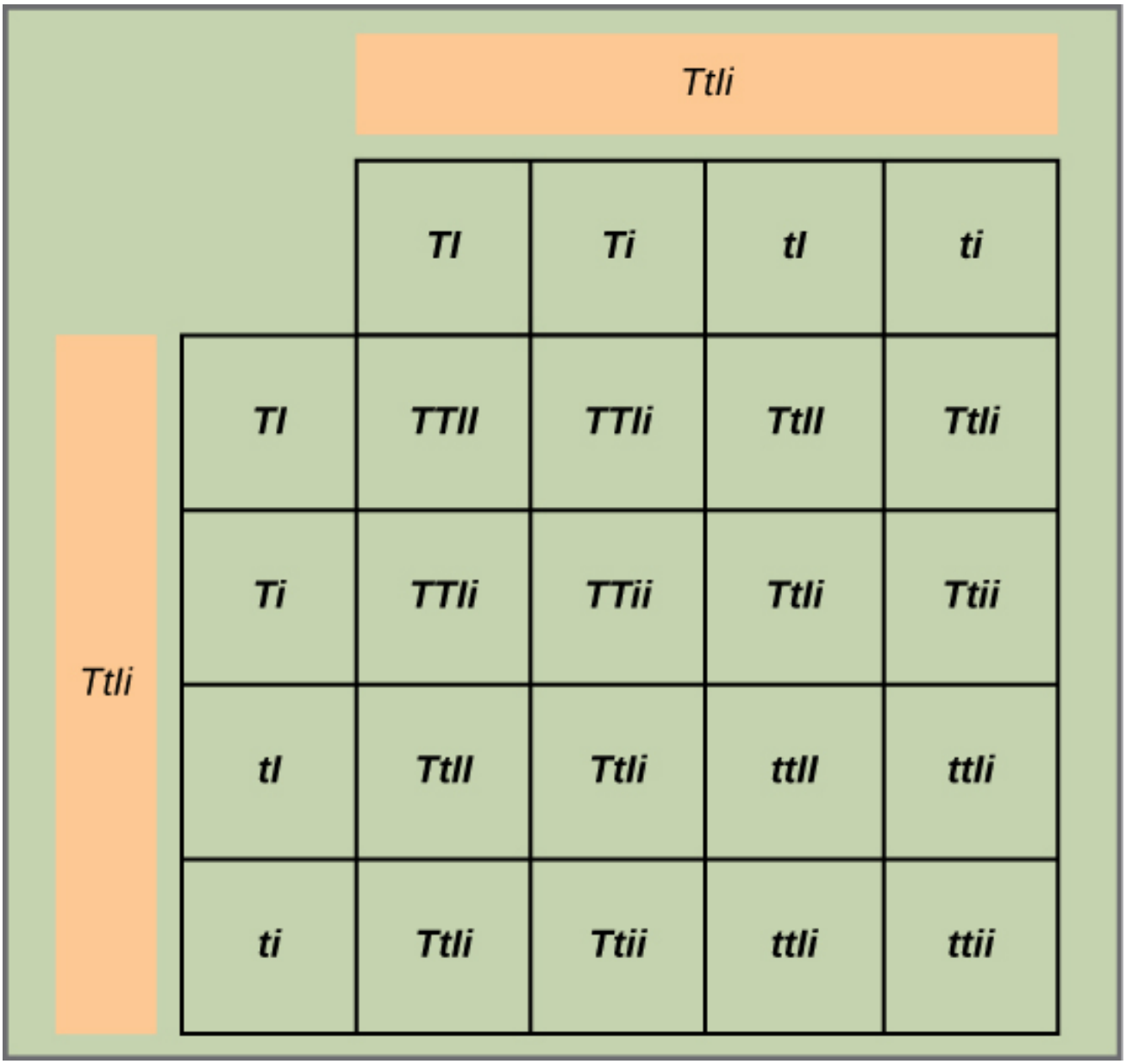


Figure 5. This figure shows all possible combinations of offspring resulting from a dihybrid cross of pea plants that are heterozygous for the tall/dwarf and inflated/constricted alleles.

Test the hypothesis: You cross the dwarf and tall plants and then self-cross the offspring. For best results, this is repeated with hundreds or even thousands of pea plants. What special precautions should be taken in the crosses and in growing the plants?

Analyze your data: You observe the following plant phenotypes in the F_2 generation: 2706 tall/inflated, 930 tall/constricted, 888 dwarf/inflated, and 300 dwarf/constricted. Reduce these findings to a ratio and determine if they are consistent with Mendelian laws.

Form a conclusion: Were the results close to the expected 9:3:3:1 phenotypic ratio? Do the results support the prediction? What might be observed if far fewer plants were used, given that alleles segregate randomly into gametes? Try to imagine growing that many pea plants, and consider the potential for experimental error. For instance, what would happen if it was extremely windy one day?

Epistasis

Mendel's studies in pea plants implied that the sum of an individual's phenotype was controlled by genes (or as he called them, unit factors), such that every characteristic was distinctly and completely controlled by a single gene. In fact, single observable characteristics are almost always under the influence of multiple genes (each with two or more alleles) acting in unison. For example, at least eight genes contribute to eye color in humans.

Link to Learning

Eye color in humans is determined by multiple genes. Use the Eye Color Calculator to predict the eye color of children from parental eye color.

In some cases, several genes can contribute to aspects of a common phenotype without their gene products ever directly interacting. In the case of organ development, for instance, genes may be expressed sequentially, with each gene adding to the complexity and specificity of the organ. Genes may function in complementary or synergistic fashions, such that two or more genes need to be expressed simultaneously to affect a phenotype. Genes may also oppose each other, with one gene modifying the expression of another.

In epistasis, the interaction between genes is antagonistic, such that one gene masks or interferes with the expression of another. "Epistasis" is a word composed of Greek roots that mean "standing upon." The alleles that are being masked or silenced are said to be hypostatic to the epistatic alleles that are doing the masking. Often the biochemical basis of epistasis is a gene pathway in which the expression of one gene is dependent on the function of a gene that precedes or follows it in the pathway.

An example of epistasis is pigmentation in mice. The wild-type coat color, agouti (*AA*), is dominant to solid-colored fur (*aa*). However, a separate gene (*C*) is necessary for pigment production. A mouse with a recessive *c* allele at this locus is unable to produce pigment and is albino regardless of the allele present at locus *A* (Figure 6). Therefore, the genotypes *AAcc*, *Aacc*, and *aacc* all produce the same albino phenotype. A cross between heterozygotes for both genes (*AaCc* x *AaCc*) would generate offspring with a phenotypic ratio of 9 agouti:3 solid color:4 albino (Figure 6). In this case, the *C* gene is epistatic to the *A* gene.

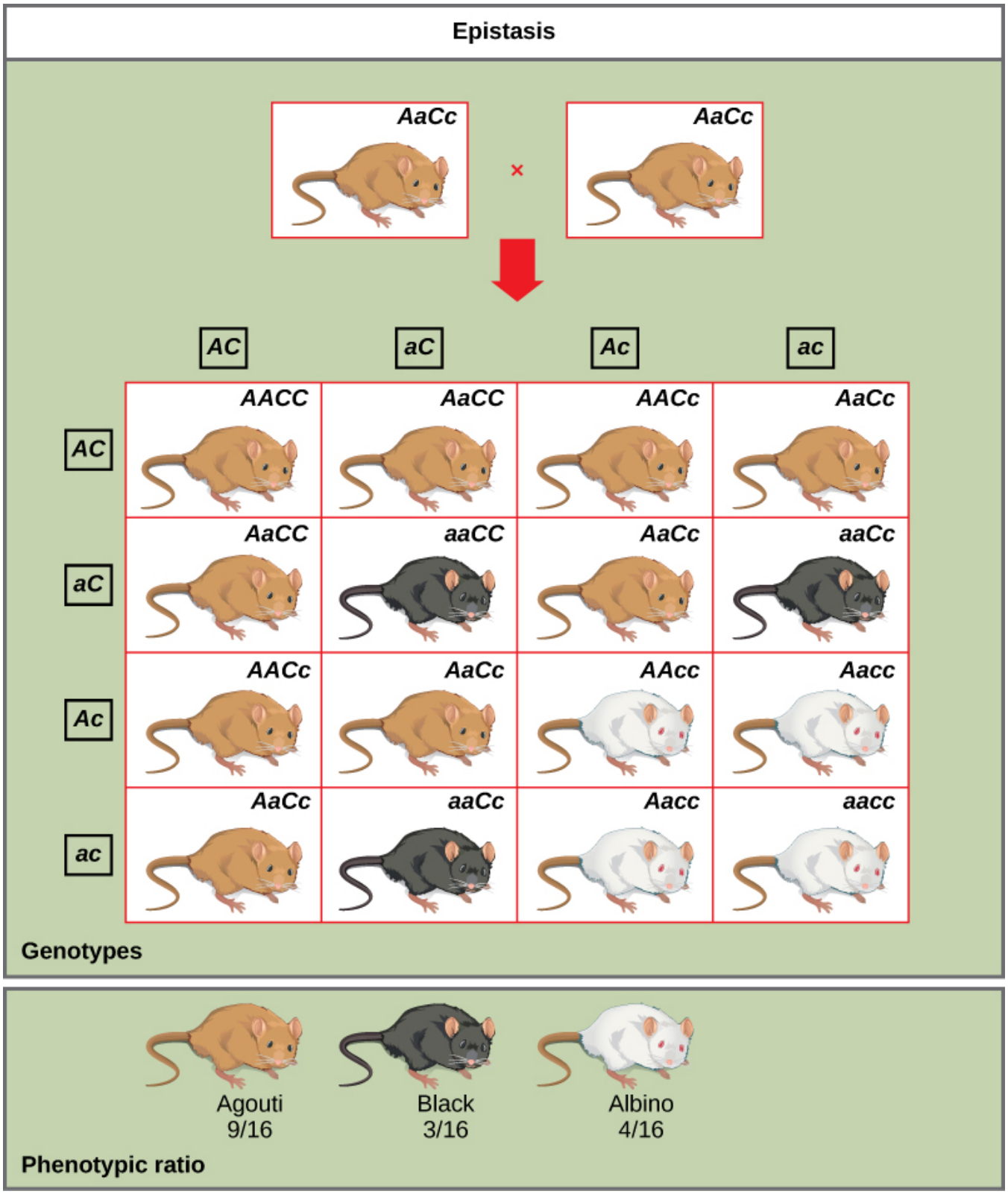


Figure 6. In mice, the mottled agouti coat color (A) is dominant to a solid coloration, such as black or gray. A gene at a separate locus (C) is responsible for pigment production. The recessive c allele does not produce pigment, and a mouse with the homozygous recessive cc genotype is albino regardless of the allele present at the A locus. Thus, the C gene is epistatic to the A gene.

Epistasis can also occur when a dominant allele masks expression at a separate gene. Fruit color in summer squash is expressed in this way. Homozygous recessive expression of the *W* gene (*ww*) coupled with

homozygous dominant or heterozygous expression of the *Y* gene (*YY* or *Yy*) generates yellow fruit, and the *wwyy* genotype produces green fruit. However, if a dominant copy of the *W* gene is present in the homozygous or heterozygous form, the summer squash will produce white fruit regardless of the *Y* alleles. A cross between white heterozygotes for both genes (*WwYy* × *WwYy*) would produce offspring with a phenotypic ratio of 12 white:3 yellow:1 green.

Finally, epistasis can be reciprocal such that either gene, when present in the dominant (or recessive) form, expresses the same phenotype. In the shepherd's purse plant (*Capsella bursa-pastoris*), the characteristic of seed shape is controlled by two genes in a dominant epistatic relationship. When the genes *A* and *B* are both homozygous recessive (*aabb*), the seeds are ovoid. If the dominant allele for either of these genes is present, the result is triangular seeds. That is, every possible genotype other than *aabb* results in triangular seeds, and a cross between heterozygotes for both genes (*AaBb* x *AaBb*) would yield offspring with a phenotypic ratio of 15 triangular:1 ovoid.

As you work through genetics problems, keep in mind that any single characteristic that results in a phenotypic ratio that totals 16 is typical of a two-gene interaction. Recall the phenotypic inheritance pattern for Mendel's dihybrid cross, which considered two non-interacting genes—9:3:3:1. Similarly, we would expect interacting gene pairs to also exhibit ratios expressed as 16 parts. Note that we are assuming the interacting genes are not linked; they are still assorting independently into gametes.

Link to Learning

For an excellent review of Mendel's experiments and to perform your own crosses and identify patterns of inheritance, visit the Mendel's Peas web lab.

Section Summary

Mendel postulated that genes (characteristics) are inherited as pairs of alleles (traits) that behave in a dominant and recessive pattern. Alleles segregate into gametes such that each gamete is equally likely to receive either one of the two alleles present in a diploid individual. In addition, genes are assorted into gametes independently of one another. That is, alleles are generally not more likely to segregate into a gamete with a particular allele of another gene. A dihybrid cross demonstrates independent assortment when the genes in question are on different chromosomes or distant from each other on the same chromosome. For crosses involving more than two genes, use the forked line or probability methods to predict offspring genotypes and phenotypes rather than a Punnett square.

Although chromosomes sort independently into gametes during meiosis, Mendel's law of independent assortment refers to genes, not chromosomes, and a single chromosome may carry more than 1,000 genes. When genes are located in close proximity on the same chromosome, their alleles tend to be inherited together. This results in offspring ratios that violate Mendel's law of independent assortment. However, recombination serves to exchange genetic material on homologous chromosomes such that maternal and paternal alleles may be recombined on the same chromosome. This is why alleles on a given chromosome are not always inherited together. Recombination is a random event occurring anywhere on a chromosome. Therefore, genes that are far apart on the same chromosome are likely to still assort independently because of recombination events that occurred in the intervening chromosomal space.

Whether or not they are sorting independently, genes may interact at the level of gene products such that the expression of an allele for one gene masks or modifies the expression of an allele for a different gene. This is called epistasis.

Visit this page in your course online to check your understanding.

Additional Self Check Questions

1. In pea plants, purple flowers (P) are dominant to white flowers (p) and yellow peas (Y) are dominant to green peas (y). What are the possible genotypes and phenotypes for a cross between PpYY and ppYy pea plants? How many squares do you need to do a Punnett square analysis of this cross?

2. Use the probability method to calculate the genotypes and genotypic proportions of a cross between *AABBCc* and *Aabbcc* parents.

3. Explain epistatis in terms of its Greek-language roots "standing upon."

4. In Section 12.3, "Laws of Inheritance," an example of epistasis was given for the summer squash. Cross white *WwYy* heterozygotes to prove the phenotypic ratio of 12 white:3 yellow:1 green that was given in the text.

Answers

1. The possible genotypes are PpYY, PpYy, ppYY, and ppYy. The former two genotypes would result in plants with purple flowers and yellow peas, while the latter two genotypes would result in plants with white flowers with yellow peas, for a 1:1 ratio of each phenotype. You only need a 2 × 2 Punnett square (four squares total) to do this analysis because two of the alleles are homozygous.

2. Considering each gene separately, the cross at *A* will produce offspring of which half are *AA* and half are *Aa*; *B* will produce all *Bb*; *C* will produce half *Cc* and half *cc*. Proportions then are (1/2) × (1) × (1/2), or 1/4 *AABbCc*; continuing for the other possibilities yields 1/4 *AABbcc*, 1/4 *AaBbCc*, and 1/4 *AaBbcc*. The proportions therefore are 1:1:1:1.

3. Epistasis describes an antagonistic interaction between genes wherein one gene masks or interferes with the expression of another. The gene that is interfering is referred to as epistatic, as if it is "standing upon" the other (hypostatic) gene to block its expression.

4. The cross can be represented as a 4 × 4 Punnett square, with the following gametes for each parent: *WY*, *Wy*, *wY*, and *wy*. For all 12 of the offspring that express a dominant *W* gene, the offspring will be white. The three offspring that are homozygous recessive for *w* but express a dominant *Y* gene will be yellow. The remaining *wwyy* offspring will be green.

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: Open Stax. **Located at**: http://cnx.org/contents/185cbf87-c72e-48f5-b51e-f14f21b5eabd@9.17:1/Biology. **License**: *CC BY: Attribution*

STUDY GUIDE: MENDELIAN GENETICS

Study Questions

Objectives: Determine the probable offspring of two parent organisms.

Use this page to check your understanding of the content.

Vocabulary

1. Phenotype
2. Genotype

3. Allele
4. Trait
5. Homozygous
6. Heterozygous
7. Dominant
8. Recessive
9. True breeding
10. Law of independent assortment
11. Law of segregation
12. Test cross

Study Guide Questions

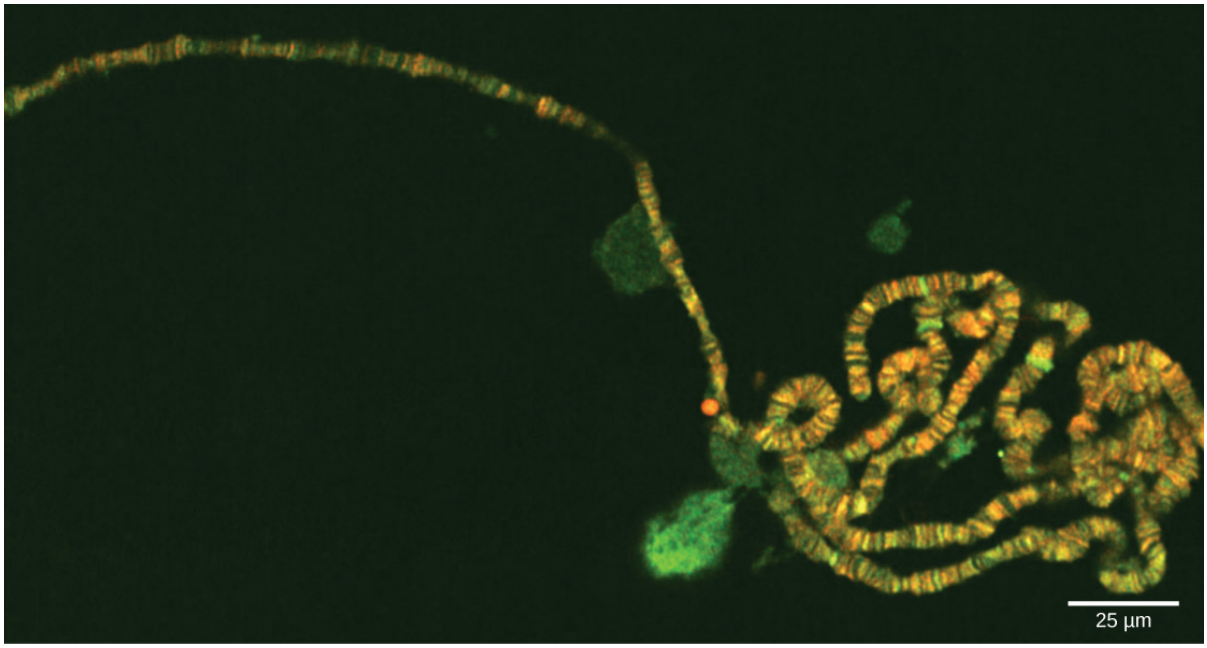
1. Understand Gregor Mendel's experiments, his results, and his conclusions.
2. Clearly relate MEIOSIS to Mendel's work.
3. Given data from a genetic cross, be able to determine information about how the trait in question is inherited.
4. Be able to successfully "do" both monohybrid and dihybrid crosses.
5. Understand the labels P, F1, F2 in reference to various genetic crosses.
6. What is a test cross? When is it used? Be able to interpret the results of a test cross.
7. Clearly connect the following 3 words: genotype, phenotype, protein

Licensing & Attributions

CC licensed content, Shared previously

- Study Guide: Mendelian Genetics. **Authored by**: Wendy Riggs. **Provided by**: College of the Redwoods. **Project**: Kaleidoscope. **License**: *CC BY: Attribution*

INTRODUCTION (MODERN UNDERSTANDING OF INHERITANCE)



Chromosomes are threadlike nuclear structures consisting of DNA and proteins that serve as the repositories for genetic information. The chromosomes depicted here were isolated from a fruit fly's salivary gland, stained with dye, and visualized under a microscope. Akin to miniature bar codes, chromosomes absorb different dyes to

produce characteristic banding patterns, which allows for their routine identification. (credit: modification of work by “LPLT”/Wikimedia Commons; scale-bar data from Matt Russell)

The gene is the physical unit of inheritance, and genes are arranged in a linear order on chromosomes. The behaviors and interactions of chromosomes during meiosis explain, at a cellular level, the patterns of inheritance that we observe in populations. Genetic disorders involving alterations in chromosome number or structure may have dramatic effects and can prevent a fertilized egg from developing altogether.

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: OpenStax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License**: *CC BY: Attribution*

CHROMOSOMAL THEORY OF INHERITANCE AND GENETIC LINKAGE

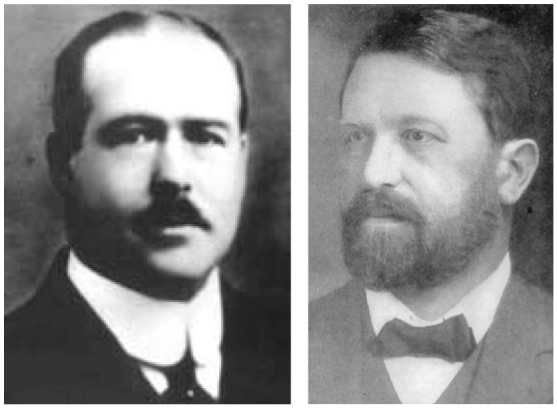
Objectives:
By the end of this section, you will be able to:

- Discuss Sutton’s Chromosomal Theory of Inheritance
- Describe genetic linkage
- Explain the process of homologous recombination, or crossing over
- Describe how chromosome maps are created
- Calculate the distances between three genes on a chromosome using a three-point test cross

Long before chromosomes were visualized under a microscope, the father of modern genetics, Gregor Mendel, began studying heredity in 1843. With the improvement of microscopic techniques during the late 1800s, cell biologists could stain and visualize subcellular structures with dyes and observe their actions during cell division and meiosis. With each mitotic division, chromosomes replicated, condensed from an amorphous (no constant shape) nuclear mass into distinct X-shaped bodies (pairs of identical sister chromatids), and migrated to separate cellular poles.

Chromosomal Theory of Inheritance

The speculation that chromosomes might be the key to understanding heredity led several scientists to examine Mendel’s publications and re-evaluate his model in terms of the behavior of chromosomes during mitosis and meiosis. In 1902, Theodor Boveri observed that proper embryonic development of sea urchins does not occur unless chromosomes are present. That same year, Walter Sutton observed the separation of chromosomes into daughter cells during meiosis (Figure). Together, these observations led to the development of the Chromosomal Theory of Inheritance, which identified chromosomes as the genetic material responsible for Mendelian inheritance.



(a) (b)

(a) Walter Sutton and (b) Theodor Boveri are credited with developing the Chromosomal Theory of Inheritance, which states that chromosomes carry the unit of heredity (genes).

The Chromosomal Theory of Inheritance was consistent with Mendel's laws and was supported by the following observations:

- During meiosis, homologous chromosome pairs migrate as discrete structures that are independent of other chromosome pairs.
- The sorting of chromosomes from each homologous pair into pre-gametes appears to be random.
- Each parent synthesizes gametes that contain only half of their chromosomal complement.
- Even though male and female gametes (sperm and egg) differ in size and morphology, they have the same number of chromosomes, suggesting equal genetic contributions from each parent.
- The gametic chromosomes combine during fertilization to produce offspring with the same chromosome number as their parents.

Despite compelling correlations between the behavior of chromosomes during meiosis and Mendel's abstract laws, the Chromosomal Theory of Inheritance was proposed long before there was any direct evidence that traits were carried on chromosomes. Critics pointed out that individuals had far more independently segregating traits than they had chromosomes. It was only after several years of carrying out crosses with the fruit fly, *Drosophila melanogaster*, that Thomas Hunt Morgan provided experimental evidence to support the Chromosomal Theory of Inheritance.

Genetic Linkage and Distances

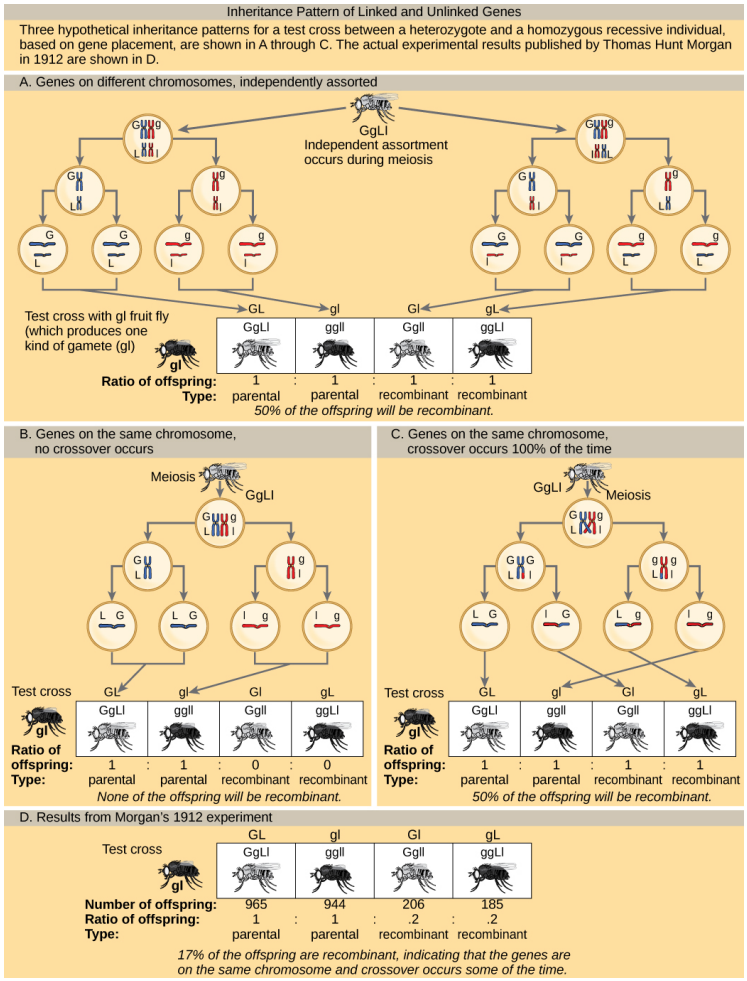
Mendel's work suggested that traits are inherited independently of each other. Morgan identified a 1:1 correspondence between a segregating trait and the X chromosome, suggesting that the random segregation of chromosomes was the physical basis of Mendel's model. This also demonstrated that linked genes disrupt Mendel's predicted outcomes. The fact that each chromosome can carry many linked genes explains how individuals can have many more traits than they have chromosomes. However, observations by researchers in Morgan's laboratory suggested that alleles positioned on the same chromosome were not always inherited together. During meiosis, linked genes somehow became unlinked.

Homologous Recombination

In 1909, Frans Janssen observed chiasmata—the point at which chromatids are in contact with each other and may exchange segments—prior to the first division of meiosis. He suggested that alleles become unlinked and chromosomes physically exchange segments. As chromosomes condensed and paired with their homologs, they appeared to interact at distinct points. Janssen suggested that these points corresponded to regions in which chromosome segments were exchanged. It is now known that the pairing and interaction between homologous chromosomes, known as synapsis, does more than simply organize the homologs for migration to separate daughter cells. When synapsed, homologous chromosomes undergo reciprocal physical exchanges at their arms in a process called homologous recombination, or more simply, "crossing over."

To better understand the type of experimental results that researchers were obtaining at this time, consider a heterozygous individual that inherited dominant maternal alleles for two genes on the same chromosome (such as *AB*) and two recessive paternal alleles for those same genes (such as *ab*). If the genes are linked, one would expect this individual to produce gametes that are either *AB* or *ab* with a 1:1 ratio. If the genes are unlinked, the individual should produce *AB*, *Ab*, *aB*, and *ab* gametes with equal frequencies, according to the Mendelian concept of independent assortment. Because they correspond to new allele combinations, the genotypes Ab and aB are nonparental types that result from homologous recombination during meiosis.Parental types are progeny that exhibit the same allelic combination as their parents. Morgan and his colleagues, however, found that when such heterozygous individuals were test crossed to a homozygous recessive parent (*AaBb* × *aabb*), both parental and nonparental cases occurred. For example, 950 offspring might be recovered that were either *AaBb* or *aabb*, but 50 offspring would also be obtained that were either *Aabb* or *aaBb*. These results suggested that linkage occurred most often, but a significant minority of offspring were the products of recombination.

ART CONNECTION



Inheritance patterns of unlinked and linked genes are shown. In (a), two genes are located on different chromosomes so independent assortment occurs during meiosis. The offspring have an equal chance of being the parental type (inheriting the same combination of traits as the parents) or a nonparental type (inheriting a different combination of traits than the parents). In (b), two genes are very close together on the same chromosome so that no crossing over occurs between them. The genes are therefore always inherited together and all of the offspring are the parental type. In (c), two genes are far apart on the chromosome such that crossing over occurs during every meiotic event. The recombination frequency will be the same as if the genes were on separate chromosomes. (d) The actual recombination frequency of fruit fly wing length and body color that Thomas Morgan observed in 1912 was 17 percent. A crossover frequency between 0 percent and 50 percent indicates that the genes are on the same chromosome and crossover occurs some of the time.

In a test cross for two characteristics such as the one shown here, can the predicted frequency of recombinant offspring be 60 percent? Why or why not?

Genetic Maps

Janssen did not have the technology to demonstrate crossing over so it remained an abstract idea that was not widely accepted. Scientists thought chiasmata were a variation on synapsis and could not understand how chromosomes could break and rejoin. Yet, the data were clear that linkage did not always occur. Ultimately, it took a young undergraduate student and an “all-nighter” to mathematically elucidate the problem of linkage and recombination.

In 1913, Alfred Sturtevant, a student in Morgan’s laboratory, gathered results from researchers in the laboratory, and took them home one night to mull them over. By the next morning, he had created the first “chromosome map,” a linear representation of gene order and relative distance on a chromosome (Figure).

ART CONNECTION

Genetic Map Based on Recombination Frequencies in *Drosophila*

MUTANT		WILD TYPE
Short aristae	0	Long aristae
Black body	48.5	Gray body
Cinnabar eyes	57.5	Red eyes
Vestigial wings	65.5	Normal wings
Brown eyes	104.5	Red eyes

Values in centimorgan (cM) map units; recombination frequency of 0.01 = 1 cM

This genetic map orders *Drosophila* genes on the basis of recombination frequency.

Which of the following statements is true?

1. Recombination of the body color and red/cinnabar eye alleles will occur more frequently than recombination of the alleles for wing length and aristae length.
2. Recombination of the body color and aristae length alleles will occur more frequently than recombination of red/brown eye alleles and the aristae length alleles.
3. Recombination of the gray/black body color and long/short aristae alleles will not occur.
4. Recombination of the red/brown eye and long/short aristae alleles will occur more frequently than recombination of the alleles for wing length and body color.

As shown in Figure, by using recombination frequency to predict genetic distance, the relative order of genes on chromosome 2 could be inferred. The values shown represent map distances in centimorgans (cM), which correspond to recombination frequencies (in percent). Therefore, the genes for body color and wing size were 65.5 – 48.5 = 17 cM apart, indicating that the maternal and paternal alleles for these genes recombine in 17 percent of offspring, on average.

To construct a chromosome map, Sturtevant assumed that genes were ordered serially on threadlike chromosomes. He also assumed that the incidence of recombination between two homologous chromosomes could occur with equal likelihood anywhere along the length of the chromosome. Operating under these assumptions, Sturtevant postulated that alleles that were far apart on a chromosome were more likely to dissociate during meiosis simply because there was a larger region over which recombination could occur. Conversely, alleles that were close to each other on the chromosome were likely to be inherited together. The average number of crossovers between two alleles—that is, their recombination frequency—correlated with their genetic distance from each other, relative to the locations of other genes on that chromosome. Considering the example cross between *AaBb* and *aabb* above, the frequency of recombination could be calculated as 50/1000 = 0.05. That is, the likelihood of a crossover between genes *A/a* and *B/b* was 0.05, or 5 percent. Such a result would indicate that the genes were definitively linked, but that they were far enough apart for crossovers to occasionally occur. Sturtevant divided his genetic map into map units, or centimorgans (cM), in which a recombination frequency of 0.01 corresponds to 1 cM.

By representing alleles in a linear map, Sturtevant suggested that genes can range from being perfectly linked (recombination frequency = 0) to being perfectly unlinked (recombination frequency = 0.5) when genes are on

different chromosomes or genes are separated very far apart on the same chromosome. Perfectly unlinked genes correspond to the frequencies predicted by Mendel to assort independently in a dihybrid cross. A recombination frequency of 0.5 indicates that 50 percent of offspring are recombinants and the other 50 percent are parental types. That is, every type of allele combination is represented with equal frequency. This representation allowed Sturtevant to additively calculate distances between several genes on the same chromosome. However, as the genetic distances approached 0.50, his predictions became less accurate because it was not clear whether the genes were very far apart on the same chromosome or on different chromosomes.

In 1931, Barbara McClintock and Harriet Creighton demonstrated the crossover of homologous chromosomes in corn plants. Weeks later, homologous recombination in *Drosophila* was demonstrated microscopically by Curt Stern. Stern observed several X-linked phenotypes that were associated with a structurally unusual and dissimilar X chromosome pair in which one X was missing a small terminal segment, and the other X was fused to a piece of the Y chromosome. By crossing flies, observing their offspring, and then visualizing the offspring's chromosomes, Stern demonstrated that every time the offspring allele combination deviated from either of the parental combinations, there was a corresponding exchange of an X chromosome segment. Using mutant flies with structurally distinct X chromosomes was the key to observing the products of recombination because DNA sequencing and other molecular tools were not yet available. It is now known that homologous chromosomes regularly exchange segments in meiosis by reciprocally breaking and rejoining their DNA at precise locations.

Link to Learning

Review Sturtevant's process to create a genetic map on the basis of recombination frequencieshere.

Mendel's Mapped Traits

Homologous recombination is a common genetic process, yet Mendel never observed it. Had he investigated both linked and unlinked genes, it would have been much more difficult for him to create a unified model of his data on the basis of probabilistic calculations. Researchers who have since mapped the seven traits investigated by Mendel onto the seven chromosomes of the pea plant genome have confirmed that all of the genes he examined are either on separate chromosomes or are sufficiently far apart as to be statistically unlinked. Some have suggested that Mendel was enormously lucky to select only unlinked genes, whereas others question whether Mendel discarded any data suggesting linkage. In any case, Mendel consistently observed independent assortment because he examined genes that were effectively unlinked.

Section Summary

The Chromosomal Theory of inheritance, proposed by Sutton and Boveri, states that chromosomes are the vehicles of genetic heredity. Neither Mendelian genetics nor gene linkage is perfectly accurate; instead, chromosome behavior involves segregation, independent assortment, and occasionally, linkage. Sturtevant devised a method to assess recombination frequency and infer the relative positions and distances of linked genes on a chromosome on the basis of the average number of crossovers in the intervening region between the genes. Sturtevant correctly presumed that genes are arranged in serial order on chromosomes and that recombination between homologs can occur anywhere on a chromosome with equal likelihood. Whereas linkage causes alleles on the same chromosome to be inherited together, homologous recombination biases alleles toward an inheritance pattern of independent assortment.

Art Connections

Figure In a test cross for two characteristics such as the one shown here, can the predicted frequency of recombinant offspring be 60 percent? Why or why not?

Figure Which of the following statements is true?

1. Recombination of the body color and red/cinnabar eye alleles will occur more frequently than recombination of the alleles for wing length and aristae length.

2. Recombination of the body color and aristae length alleles will occur more frequently than recombination of red/brown eye alleles and the aristae length alleles.
3. Recombination of the gray/black body color and long/short aristae alleles will not occur.
4. Recombination of the red/brown eye and long/short aristae alleles will occur more frequently than recombination of the alleles for wing length and body color.

Review Questions

X-linked recessive traits in humans (or in *Drosophila*) are observed _____.

1. in more males than females
2. in more females than males
3. in males and females equally
4. in different distributions depending on the trait

The first suggestion that chromosomes may physically exchange segments came from the microscopic identification of _____.

1. synapsis
2. sister chromatids
3. chiasmata
4. alleles

Which recombination frequency corresponds to independent assortment and the absence of linkage?

1. 0
2. 0.25
3. 0.50
4. 0.75

Which recombination frequency corresponds to perfect linkage and violates the law of independent assortment?

1. 0
2. 0.25
3. 0.50
4. 0.75

Free Response

Explain how the Chromosomal Theory of Inheritance helped to advance our understanding of genetics.

Glossary

centimorgan (cM) (also, map unit) relative distance that corresponds to a recombination frequency of 0.01

Chromosomal Theory of Inheritance theory proposing that chromosomes are the vehicles of genes and that their behavior during meiosis is the physical basis of the inheritance patterns that Mendel observed

homologous recombination process by which homologous chromosomes undergo reciprocal physical exchanges at their arms, also known as crossing over

nonparental (recombinant) type progeny resulting from homologous recombination that exhibits a different allele combination compared with its parents

parental types progeny that exhibits the same allelic combination as its parents

recombination frequency average number of crossovers between two alleles; observed as the number of nonparental types in a population of progeny

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: OpenStax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License**: *CC BY: Attribution*

CHROMOSOMAL BASIS OF INHERITED DISORDERS

Learning Objectives

By the end of this section, you will be able to:

- Describe how a karyogram is created
- Explain how nondisjunction leads to disorders in chromosome number
- Compare disorders caused by aneuploidy
- Describe how errors in chromosome structure occur through inversions and translocations

Inherited disorders can arise when chromosomes behave abnormally during meiosis. Chromosome disorders can be divided into two categories: abnormalities in chromosome number and chromosomal structural rearrangements. Because even small segments of chromosomes can span many genes, chromosomal disorders are characteristically dramatic and often fatal.

Identification of Chromosomes

The isolation and microscopic observation of chromosomes forms the basis of cytogenetics and is the primary method by which clinicians detect chromosomal abnormalities in humans. A karyotype is the number and appearance of chromosomes, and includes their length, banding pattern, and centromere position. To obtain a view of an individual's karyotype, cytologists photograph the chromosomes and then cut and paste each chromosome into a chart, or karyogram, also known as an ideogram (Figure 1).

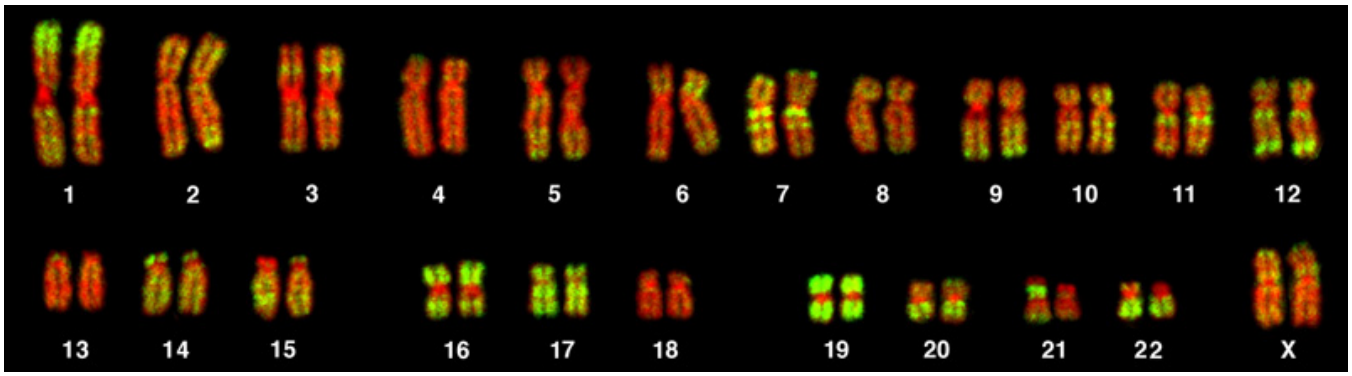


Figure 1. This karyotype is of a female human. Notice that homologous chromosomes are the same size, and have the same centromere positions and banding patterns. A human male would have an XY chromosome pair instead of the XX pair shown. (credit: Andreas Blozer et al)

In a given species, chromosomes can be identified by their number, size, centromere position, and banding pattern. In a human karyotype, autosomes or "body chromosomes" (all of the non–sex chromosomes) are generally organized in approximate order of size from largest (chromosome 1) to smallest (chromosome 22). The X and Y chromosomes are not autosomes. However, chromosome 21 is actually shorter than chromosome 22. This was discovered after the naming of Down syndrome as trisomy 21, reflecting how this disease results from

possessing one extra chromosome 21 (three total). Not wanting to change the name of this important disease, chromosome 21 retained its numbering, despite describing the shortest set of chromosomes. The chromosome "arms" projecting from either end of the centromere may be designated as short or long, depending on their relative lengths. The short arm is abbreviated *p* (for "petite"), whereas the long arm is abbreviated *q* (because it follows "p" alphabetically). Each arm is further subdivided and denoted by a number. Using this naming system, locations on chromosomes can be described consistently in the scientific literature.

Career Connection

Geneticists Use Karyograms to Identify Chromosomal Aberrations

Although Mendel is referred to as the "father of modern genetics," he performed his experiments with none of the tools that the geneticists of today routinely employ. One such powerful cytological technique is karyotyping, a method in which traits characterized by chromosomal abnormalities can be identified from a single cell. To observe an individual's karyotype, a person's cells (like white blood cells) are first collected from a blood sample or other tissue. In the laboratory, the isolated cells are stimulated to begin actively dividing. A chemical called colchicine is then applied to cells to arrest condensed chromosomes in metaphase. Cells are then made to swell using a hypotonic solution so the chromosomes spread apart. Finally, the sample is preserved in a fixative and applied to a slide.

The geneticist then stains chromosomes with one of several dyes to better visualize the distinct and reproducible banding patterns of each chromosome pair. Following staining, the chromosomes are viewed using bright-field microscopy. A common stain choice is the Giemsa stain. Giemsa staining results in approximately 400–800 bands (of tightly coiled DNA and condensed proteins) arranged along all of the 23 chromosome pairs; an experienced geneticist can identify each band. In addition to the banding patterns, chromosomes are further identified on the basis of size and centromere location. To obtain the classic depiction of the karyotype in which homologous pairs of chromosomes are aligned in numerical order from longest to shortest, the geneticist obtains a digital image, identifies each chromosome, and manually arranges the chromosomes into this pattern (Figure).

At its most basic, the karyogram may reveal genetic abnormalities in which an individual has too many or too few chromosomes per cell. Examples of this are Down Syndrome, which is identified by a third copy of chromosome 21, and Turner Syndrome, which is characterized by the presence of only one X chromosome in women instead of the normal two. Geneticists can also identify large deletions or insertions of DNA. For instance, Jacobsen Syndrome—which involves distinctive facial features as well as heart and bleeding defects—is identified by a deletion on chromosome 11. Finally, the karyotype can pinpoint translocations, which occur when a segment of genetic material breaks from one chromosome and reattaches to another chromosome or to a different part of the same chromosome. Translocations are implicated in certain cancers, including chronic myelogenous leukemia.

During Mendel's lifetime, inheritance was an abstract concept that could only be inferred by performing crosses and observing the traits expressed by offspring. By observing a karyogram, today's geneticists can actually visualize the chromosomal composition of an individual to confirm or predict genetic abnormalities in offspring, even before birth.

Disorders in Chromosome Number

Of all of the chromosomal disorders, abnormalities in chromosome number are the most obviously identifiable from a karyogram. Disorders of chromosome number include the duplication or loss of entire chromosomes, as well as changes in the number of complete sets of chromosomes. They are caused by nondisjunction, which occurs when pairs of homologous chromosomes or sister chromatids fail to separate during meiosis. Misaligned or incomplete synapsis, or a dysfunction of the spindle apparatus that facilitates chromosome migration, can cause nondisjunction. The risk of nondisjunction occurring increases with the age of the parents.

Nondisjunction can occur during either meiosis I or II, with differing results (Figure 2). If homologous chromosomes fail to separate during meiosis I, the result is two gametes that lack that particular chromosome and

two gametes with two copies of the chromosome. If sister chromatids fail to separate during meiosis II, the result is one gamete that lacks that chromosome, two normal gametes with one copy of the chromosome, and one gamete with two copies of the chromosome.

Art Connection

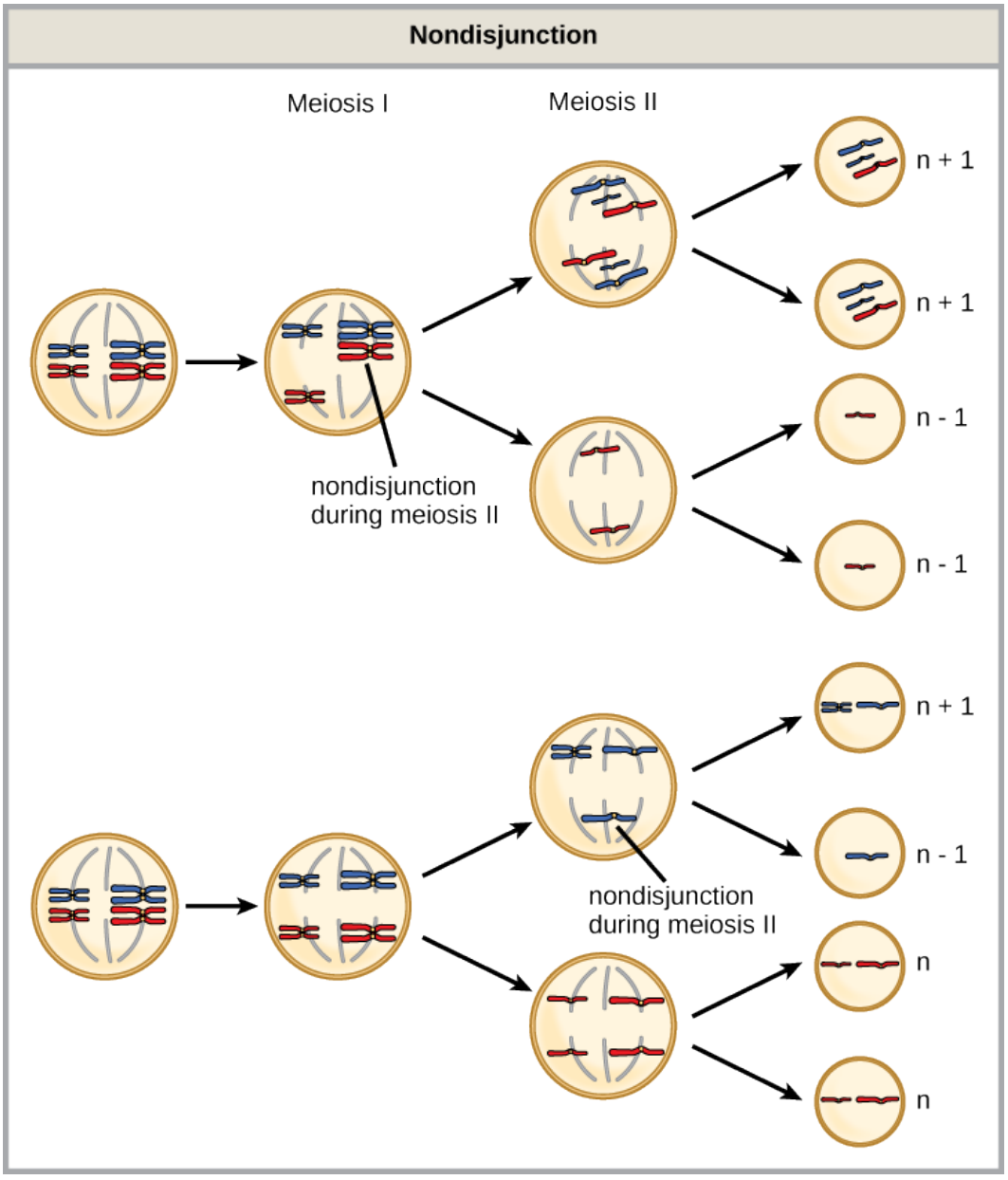


Figure 2. Nondisjunction occurs when homologous chromosomes or sister chromatids fail to separate during meiosis, resulting in an abnormal chromosome number. Nondisjunction may occur during meiosis I or meiosis II.

Which of the following statements about nondisjunction is true?

1. Nondisjunction only results in gametes with n+1 or n–1 chromosomes.
2. Nondisjunction occurring during meiosis II results in 50 percent normal gametes.
3. Nondisjunction during meiosis I results in 50 percent normal gametes.

4. Nondisjunction always results in four different kinds of gametes.

Aneuploidy

An individual with the appropriate number of chromosomes for their species is called euploid; in humans, euploidy corresponds to 22 pairs of autosomes and one pair of sex chromosomes. An individual with an error in chromosome number is described as aneuploid, a term that includes monosomy (loss of one chromosome) or trisomy (gain of an extraneous chromosome). Monosomic human zygotes missing any one copy of an autosome invariably fail to develop to birth because they lack essential genes. This underscores the importance of "gene dosage" in humans. Most autosomal trisomies also fail to develop to birth; however, duplications of some of the smaller chromosomes (13, 15, 18, 21, or 22) can result in offspring that survive for several weeks to many years. Trisomic individuals suffer from a different type of genetic imbalance: an excess in gene dose. Individuals with an extra chromosome may synthesize an abundance of the gene products encoded by that chromosome. This extra dose (150 percent) of specific genes can lead to a number of functional challenges and often precludes development. The most common trisomy among viable births is that of chromosome 21, which corresponds to Down Syndrome. Individuals with this inherited disorder are characterized by short stature and stunted digits, facial distinctions that include a broad skull and large tongue, and significant developmental delays. The incidence of Down syndrome is correlated with maternal age; older women are more likely to become pregnant with fetuses carrying the trisomy 21 genotype (Figure 3).

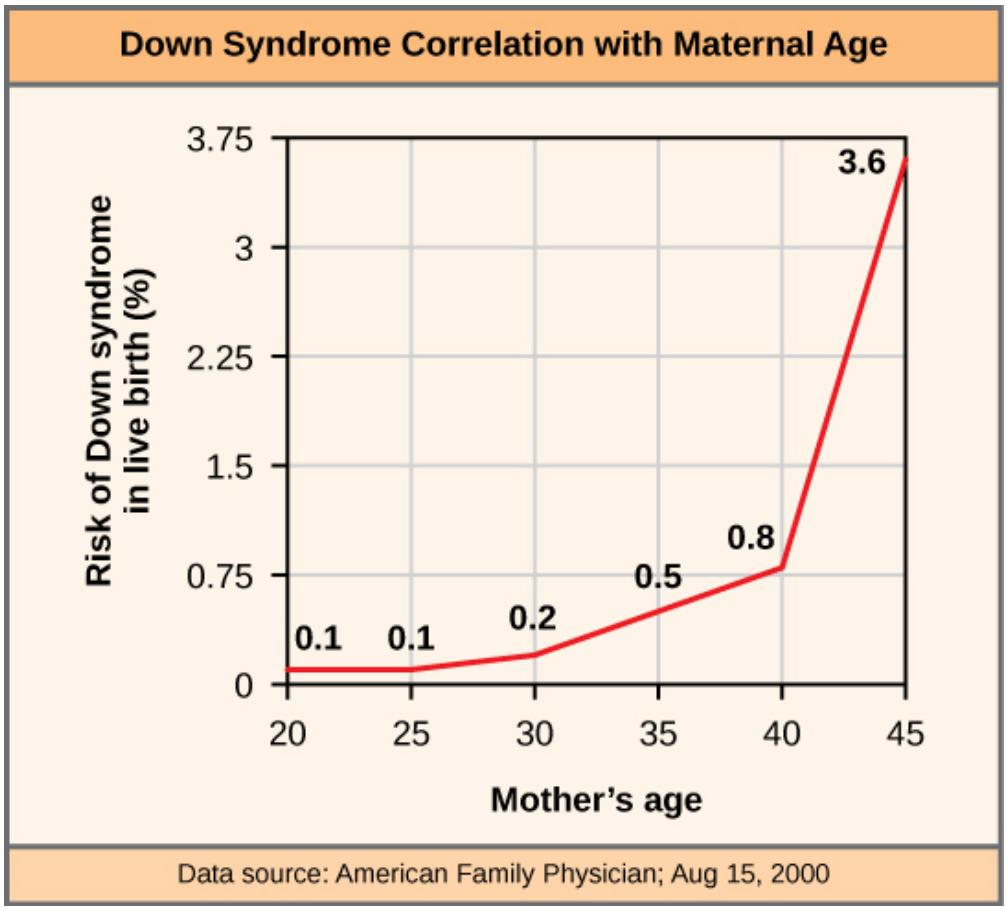


Figure 3. The incidence of having a fetus with trisomy 21 increases dramatically with maternal age.

Link to Learning

Visualize the addition of a chromosome that leads to Down syndrome in this video simulation.

Polyploidy

An individual with more than the correct number of chromosome sets (two for diploid species) is called polyploid. For instance, fertilization of an abnormal diploid egg with a normal haploid sperm would yield a triploid zygote. Polyploid animals are extremely rare, with only a few examples among the flatworms, crustaceans, amphibians, fish, and lizards. Polyploid animals are sterile because meiosis cannot proceed normally and instead produces mostly aneuploid daughter cells that cannot yield viable zygotes. Rarely, polyploid animals can reproduce asexually by haplodiploidy, in which an unfertilized egg divides mitotically to produce offspring. In contrast, polyploidy is very common in the plant kingdom, and polyploid plants tend to be larger and more robust than euploids of their species (Figure 4).

Figure 4. As with many polyploid plants, this triploid orange daylily (Hemerocallis fulva) is particularly large and robust, and grows flowers with triple the number of petals of its diploid counterparts. (credit: Steve Karg)

Sex Chromosome Nondisjunction in Humans

Humans display dramatic deleterious effects with autosomal trisomies and monosomies. Therefore, it may seem counterintuitive that human females and males can function normally, despite carrying different numbers of the X chromosome. Rather than a gain or loss of autosomes, variations in the number of sex chromosomes are associated with relatively mild effects. In part, this occurs because of a molecular process called X inactivation. Early in development, when female mammalian embryos consist of just a few thousand cells (relative to trillions in the newborn), one X chromosome in each cell inactivates by tightly condensing into a quiescent (dormant) structure called a Barr body. The chance that an X chromosome (maternally or paternally derived) is inactivated in each cell is random, but once the inactivation occurs, all cells derived from that one will have the same inactive X chromosome or Barr body. By this process, females compensate for their double genetic dose of X chromosome. In so-called “tortoiseshell” cats, embryonic X inactivation is observed as color variegation (Figure 5). Females that are heterozygous for an X-linked coat color gene will express one of two different coat colors over different regions of their body, corresponding to whichever X chromosome is inactivated in the embryonic cell progenitor of that region.

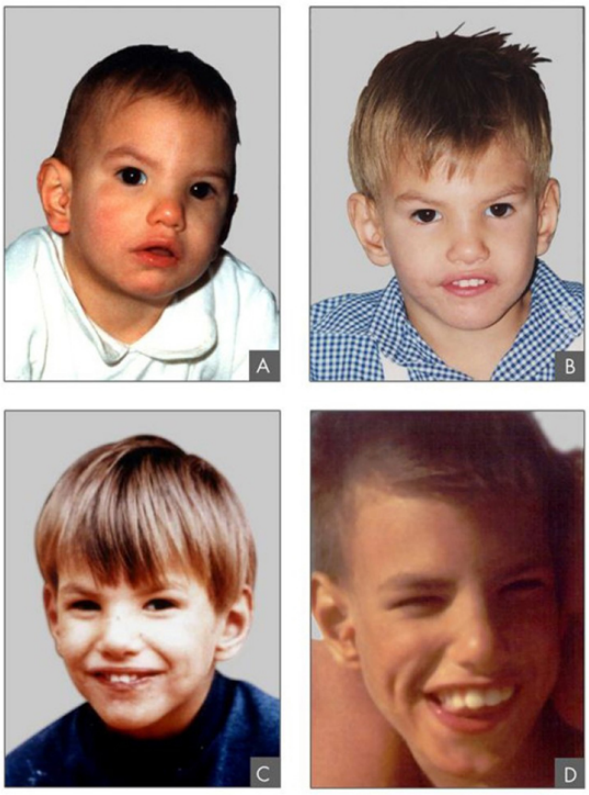
An individual carrying an abnormal number of X chromosomes will inactivate all but one X chromosome in each of her cells. However, even inactivated X chromosomes continue to express a few genes, and X chromosomes must reactivate for the proper maturation of female ovaries. As a result, X-chromosomal abnormalities are typically associated with mild mental and physical defects, as well as sterility. If the X chromosome is absent altogether, the individual will not develop in utero.

Several errors in sex chromosome number have been characterized. Individuals with three X chromosomes, called triplo-X, are phenotypically female but express developmental delays and reduced fertility. The XXY genotype, corresponding to one type of Klinefelter syndrome, corresponds to phenotypically male individuals with small testes, enlarged breasts, and reduced body hair. More complex types of Klinefelter syndrome exist in which the individual has as many as five X chromosomes. In all types, every X chromosome except one undergoes inactivation to compensate for the excess genetic dosage. This can be seen as several Barr bodies in each cell nucleus. Turner syndrome, characterized as an X0 genotype (i.e., only a single sex chromosome), corresponds to a phenotypically female individual with short stature, webbed skin in the neck region, hearing and cardiac impairments, and sterility.

Figure 5. In cats, the gene for coat color is located on the X chromosome. In the embryonic development of female cats, one of the two X chromosomes is randomly inactivated in each cell, resulting in a tortoiseshell pattern if the cat has two different alleles for coat color. Male cats, having only one X chromosome, never exhibit a tortoiseshell coat color. (credit: Michael Bodega)

Duplications and Deletions

In addition to the loss or gain of an entire chromosome, a chromosomal segment may be duplicated or lost. Duplications and deletions often produce offspring that survive but exhibit physical and mental abnormalities. Duplicated chromosomal segments may fuse to existing chromosomes or may be free in the nucleus. Cri-du-chat (from the French for “cry of the cat”) is a syndrome associated with nervous system abnormalities and identifiable physical features that result from a deletion of most of 5p (the small arm of chromosome 5) (Figure 6). Infants with this genotype emit a characteristic high-pitched cry on which the disorder’s name is based.



Chromosomal Structural Rearrangements

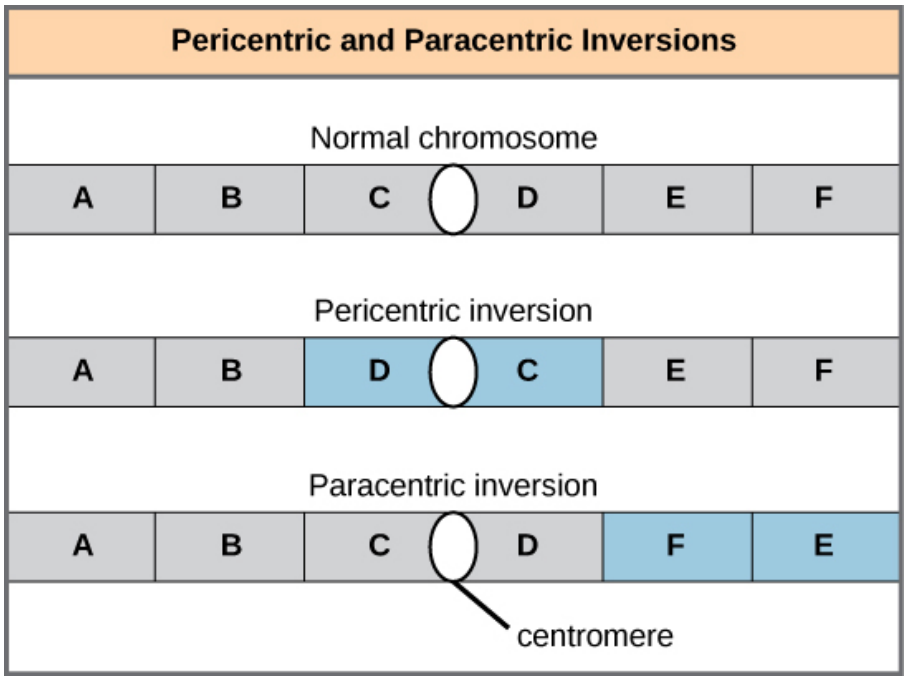
Figure 6. This individual with cri-du-chat syndrome is shown at two, four, nine, and 12 years of age. (credit: Paola Cerruti Mainardi)

Cytologists have characterized numerous structural rearrangements in chromosomes, but chromosome inversions and translocations are the most common. Both are identified during meiosis by the adaptive pairing of rearranged chromosomes with their former homologs to maintain appropriate gene alignment. If the genes carried on two homologs are not oriented correctly, a recombination event could result in the loss of genes from one chromosome and the gain of genes on the other. This would produce aneuploid gametes.

Chromosome Inversions

A chromosome inversion is the detachment, 180° rotation, and reinsertion of part of a chromosome. Inversions may occur in nature as a result of mechanical shear, or from the action of transposable elements (special DNA sequences capable of facilitating the rearrangement of chromosome segments with the help of enzymes that cut and paste DNA sequences). Unless they disrupt a gene sequence, inversions only change the orientation of genes and are likely to have more mild effects than aneuploid errors. However, altered gene orientation can result in functional changes because regulators of gene expression could be moved out of position with respect to their targets, causing aberrant levels of gene products.

An inversion can be pericentric and include the centromere, or paracentric and occur outside of the centromere (Figure 7). A pericentric inversion that is asymmetric about the centromere can change the relative lengths of the chromosome arms, making these inversions easily identifiable.



Pericentric inversions include the centromere, and paracentric inversions do not. A pericentric inversion can change the relative lengths of the chromosome arms; a paracentric inversion cannot.

When one homologous chromosome undergoes an inversion but the other does not, the individual is described as an inversion heterozygote. To maintain point-for-point synapsis during meiosis, one homolog must form a loop, and the other homolog must mold around it. Although this topology can ensure that the genes are correctly aligned, it also forces the homologs to stretch and can be associated with regions of imprecise synapsis (Figure 8).

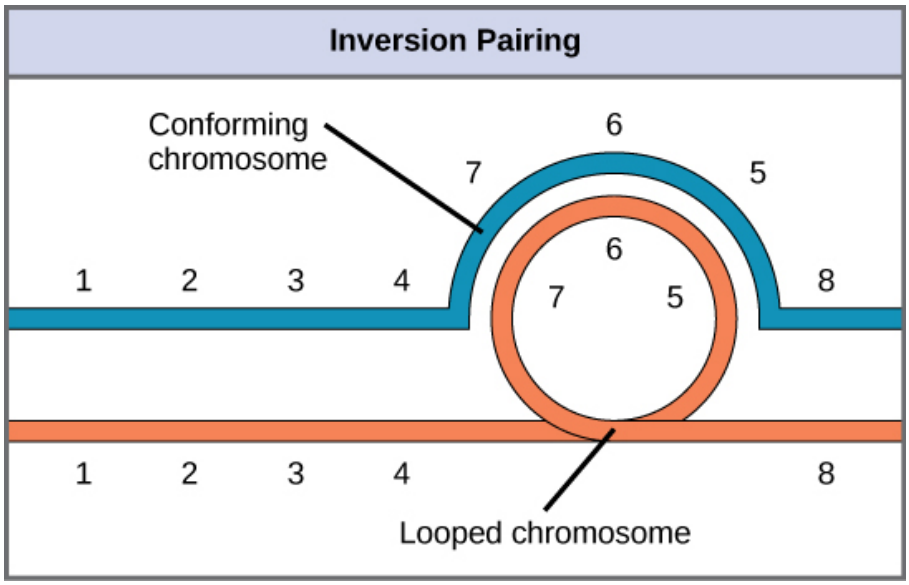


Figure 8. When one chromosome undergoes an inversion but the other does not, one chromosome must form an inverted loop to retain point-for-point interaction during synapsis. This inversion pairing is essential to maintaining gene alignment during meiosis and to allow for recombination.

Evolution Connection

The Chromosome 18 Inversion

Not all structural rearrangements of chromosomes produce nonviable, impaired, or infertile individuals. In rare instances, such a change can result in the evolution of a new species. In fact, a pericentric inversion in chromosome 18 appears to have contributed to the evolution of humans. This inversion is not present in our closest genetic relatives, the chimpanzees. Humans and chimpanzees differ cytogenetically by pericentric inversions on several chromosomes and by the fusion of two separate chromosomes in chimpanzees that correspond to chromosome two in humans.

The pericentric chromosome 18 inversion is believed to have occurred in early humans following their divergence from a common ancestor with chimpanzees approximately five million years ago. Researchers characterizing this inversion have suggested that approximately 19,000 nucleotide bases were duplicated on 18p, and the duplicated region inverted and reinserted on chromosome 18 of an ancestral human.

A comparison of human and chimpanzee genes in the region of this inversion indicates that two genes—*ROCK1* and *USP14*—that are adjacent on chimpanzee chromosome 17 (which corresponds to human chromosome 18) are more distantly positioned on human chromosome 18. This suggests that one of the inversion breakpoints occurred between these two genes. Interestingly, humans and chimpanzees express *USP14* at distinct levels in specific cell types, including cortical cells and fibroblasts. Perhaps the chromosome 18 inversion in an ancestral human repositioned specific genes and reset their expression levels in a useful way. Because both *ROCK1* and *USP14* encode cellular enzymes, a change in their expression could alter cellular function. It is not known how this inversion contributed to hominid evolution, but it appears to be a significant factor in the divergence of humans from other primates. ((Note: Violaine Goidts et al., "Segmental duplication associated with the human-specific inversion of chromosome 18: a further example of the impact of segmental duplications on karyotype and genome evolution in primates," *Human Genetics*. 115 (2004):116–122))

Translocations

A translocation occurs when a segment of a chromosome dissociates and reattaches to a different, nonhomologous chromosome. Translocations can be benign or have devastating effects depending on how the positions of genes are altered with respect to regulatory sequences. Notably, specific translocations have been associated with several cancers and with schizophrenia. Reciprocal translocations result from the exchange of chromosome segments between two nonhomologous chromosomes such that there is no gain or loss of genetic information (Figure 9).

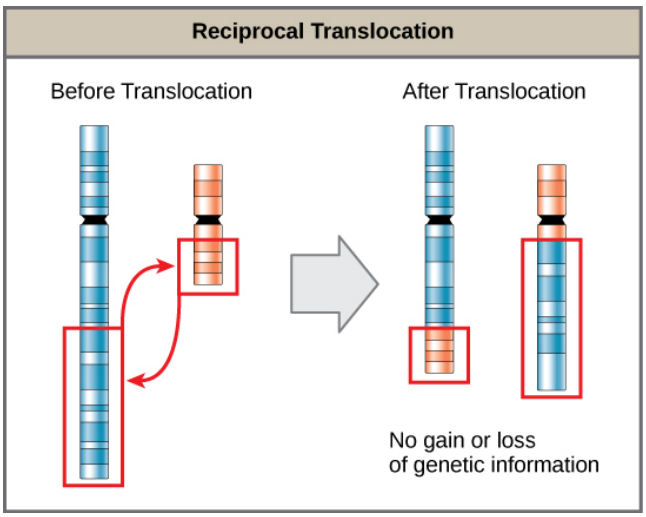


Figure 9. A reciprocal translocation occurs when a segment of DNA is transferred from one chromosome to another, nonhomologous chromosome. (credit: modification of work by National Human Genome Research/USA)

Section Summary

The number, size, shape, and banding pattern of chromosomes make them easily identifiable in a karyogram and allows for the assessment of many chromosomal abnormalities. Disorders in chromosome number, or aneuploidies, are typically lethal to the embryo, although a few trisomic genotypes are viable. Because of X inactivation, aberrations in sex chromosomes typically have milder phenotypic effects. Aneuploidies also include instances in which segments of a chromosome are duplicated or deleted. Chromosome structures may also be rearranged, for example by inversion or translocation. Both of these aberrations can result in problematic phenotypic effects. Because they force chromosomes to assume unnatural topologies during meiosis, inversions and translocations are often associated with reduced fertility because of the likelihood of nondisjunction.

Self Check Questions

1. Which of the following statements about nondisjunction is true?

    A. Nondisjunction only results in gametes with n+1 or n–1 chromosomes.

    B. Nondisjunction occurring during meiosis II results in 50 percent normal gametes.

    C. Nondisjunction during meiosis I results in 50 percent normal gametes.

    D. Nondisjunction always results in four different kinds of gametes.

2. Using diagrams, illustrate how nondisjunction can result in an aneuploid zygote.

Answers

1. B

2. Exact diagram style will vary; diagram should look like Figure.

Glossary

aneuploid: individual with an error in chromosome number; includes deletions and duplications of chromosome segments

autosome: any of the non-sex chromosomes

chromosome inversion: detachment, 180° rotation, and reinsertion of a chromosome arm

euploid: individual with the appropriate number of chromosomes for their species

karyogram: photographic image of a karyotype

karyotype: number and appearance of an individuals chromosomes; includes the size, banding patterns, and centromere position

monosomy: otherwise diploid genotype in which one chromosome is missing

nondisjunction: failure of synapsed homologs to completely separate and migrate to separate poles during the first cell division of meiosis

paracentric: inversion that occurs outside of the centromere

pericentric: inversion that involves the centromere

polyploid: individual with an incorrect number of chromosome sets

translocation: process by which one segment of a chromosome dissociates and reattaches to a different, nonhomologous chromosome

trisomy: otherwise diploid genotype in which one entire chromosome is duplicated

X inactivation: condensation of X chromosomes into Barr bodies during embryonic development in females to compensate for the double genetic dose

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: Open Stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License**: *CC BY: Attribution*

STUDY GUIDE: CHROMOSOMAL INHERITANCE

Study Questions

Objective: Solve problems involving heritable conditions related to chromosomal inheritance.

Use this page to check your understanding of the content.

Vocabulary

1. Karyotype
2. Autosomes
3. Sex chromosomes
4. Aneuploidy

Study Guide Questions

1. Be able to read and interpret any karyotype, including determining the chromosomal gender of the “patient”.

2. Successfully solve heredity problems involving sex-linked characteristics.
3. Example: Hemophilia in humans is a condition resulting in a failure to form blood clots, resulting in excessive bleeding. It is due to a mutation in a gene found on the X-chromosome. What possible offspring can result from a baby-making union between a non-carrier female and a hemophilac male?
4. For more practice, visit: http://www.biology.arizona.edu/mendelian_genetics/problem_sets/sex_linked_inheritance/sex_linked_inheritance.html (Links to an external site.)
5. When given an example of aneuploidy, be able to hypothesize how it occurred. (Did non-disjunction take place during meiosis I or meiosis II? How can you tell? Can you always tell?)
6. Be able to read and interpret pedigrees.
    1. In this pedigree, the shaded individuals are homozygous recessive. What is the genotype of individual B?
    2. What is the genotype of individual E?
    3. If E married an individual who is homozygous recessive, what is the probability that their first child will be homozygous recessive?
    4. For practice, see: http://bio3400.nicerweb.com/med/QUIZ/pedigree_q.html
    5. Why are pedigrees necessary when studying human inheritance patterns?

Licensing & Attributions

CC licensed content, Shared previously

- Study Guide: Chromosomal Inheritance. **Authored by**: Wendy Riggs. **Provided by**: College of the Redwoods. **Project**: Kaleidoscope. **License**: *CC BY: Attribution*

STUDY GUIDE: NON-MENDELIAN GENETICS

Study Questions

Objective: Describe inheritance patterns that do not follow Mendelian patterns.

Use this page to check your understanding of the content.

Vocabulary

1. Epistasis
2. Pleiotropy
3. Polygenic inheritance
4. Penetrance
5. Incomplete dominance
6. Codominance
7. Epigenetics

Study Guide Questions

1. What is incomplete dominance? How does this affect the phenotypic ratios of the offspring in various crosses?
2. Give an example of incomplete dominance.
3. What is co-dominance? How does this affect the expected phenotypic ratios of the offspring in various crosses?
4. Give an example of co-dominance.
5. Understand the genetics of blood types.
6. What is penetrance? How does this affect the expected phenotypic ratios of the offspring in various crosses?
7. What is epigenetics? What significance does this have in the study of inheritance?
8. Compare and contrast all of the following situations, and give an example of each:
    1. Multiple alleles
    2. polygenic inheritance
    3. epistasis

4. pleiotropy
5. the effect of the environment on phenotype

Licensing & Attributions

CC licensed content, Shared previously

- Study Guide: Non-Mendelian Genetics. **Authored by**: Wendy Riggs. **Provided by**: College of the Redwoods. **Project**: Kaleidoscope. **License**: *CC BY: Attribution*

STARGENETICS

StarGenetics provides a set of tools for analyzing genetic traits. This software simulates mating experiments between organisms that are genetically different across a range of traits and allows students to analyze the nature of the traits in question.
Get Started with StarGenetics.

Licensing & Attributions

CC licensed content, Shared previously

- StarGenetics. **Authored by**: Chris Kaiser. **Provided by**: OEIT. **Located at**: http://oeit.mit.edu/gallery/projects/stargenetics. **Project**: Gallery of Educational Innovation. **License**: *CC BY-NC: Attribution-NonCommercial*

INTRODUCTION TO DARWIN AND EVOLUTION

INTRODUCTION: DARWIN AND EVOLUTION

All species of living organisms, from bacteria to baboons to blueberries, evolved at some point from a different species. Although it may seem that living things today stay much the same, that is not the case—evolution is an ongoing process.

The theory of evolution is the unifying theory of biology, meaning it is the framework within which biologists ask questions about the living world. Its power is that it provides direction for predictions about living things that are borne out in experiment after experiment. The Ukrainian-born American geneticist Theodosius Dobzhansky famously wrote that “nothing makes sense in biology except in the light of evolution.” ((Note: Theodosius Dobzhansky. “Biology, Molecular and Organismic.” *American Zoologist* 4, no. 4 (1964): 449.)) He meant that the tenet that all life has evolved and diversified from a common ancestor is the foundation from which we approach all questions in biology.

Figure 1. All organisms are products of evolution adapted to their environment. (a) Saguaro (Carnegiea gigantea) can soak up 750 liters of water in a single rain storm, enabling these cacti to survive the dry conditions of the Sonora desert in Mexico and the Southwestern United States. (b) The Andean semiaquatic lizard (Potamites montanicola) discovered in Peru in 2010 lives between 1,570 to 2,100 meters in elevation, and, unlike most lizards, is nocturnal and swims. Scientists still do no know how these cold-blood animals are able to move in the cold (10 to 15°C) temperatures of the Andean night. (credit a: modification of work by Gentry George, U.S. Fish and Wildlife Service; credit b: modification of work by Germán Chávez and Diego Vásquez, ZooKeys)

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: Open Stax. **Located at**: http://cnx.org/contents/185cbf87-c72e-48f5-b51e-f14f21b5eabd@9.17:1/Biology. **License**: *CC BY: Attribution*

UNDERSTANDING EVOLUTION

Learning Objectives

By the end of this section, you will be able to:

- Describe how the present-day theory of evolution was developed
- Define adaptation
- Explain convergent and divergent evolution
- Describe homologous and vestigial structures
- Discuss misconceptions about the theory of evolution

Evolution by natural selection describes a mechanism for how species change over time. That species change had been suggested and debated well before Darwin began to explore this idea. The view that species were static and unchanging was grounded in the writings of Plato, yet there were also ancient Greeks who expressed evolutionary ideas. In the eighteenth century, ideas about the evolution of animals were reintroduced by the naturalist Georges-Louis Leclerc Comte de Buffon who observed that various geographic regions have different plant and animal populations, even when the environments are similar. It was also accepted that there were extinct species.

During this time, James Hutton, a Scottish naturalist, proposed that geological change occurred gradually by the accumulation of small changes from processes operating like they are today over long periods of time. This contrasted with the predominant view that the geology of the planet was a consequence of catastrophic events occurring during a relatively brief past. Hutton's view was popularized in the nineteenth century by the geologist Charles Lyell who became a friend to Darwin. Lyell's ideas were influential on Darwin's thinking: Lyell's notion of the greater age of Earth gave more time for gradual change in species, and the process of change provided an analogy for gradual change in species. In the early nineteenth century, Jean-Baptiste Lamarck published a book that detailed a mechanism for evolutionary change. This mechanism is now referred to as an inheritance of acquired characteristics by which modifications in an individual are caused by its environment, or the use or disuse of a structure during its lifetime, could be inherited by its offspring and thus bring about change in a species. While this mechanism for evolutionary change was discredited, Lamarck's ideas were an important influence on evolutionary thought.

Charles Darwin and Natural Selection

In the mid-nineteenth century, the actual mechanism for evolution was independently conceived of and described by two naturalists: Charles Darwin and Alfred Russel Wallace. Importantly, each naturalist spent time exploring the natural world on expeditions to the tropics. From 1831 to 1836, Darwin traveled around the world on *H.M.S. Beagle*, including stops in South America, Australia, and the southern tip of Africa. Wallace traveled to Brazil to collect insects in the Amazon rainforest from 1848 to 1852 and to the Malay Archipelago from 1854 to 1862. Darwin's journey, like Wallace's later journeys to the Malay Archipelago, included stops at several island chains, the last being the Galápagos Islands west of Ecuador. On these islands, Darwin observed species of organisms on different islands that were clearly similar, yet had distinct differences.

For example, the ground finches inhabiting the Galápagos Islands comprised several species with a unique beak shape (Figure 1). The species on the islands had a graded series of beak sizes and shapes with very small differences between the most similar. He observed that these finches closely resembled another finch species on the mainland of South America. Darwin imagined that the island species might be species modified from one of the original mainland species. Upon further study, he realized that the varied beaks of each finch helped the birds acquire a specific type of food. For example, seed-eating finches had stronger, thicker beaks for breaking seeds, and insect-eating finches had spear-like beaks for stabbing their prey.

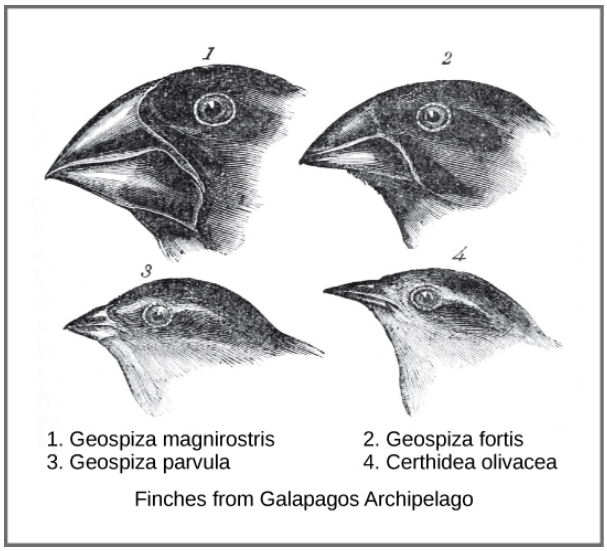


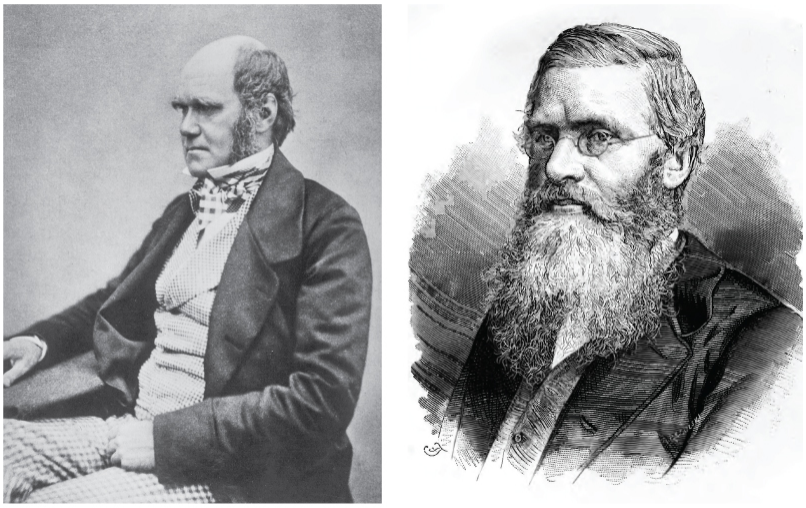
Figure 1. Darwin observed that beak shape varies among finch species. He postulated that the beak of an ancestral species had adapted over time to equip the finches to acquire different food sources.

Wallace and Darwin both observed similar patterns in other organisms and they independently developed the same explanation for how and why such changes could take place. Darwin called this mechanism natural selection. **Natural selection**, also known as "survival of the fittest," is the more prolific reproduction of individuals with favorable traits that survive environmental change because of those traits; this leads to evolutionary change.

For example, a population of giant tortoises found in the Galapagos Archipelago was observed by Darwin to have longer necks than those that lived on other islands with dry lowlands. These tortoises were "selected" because they could reach more leaves and access more food than those with short necks. In times of drought when fewer leaves would be available, those that could reach more leaves had a better chance to eat and survive than those that couldn't reach the food source. Consequently, long-necked tortoises would be more likely to be reproductively successful and pass the long-necked trait to their offspring. Over time, only long-necked tortoises would be present in the population.

Natural selection, Darwin argued, was an inevitable outcome of three principles that operated in nature. First, most characteristics of organisms are inherited, or passed from parent to offspring. Although no one, including Darwin and Wallace, knew how this happened at the time, it was a common understanding. Second, more offspring are produced than are able to survive, so resources for survival and reproduction are limited. The capacity for reproduction in all organisms outstrips the availability of resources to support their numbers. Thus, there is competition for those resources in each generation. Both Darwin and Wallace's understanding of this principle came from reading an essay by the economist Thomas Malthus who discussed this principle in relation to human populations. Third, offspring vary among each other in regard to their characteristics and those variations are inherited. Darwin and Wallace reasoned that offspring with inherited characteristics which allow them to best compete for limited resources will survive and have more offspring than those individuals with variations that are less able to compete. Because characteristics are inherited, these traits will be better represented in the next generation. This will lead to change in populations over generations in a process that Darwin called descent with modification. Ultimately, natural selection leads to greater adaptation of the population to its local environment; it is the only mechanism known for adaptive evolution.

Papers by Darwin and Wallace (Figure 2) presenting the idea of natural selection were read together in 1858 before the Linnean Society in London. The following year Darwin's book, *On the Origin of Species,* was published. His book outlined in considerable detail his arguments for evolution by natural selection.



(a) (b)

Figure 2. Both (a) Charles Darwin and (b) Alfred Wallace wrote scientific papers on natural selection that were presented together before the Linnean Society in 1858.

Demonstrations of evolution by natural selection are time consuming and difficult to obtain. One of the best examples has been demonstrated in the very birds that helped to inspire Darwin's theory: the Galápagos finches. Peter and Rosemary Grant and their colleagues have studied Galápagos finch populations every year since 1976 and have provided important demonstrations of natural selection. The Grants found changes from one generation to the next in the distribution of beak shapes with the medium ground finch on the Galápagos island of Daphne Major. The birds have inherited variation in the bill shape with some birds having wide deep bills and others having thinner bills. During a period in which rainfall was higher than normal because of an El Niño, the large hard seeds that large-billed birds ate were reduced in number; however, there was an abundance of the small soft seeds which the small-billed birds ate. Therefore, survival and reproduction were much better in the following years for the small-billed birds. In the years following this El Niño, the Grants measured beak sizes in the population and found that the average bill size was smaller. Since bill size is an inherited trait, parents with smaller bills had more offspring and the size of bills had evolved to be smaller. As conditions improved in 1987 and larger seeds became more available, the trend toward smaller average bill size ceased.

Career Connection

Field Biologist

Many people hike, explore caves, scuba dive, or climb mountains for recreation. People often participate in these activities hoping to see wildlife. Experiencing the outdoors can be incredibly enjoyable and invigorating. What if your job was to be outside in the wilderness? Field biologists by definition work outdoors in the "field." The term field in this case refers to any location outdoors, even under water. A field biologist typically focuses research on a certain species, group of organisms, or a single habitat (Figure 3).

Figure 3. A field biologist tranquilizes a polar bear for study. (credit: Karen Rhode)

One objective of many field biologists includes discovering new species that have never been recorded. Not only do such findings expand our understanding of the natural world, but they also lead to important innovations in fields such as medicine and agriculture. Plant and microbial species, in particular, can reveal new medicinal and nutritive knowledge. Other organisms can play key roles in ecosystems or be considered rare and in need of protection. When discovered, these important species can be used as evidence for environmental regulations and laws.

Processes and Patterns of Evolution

Natural selection can only take place if there is **variation**, or differences, among individuals in a population. Importantly, these differences must have some genetic basis; otherwise, the selection will not lead to change in the next generation. This is critical because variation among individuals can be caused by non-genetic reasons such as an individual being taller because of better nutrition rather than different genes.

Genetic diversity in a population comes from two main mechanisms: mutation and sexual reproduction. Mutation, a change in DNA, is the ultimate source of new alleles, or new genetic variation in any population. The genetic changes caused by mutation can have one of three outcomes on the phenotype. A mutation affects the phenotype of the organism in a way that gives it reduced fitness—lower likelihood of survival or fewer offspring. A mutation may produce a phenotype with a beneficial effect on fitness. And, many mutations will also have no effect on the fitness of the phenotype; these are called neutral mutations. Mutations may also have a whole range of effect sizes on the fitness of the organism that expresses them in their phenotype, from a small effect to a great effect. Sexual reproduction also leads to genetic diversity: when two parents reproduce, unique combinations of alleles assemble to produce the unique genotypes and thus phenotypes in each of the offspring.

A heritable trait that helps the survival and reproduction of an organism in its present environment is called an **adaptation**. Scientists describe groups of organisms becoming adapted to their environment when a change in the range of genetic variation occurs over time that increases or maintains the "fit" of the population to its environment. The webbed feet of platypuses are an adaptation for swimming. The snow leopards' thick fur is an adaptation for living in the cold. The cheetahs' fast speed is an adaptation for catching prey.

Whether or not a trait is favorable depends on the environmental conditions at the time. The same traits are not always selected because environmental conditions can change. For example, consider a species of plant that grew in a moist climate and did not need to conserve water. Large leaves were selected because they allowed the plant to obtain more energy from the sun. Large leaves require more water to maintain than small leaves, and the moist environment provided favorable conditions to support large leaves. After thousands of years, the climate changed, and the area no longer had excess water. The direction of natural selection shifted so that plants with small leaves were selected because those populations were able to conserve water to survive the new environmental conditions.

The evolution of species has resulted in enormous variation in form and function. Sometimes, evolution gives rise to groups of organisms that become tremendously different from each other. When two species evolve in diverse directions from a common point, it is called **divergent evolution**. Such divergent evolution can be seen in the forms of the reproductive organs of flowering plants which share the same basic anatomies; however, they can look very different as a result of selection in different physical environments and adaptation to different kinds of pollinators (Figure 4).

(a) (b)

Figure 4. Flowering plants evolved from a common ancestor. Notice that the (a) dense blazing star (Liatrus spicata) and the (b) purple coneflower (Echinacea purpurea) vary in appearance, yet both share a similar basic morphology. (credit a: modification of work by Drew Avery; credit b: modification of work by Cory Zanker)

In other cases, similar phenotypes evolve independently in distantly related species. For example, flight has evolved in both bats and insects, and they both have structures we refer to as wings, which are adaptations to flight. However, the wings of bats and insects have evolved from very different original structures. This phenomenon is called **convergent evolution**, where similar traits evolve independently in species that do not share a recent common ancestry. The two species came to the same function, flying, but did so separately from each other.

These physical changes occur over enormous spans of time and help explain how evolution occurs. Natural selection acts on individual organisms, which in turn can shape an entire species. Although natural selection may work in a single generation on an individual, it can take thousands or even millions of years for the genotype of an entire species to evolve. It is over these large time spans that life on earth has changed and continues to change.

Evidence of Evolution

The evidence for evolution is compelling and extensive. Looking at every level of organization in living systems, biologists see the signature of past and present evolution. Darwin dedicated a large portion of his book, *On the Origin of Species*, to identifying patterns in nature that were consistent with evolution, and since Darwin, our understanding has become clearer and broader.

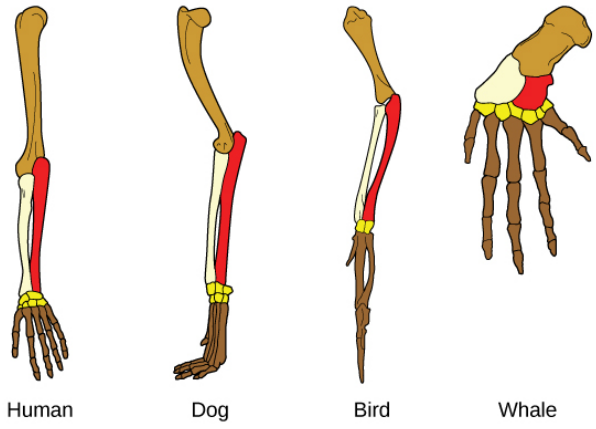
Fossils

Fossils provide solid evidence that organisms from the past are not the same as those found today, and fossils show a progression of evolution. Scientists determine the age of fossils and categorize them from all over the world to determine when the organisms lived relative to each other. The resulting fossil record tells the story of the past and shows the evolution of form over millions of years (Figure). For example, scientists have recovered highly detailed records showing the evolution of humans and horses (Figure). The whale flipper shares a similar morphology to appendages of birds and mammals (Figure 5) indicating that these species share a common ancestor.

(a) (b)

Figure 5. In this (a) display, fossil hominids are arranged from oldest (bottom) to newest (top). As hominids evolved, the shape of the skull changed. An artist's rendition of (b) extinct species of the genus Equus reveals that these ancient species resembled the modern horse (Equus ferus) but varied in size.

Anatomy and Embryology



Another type of evidence for evolution is the presence of structures in organisms that share the same basic form. For example, the bones in the appendages of a human, dog, bird, and whale all share the same overall construction (Figure) resulting from their origin in the appendages of a common ancestor. Over time, evolution led to changes in the shapes and sizes of these bones in different species, but they have maintained the same overall layout. Scientists call these synonymous parts **homologous structures**.

Figure 6. The similar construction of these appendages indicates that these organisms share a common ancestor.

Some structures exist in organisms that have no apparent function at all, and appear to be residual parts from a past common ancestor. These unused structures without function are called **vestigial structures**. Other examples of vestigial structures are wings on flightless birds, leaves on some cacti, and hind leg bones in whales.

Link to Learning

Visit this interactive site to guess which bones structures are homologous and which are analogous, and see examples of evolutionary adaptations to illustrate these concepts.

Another evidence of evolution is the convergence of form in organisms that share similar environments. For example, species of unrelated animals, such as the arctic fox and ptarmigan, living in the arctic region have been selected for seasonal white phenotypes during winter to blend with the snow and ice (Figure 7). These similarities occur not because of common ancestry, but because of similar selection pressures—the benefits of not being seen by predators.

Embryology, the study of the development of the anatomy of an organism to its adult form, also provides evidence of relatedness between now widely divergent groups of organisms. Mutational tweaking in the embryo can have such magnified consequences in the adult that embryo formation tends to be conserved. As a result, structures that are absent in some groups often appear in their embryonic forms and disappear by the time the adult or juvenile form is reached. For example, all vertebrate embryos, including humans, exhibit gill slits and tails at some point in their early development. These disappear in the adults of terrestrial groups but are maintained in adult forms of aquatic groups such as fish and some amphibians. Great ape embryos, including humans, have a tail structure during their development that is lost by the time of birth.

(a) (b)

Figure 7. The white winter coat of the (a) arctic fox and the (b) ptarmigan's plumage are adaptations to their environments. (credit a: modification of work by Keith Morehouse)

Biogeography

The geographic distribution of organisms on the planet follows patterns that are best explained by evolution in conjunction with the movement of tectonic plates over geological time. Broad groups that evolved before the breakup of the supercontinent Pangaea (about 200 million years ago) are distributed worldwide. Groups that evolved since the breakup appear uniquely in regions of the planet, such as the unique flora and fauna of northern continents that formed from the supercontinent Laurasia and of the southern continents that formed from the supercontinent Gondwana. The presence of members of the plant family Proteaceae in Australia, southern Africa, and South America is best by their presence prior to the southern supercontinent Gondwana breaking up.

The great diversification of marsupials in Australia and the absence of other mammals reflect Australia's long isolation. Australia has an abundance of endemic species—species found nowhere else—which is typical of islands whose isolation by expanses of water prevents species to migrate. Over time, these species diverge evolutionarily into new species that look very different from their ancestors that may exist on the mainland. The marsupials of Australia, the finches on the Galápagos, and many species on the Hawaiian Islands are all unique to their one point of origin, yet they display distant relationships to ancestral species on mainlands.

Molecular Biology

Like anatomical structures, the structures of the molecules of life reflect descent with modification. Evidence of a common ancestor for all of life is reflected in the universality of DNA as the genetic material and in the near universality of the genetic code and the machinery of DNA replication and expression. Fundamental divisions in life between the three domains are reflected in major structural differences in otherwise conservative structures such as the components of ribosomes and the structures of membranes. In general, the relatedness of groups of organisms is reflected in the similarity of their DNA sequences—exactly the pattern that would be expected from descent and diversification from a common ancestor.

DNA sequences have also shed light on some of the mechanisms of evolution. For example, it is clear that the evolution of new functions for proteins commonly occurs after gene duplication events that allow the free modification of one copy by mutation, selection, or drift (changes in a population's gene pool resulting from chance), while the second copy continues to produce a functional protein.

Misconceptions of Evolution

Although the theory of evolution generated some controversy when it was first proposed, it was almost universally accepted by biologists, particularly younger biologists, within 20 years after publication of *On the Origin of Species*. Nevertheless, the theory of evolution is a difficult concept and misconceptions about how it works abound.

> Link to Learning
> This site addresses some of the main misconceptions associated with the theory of evolution.

Evolution Is Just a Theory

Critics of the theory of evolution dismiss its importance by purposefully confounding the everyday usage of the word "theory" with the way scientists use the word. In science, a "theory" is understood to be a body of thoroughly tested and verified explanations for a set of observations of the natural world. Scientists have a theory of the atom, a theory of gravity, and the theory of relativity, each of which describes understood facts about the world. In the same way, the theory of evolution describes facts about the living world. As such, a theory in science has survived significant efforts to discredit it by scientists. In contrast, a "theory" in common vernacular is a word meaning a guess or suggested explanation; this meaning is more akin to the scientific concept of "hypothesis." When critics of evolution say evolution is "just a theory," they are implying that there is little evidence supporting it and that it is still in the process of being rigorously tested. This is a mischaracterization.

Individuals Evolve

Evolution is the change in genetic composition of a population over time, specifically over generations, resulting from differential reproduction of individuals with certain alleles. Individuals do change over their lifetime, obviously, but this is called development and involves changes programmed by the set of genes the individual acquired at birth in coordination with the individual's environment. When thinking about the evolution of a characteristic, it is probably best to think about the change of the average value of the characteristic in the population over time. For example, when natural selection leads to bill-size change in medium-ground finches in the Galápagos, this does not mean that individual bills on the finches are changing. If one measures the average bill size among all

individuals in the population at one time and then measures the average bill size in the population several years later, this average value will be different as a result of evolution. Although some individuals may survive from the first time to the second, they will still have the same bill size; however, there will be many new individuals that contribute to the shift in average bill size.

Evolution Explains the Origin of Life

It is a common misunderstanding that evolution includes an explanation of life's origins. Conversely, some of the theory's critics believe that it cannot explain the origin of life. The theory does not try to explain the origin of life. The theory of evolution explains how populations change over time and how life diversifies the origin of species. It does not shed light on the beginnings of life including the origins of the first cells, which is how life is defined. The mechanisms of the origin of life on Earth are a particularly difficult problem because it occurred a very long time ago, and presumably it just occurred once. Importantly, biologists believe that the presence of life on Earth precludes the possibility that the events that led to life on Earth can be repeated because the intermediate stages would immediately become food for existing living things.

However, once a mechanism of inheritance was in place in the form of a molecule like DNA either within a cell or pre-cell, these entities would be subject to the principle of natural selection. More effective reproducers would increase in frequency at the expense of inefficient reproducers. So while evolution does not explain the origin of life, it may have something to say about some of the processes operating once pre-living entities acquired certain properties.

Organisms Evolve on Purpose

Statements such as "organisms evolve in response to a change in an environment" are quite common, but such statements can lead to two types of misunderstandings. First, the statement must not be understood to mean that individual organisms evolve. The statement is shorthand for "a population evolves in response to a changing environment." However, a second misunderstanding may arise by interpreting the statement to mean that the evolution is somehow intentional. A changed environment results in some individuals in the population, those with particular phenotypes, benefiting and therefore producing proportionately more offspring than other phenotypes. This results in change in the population if the characteristics are genetically determined.

It is also important to understand that the variation that natural selection works on is already in a population and does not arise in response to an environmental change. For example, applying antibiotics to a population of bacteria will, over time, select a population of bacteria that are resistant to antibiotics. The resistance, which is caused by a gene, did not arise by mutation because of the application of the antibiotic. The gene for resistance was already present in the gene pool of the bacteria, likely at a low frequency. The antibiotic, which kills the bacterial cells without the resistance gene, strongly selects individuals that are resistant, since these would be the only ones that survived and divided. Experiments have demonstrated that mutations for antibiotic resistance do not arise as a result of antibiotic.

In a larger sense, evolution is not goal directed. Species do not become "better" over time; they simply track their changing environment with adaptations that maximize their reproduction in a particular environment at a particular time. Evolution has no goal of making faster, bigger, more complex, or even smarter species, despite the commonness of this kind of language in popular discourse. What characteristics evolve in a species are a function of the variation present and the environment, both of which are constantly changing in a non-directional way. What trait is fit in one environment at one time may well be fatal at some point in the future. This holds equally well for a species of insect as it does the human species.

Section Summary

Evolution is the process of adaptation through mutation which allows more desirable characteristics to be passed to the next generation. Over time, organisms evolve more characteristics that are beneficial to their survival. For living organisms to adapt and change to environmental pressures, genetic variation must be present. With genetic variation, individuals have differences in form and function that allow some to survive certain conditions better than others. These organisms pass their favorable traits to their offspring. Eventually, environments change, and

what was once a desirable, advantageous trait may become an undesirable trait and organisms may further evolve. Evolution may be convergent with similar traits evolving in multiple species or divergent with diverse traits evolving in multiple species that came from a common ancestor. Evidence of evolution can be observed by means of DNA code and the fossil record, and also by the existence of homologous and vestigial structures.

Visit this page in your course online to check your understanding.

Self Check Questions

1. If a person scatters a handful of garden pea plant seeds in one area, how would natural selection work in this situation?

2. Why do scientists consider vestigial structures evidence for evolution?

3. How does the scientific meaning of “theory” differ from the common vernacular meaning?

4. Explain why the statement that a monkey is more evolved than a mouse is incorrect.

Answers

1. The plants that can best use the resources of the area, including competing with other individuals for those resources will produce more seeds themselves and those traits that allowed them to better use the resources will increase in the population of the next generation.

2. Vestigial structures are considered evidence for evolution because most structures do not exist in an organism without serving some function either presently or in the past. A vestigial structure indicates a past form or function that has since changed, but the structure remains present because it had a function in the ancestor.
3. In science, a theory is a thoroughly tested and verified set of explanations for a body of observations of nature. It is the strongest form of knowledge in science. In contrast, a theory in common vernacular can mean a guess or speculation about something, meaning that the knowledge implied by the theory is very weak.
4. The statement implies that there is a goal to evolution and that the monkey represents greater progress to that goal than the mouse. Both species are likely to be well adapted to their particular environments, which is the outcome of natural selection.

Glossary

adaptation: heritable trait or behavior in an organism that aids in its survival and reproduction in its present environment

convergent evolution: process by which groups of organisms independently evolve to similar forms

divergent evolution: process by which groups of organisms evolve in diverse directions from a common point

homologous structures: parallel structures in diverse organisms that have a common ancestor

natural selection: reproduction of individuals with favorable genetic traits that survive environmental change because of those traits, leading to evolutionary change

variation: genetic differences among individuals in a population

vestigial structure: physical structure present in an organism but that has no apparent function and appears to be from a functional structure in a distant ancestor

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by:** Open Stax. **Located at:** http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License:** *CC BY: Attribution*

POPULATION GENETICS

Learning Objectives

By the end of this section, you will be able to:

- Describe the different types of variation in a population
- Explain why only heritable variation can be acted upon by natural selection
- Describe genetic drift and the bottleneck effect
- Explain how each evolutionary force can influence the allele frequencies of a population

Individuals of a population often display different phenotypes, or express different alleles of a particular gene, referred to as polymorphisms. Populations with two or more variations of particular characteristics are called polymorphic. The distribution of phenotypes among individuals, known as the population variation, is influenced by a number of factors, including the population's genetic structure and the environment (Figure 1). Understanding the sources of a phenotypic variation in a population is important for determining how a population will evolve in response to different evolutionary pressures.

Figure 1. The distribution of phenotypes in this litter of kittens illustrates population variation. (credit: Pieter Lanser)

Genetic Variance

Natural selection and some of the other evolutionary forces can only act on heritable traits, namely an organism's genetic code. Because alleles are passed from parent to offspring, those that confer beneficial traits or behaviors may be selected for, while deleterious alleles may be selected against. Acquired traits, for the most part, are not heritable. For example, if an athlete works out in the gym every day, building up muscle strength, the athlete's offspring will not necessarily grow up to be a body builder. If there is a genetic basis for the ability to run fast, on the other hand, this may be passed to a child.

Link to Learning

Before Darwinian evolution became the prevailing theory of the field, French naturalist Jean-Baptiste Lamarck theorized that acquired traits could, in fact, be inherited; while this hypothesis has largely been unsupported, scientists have recently begun to realize that Lamarck was not completely wrong. Visit this site to learn more.

Heritability is the fraction of phenotype variation that can be attributed to genetic differences, or genetic variance, among individuals in a population. The greater the hereditability of a population's phenotypic variation, the more susceptible it is to the evolutionary forces that act on heritable variation.

The diversity of alleles and genotypes within a population is called **genetic variance**. When scientists are involved in the breeding of a species, such as with animals in zoos and nature preserves, they try to increase a population's genetic variance to preserve as much of the phenotypic diversity as they can. This also helps reduce the risks associated with **inbreeding**, the mating of closely related individuals, which can have the undesirable effect of bringing together deleterious recessive mutations that can cause abnormalities and susceptibility to disease. For example, a disease that is caused by a rare, recessive allele might exist in a population, but it will

only manifest itself when an individual carries two copies of the allele. Because the allele is rare in a normal, healthy population with unrestricted habitat, the chance that two carriers will mate is low, and even then, only 25 percent of their offspring will inherit the disease allele from both parents. While it is likely to happen at some point, it will not happen frequently enough for natural selection to be able to swiftly eliminate the allele from the population, and as a result, the allele will be maintained at low levels in the gene pool. However, if a family of carriers begins to interbreed with each other, this will dramatically increase the likelihood of two carriers mating and eventually producing diseased offspring, a phenomenon known as **inbreeding depression**.

Changes in allele frequencies that are identified in a population can shed light on how it is evolving. In addition to natural selection, there are other evolutionary forces that could be in play: genetic drift, gene flow, mutation, nonrandom mating, and environmental variances.

Genetic Drift

The theory of natural selection stems from the observation that some individuals in a population are more likely to survive longer and have more offspring than others; thus, they will pass on more of their genes to the next generation. A big, powerful male gorilla, for example, is much more likely than a smaller, weaker one to become the population's silverback, the pack's leader who mates far more than the other males of the group. The pack leader will father more offspring, who share half of his genes, and are likely to also grow bigger and stronger like their father. Over time, the genes for bigger size will increase in frequency in the population, and the population will, as a result, grow larger on average. That is, this would occur if this particular **selection pressure**, or driving selective force, were the only one acting on the population. In other examples, better camouflage or a stronger resistance to drought might pose a selection pressure.

Another way a population's allele and genotype frequencies can change is **genetic drift** (Figure 2), which is simply the effect of chance. By chance, some individuals will have more offspring than others—not due to an advantage conferred by some genetically-encoded trait, but just because one male happened to be in the right place at the right time (when the receptive female walked by) or because the other one happened to be in the wrong place at the wrong time (when a fox was hunting).

Art Connection

Figure 2 is too large to fit on one or two pages. Visit your course online to view Figure 2.

Figure 2. The genetic drift of a rabbit population.

Genetic drift in a population can lead to the elimination of an allele from a population by chance. In Figure 2, rabbits with the brown coat color allele (*B*) are dominant over rabbits with the white coat color allele (*b*). In the first generation, the two alleles occur with equal frequency in the population, resulting in p and q values of .5. Only half of the individuals reproduce, resulting in a second generation with p and q values of .7 and .3, respectively. Only two individuals in the second generation reproduce, and by chance these individuals are homozygous dominant for brown coat color. As a result, in the third generation the recessive *b* allele is lost.

Do you think genetic drift would happen more quickly on an island or on the mainland?

Small populations are more susceptible to the forces of genetic drift. Large populations, on the other hand, are buffered against the effects of chance. If one individual of a population of 10 individuals happens to die at a young age before it leaves any offspring to the next generation, all of its genes—1/10 of the population's gene pool—will be suddenly lost. In a population of 100, that's only 1 percent of the overall gene pool; therefore, it is much less impactful on the population's genetic structure.

Link to Learning

Go to this site to watch an animation of random sampling and genetic drift in action.

Genetic drift can also be magnified by natural events, such as a natural disaster that kills—at random—a large portion of the population. Known as the **bottleneck effect**, it results in a large portion of the genome suddenly being wiped out (Figure 3). In one fell swoop, the genetic structure of the survivors becomes the genetic structure of the entire population, which may be very different from the pre-disaster population.

Another scenario in which populations might experience a strong influence of genetic drift is if some portion of the population leaves to start a new population in a new location or if a population gets divided by a physical barrier of some kind. In this situation, those individuals are unlikely to be representative of the entire population, which results in the founder effect. The founder effect occurs when the genetic structure changes to match that of the new population's founding fathers and mothers. The founder effect is believed to have been a key factor in the genetic history of the Afrikaner population of Dutch settlers in South Africa, as evidenced by mutations that are common in Afrikaners but rare in most other populations. This is likely due to the fact that a higher-than-normal proportion of the founding colonists carried these mutations. As a result, the population expresses unusually high incidences of Huntington's disease (HD) and Fanconi anemia (FA), a genetic disorder known to cause blood marrow and congenital abnormalities—even cancer. ((Note: A. J. Tipping et al., "Molecular and Genealogical Evidence for a Founder Effect in Fanconi Anemia Families of the Afrikaner Population of South Africa," *PNAS* 98, no. 10 (2001): 5734-5739, doi: 10.1073/pnas.091402398.))

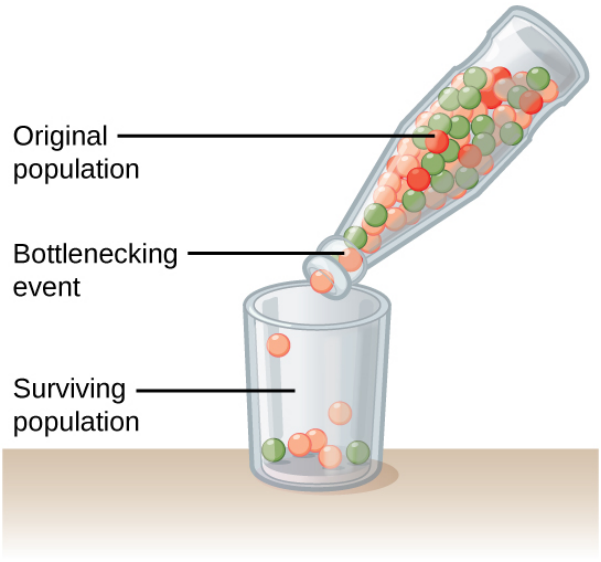


Figure 3. A chance event or catastrophe can reduce the genetic variability within a population.

Link to Learning

Watch this short video to learn more about the founder and bottleneck effects.
Watch this video online: https://youtu.be/hEYV9WEvwaI

Scientific Method Connection

Testing the Bottleneck Effect

Question: How do natural disasters affect the genetic structure of a population?

Background: When much of a population is suddenly wiped out by an earthquake or hurricane, the individuals that survive the event are usually a random sampling of the original group. As a result, the genetic makeup of the population can change dramatically. This phenomenon is known as the bottleneck effect.

Hypothesis: Repeated natural disasters will yield different population genetic structures; therefore, each time this experiment is run, the results will vary.

Test the hypothesis: Count out the original population using different colored beads. For example, red, blue, and yellow beads might represent red, blue, and yellow individuals. After recording the number of each individual in the original population, place them all in a bottle with a narrow neck that will only allow a few beads out at a time. Then, pour 1/3 of the bottle's contents into a bowl. This represents the surviving individuals after a natural disaster kills a majority of the population. Count the number of the different colored beads in the bowl, and record it. Then, place all of the beads back in the bottle and repeat the experiment four more times.

Analyze the data: Compare the five populations that resulted from the experiment. Do the populations all contain the same number of different colored beads, or do they vary? Remember, these populations all came from the same exact parent population.

Form a conclusion: Most likely, the five resulting populations will differ quite dramatically. This is because natural disasters are not selective—they kill and spare individuals at random. Now think about how this might affect a real population. What happens when a hurricane hits the Mississippi Gulf Coast? How do the seabirds that live on the beach fare?

Gene Flow

Another important evolutionary force is gene flow: the flow of alleles in and out of a population due to the migration of individuals or gametes (Figure 4). While some populations are fairly stable, others experience more flux. Many plants, for example, send their pollen far and wide, by wind or by bird, to pollinate other populations of the same species some distance away. Even a population that may initially appear to be stable, such as a pride of lions, can experience its fair share of immigration and emigration as developing males leave their mothers to seek out a new pride with genetically unrelated females. This variable flow of individuals in and out of the group not only changes the gene structure of the population, but it can also introduce new genetic variation to populations in different geological locations and habitats.

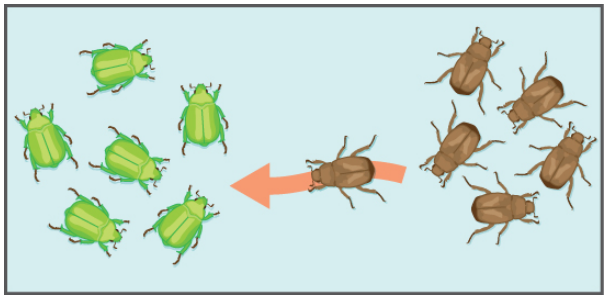


Figure 4. Gene flow can occur when an individual travels from one geographic location to another.

Mutation

Mutations are changes to an organism's DNA and are an important driver of diversity in populations. Species evolve because of the accumulation of mutations that occur over time. The appearance of new mutations is the most common way to introduce novel genotypic and phenotypic variance. Some mutations are unfavorable or harmful and are quickly eliminated from the population by natural selection. Others are beneficial and will spread through the population. Whether or not a mutation is beneficial or harmful is determined by whether it helps an organism survive to sexual maturity and reproduce. Some mutations do not do anything and can linger, unaffected by natural selection, in the genome. Some can have a dramatic effect on a gene and the resulting phenotype.

Nonrandom Mating

If individuals nonrandomly mate with their peers, the result can be a changing population. There are many reasons **nonrandom mating** occurs. One reason is simple mate choice; for example, female peahens may prefer peacocks with bigger, brighter tails. Traits that lead to more matings for an individual become selected for by natural selection. One common form of mate choice, called **assortative mating**, is an individual's preference to mate with partners who are phenotypically similar to themselves.

Another cause of nonrandom mating is physical location. This is especially true in large populations spread over large geographic distances where not all individuals will have equal access to one another. Some might be miles apart through woods or over rough terrain, while others might live immediately nearby.

Environmental Variance

Genes are not the only players involved in determining population variation. Phenotypes are also influenced by other factors, such as the environment (Figure 5). A beachgoer is likely to have darker skin than a city dweller, for example, due to regular exposure to the sun, an environmental factor. Some major characteristics, such as gender, are determined by the environment for some species. For example, some turtles and other reptiles have temperature-dependent sex determination (TSD). TSD means that individuals develop into males if their eggs are incubated within a certain temperature range, or females at a different temperature range.

Figure 5. The sex of the American alligator (Alligator mississippiensis) is determined by the temperature at which the eggs are incubated. Eggs incubated at 30°C produce females, and eggs incubated at 33°C produce males. (credit: Steve Hillebrand, USFWS)

Geographic separation between populations can lead to differences in the phenotypic variation between those populations. Such geographical variation is seen between most populations and can be significant. One type of geographic variation, called a cline, can be seen as populations of a given species vary gradually across an ecological gradient. Species of warm-blooded animals, for example, tend to have larger bodies in the cooler climates closer to the earth's poles, allowing them to better conserve heat. This is considered a latitudinal cline. Alternatively, flowering plants tend to bloom at different times depending on where they are along the slope of a mountain, known as an altitudinal cline.

If there is gene flow between the populations, the individuals will likely show gradual differences in phenotype along the cline. Restricted gene flow, on the other hand, can lead to abrupt differences, even speciation.

Section Summary

Both genetic and environmental factors can cause phenotypic variation in a population. Different alleles can confer different phenotypes, and different environments can also cause individuals to look or act differently. Only those differences encoded in an individual's genes, however, can be passed to its offspring and, thus, be a target of natural selection. Natural selection works by selecting for alleles that confer beneficial traits or behaviors, while selecting against those for deleterious qualities. Genetic drift stems from the chance occurrence that some individuals in the germ line have more offspring than others. When individuals leave or join the population, allele frequencies can change as a result of gene flow. Mutations to an individual's DNA may introduce new variation into a population. Allele frequencies can also be altered when individuals do not randomly mate with others in the group.

Visit this page in your course online to check your understanding.

Additional Self Check Questions
1. Do you think genetic drift would happen more quickly on an island or on the mainland?
2. Describe a situation in which a population would undergo the bottleneck effect and explain what impact that would have on the population's gene pool.
3. Describe natural selection and give an example of natural selection at work in a population.

4. Explain what a cline is and provide examples.

Answers

1. Genetic drift is likely to occur more rapidly on an island where smaller populations are expected to occur.

2. A hurricane kills a large percentage of a population of sand-dwelling crustaceans—only a few individuals survive. The alleles carried by those surviving individuals would represent the entire population's gene pool. If those surviving individuals are not representative of the original population, the post-hurricane gene pool will differ from the original gene pool.

3. The theory of natural selection stems from the observation that some individuals in a population survive longer and have more offspring than others: thus, more of their genes are passed to the next generation. For example, a big, powerful male gorilla is much more likely than a smaller, weaker one to become the population's silverback: the pack's leader who mates far more than the other males of the group. Therefore, the pack leader will father more offspring who share half of his genes and are likely to grow bigger and stronger like their father. Over time, the genes for bigger size will increase in frequency in the population, and the average body size, as a result, grow larger on average.

4. A cline is a type of geographic variation that is seen in populations of a given species that vary gradually across an ecological gradient. For example, warm-blooded animals tend to have larger bodies in the cooler climates closer to the earth's poles, allowing them to better conserve heat. This is considered a latitudinal cline. Flowering plants tend to bloom at different times depending on where they are along the slope of a mountain. This is known as an altitudinal cline.

Glossary

assortative mating: when individuals tend to mate with those who are phenotypically similar to themselves

bottleneck effect: magnification of genetic drift as a result of natural events or catastrophes

cline: gradual geographic variation across an ecological gradient

gene flow: flow of alleles in and out of a population due to the migration of individuals or gametes

genetic drift: effect of chance on a population's gene pool

genetic variance: diversity of alleles and genotypes in a population

geographical variation: differences in the phenotypic variation between populations that are separated geographically

heritability: fraction of population variation that can be attributed to its genetic variance

inbreeding: mating of closely related individuals

inbreeding depression: increase in abnormalities and disease in inbreeding populations

nonrandom mating: changes in a population's gene pool due to mate choice or other forces that cause individuals to mate with certain phenotypes more than others

population variation: distribution of phenotypes in a population

selective pressure: environmental factor that causes one phenotype to be better than another

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: Open Stax. **Located at**: http://cnx.org/contents/185cbf87-c72e-48f5-b51e-f14f21b5eabd@9.17:1/Biology. **License**: *CC BY: Attribution*

All rights reserved content

- Founder and Bottleneck Effect (Evolution). **Authored by**: Greg Korchnak. **Located at**: https://youtu.be/hEYV9WEvwaI. **License**: *All Rights Reserved*. **License Terms**: Standard YouTube License

STUDY GUIDE: INTRODUCTION TO EVOLUTION

Study Questions

Objective: Define evolution.

Use this page to check your understanding of the content.

Vocabulary

1. Evolution
2. Population
3. Allele frequency
4. Evolutionary tree (aka cladogram)
5. Extant
6. Extinct
7. Common ancestor
8. Gene pool

Study Guide Questions

1. Compare and contrast “species” and “populations”.
2. Compare and contrast microevolution and macroevolution.
3. What is the difference between microevolution and macroevolution? Please don’t just memorize the definitions…be able to APPLY your definitions to different scenarios! For good practice, think of examples of each!
4. Relate the concept of a gene pool to evolution.
5. Define speciation.
6. Be able to calculate the allele frequency of a population given the following information:

25% of the population had the bb genotype

25% had the Bb genotype

50% had the BB genotype,

Licensing & Attributions

CC licensed content, Shared previously

- Study Guide: Introduction to Evolution. **Authored by**: Wendy Riggs. **Provided by**: College of the Redwoods. **Project**: Kaleidoscope. **License**: *CC BY: Attribution*

STUDY GUIDE: MECHANISMS OF EVOLUTION

Study Questions

Objective: Compare and contrast the many mechanisms by which evolutionary change occurs.

Use this page to check your understanding of the content.

Vocabulary

1. Evolution.
2. Natural selection
3. Mutation
4. Gene flow
5. Genetic drift
6. Sexual selection

Study Guide Questions

1. Be able to identify, compare, contrast, and discuss the various mechanisms of microevolution, including:
    1. Mutation
    2. Gene flow
    3. Genetic drift
    4. Sexual selection
    5. Natural selection

1. What are the observations that led to Darwin's conclusions regarding natural selection?
2. Compare and contrast sexual selection and natural selection.
3. What is the difference between microevolution and macroevolution? Please don't just memorize the definitions…be able to APPLY your definitions to different scenarios! For good practice, think of examples of each!
4. Clearly explain HOW speciation occurs…
5. Clearly describe each of the following forms of reproductive isolation. Be able to compare and contrast each form.
    1. Geographic
    2. Ecological
    3. Temporal
    4. Behavioral
    5. Mechanical
    6. Gametic isolation
    7. Hybrid inviability

1. Given any scenario, be able to determine the TYPE of reproductive isolation that is occurring. Make up scenarios and practice!
2. Explain the connection between reproductive isolation, speciation, and microevolution.

Licensing & Attributions

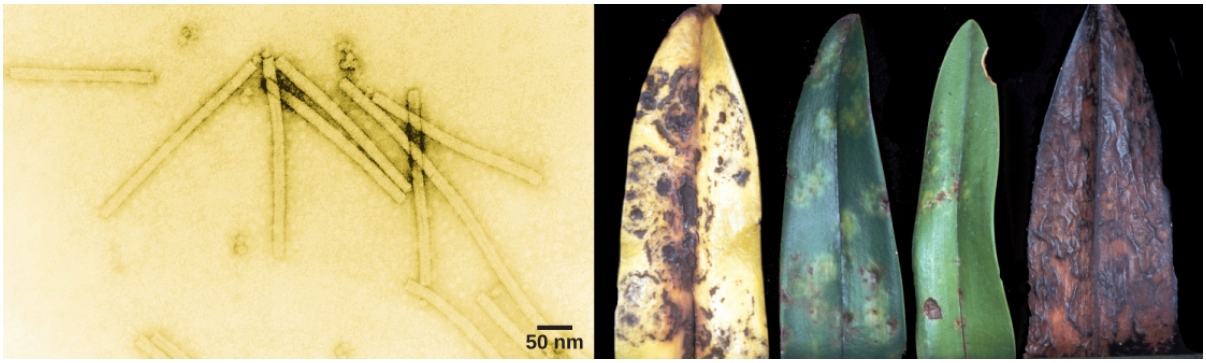
CC licensed content, Shared previously

- Study Guide: Mechanisms of Evolution. **Authored by**: Wendy Riggs. **Provided by**: College of the Redwoods. **Project**: Kaleidoscope. **License**: *CC BY: Attribution*

VIRUSES

INTRODUCTION: VIRUSES

Chapter 21. Introduction



The tobacco mosaic virus (left), seen here by transmission electron microscopy, was the first virus to be discovered. The virus causes disease in tobacco and other plants, such as the orchid (right). (credit a: USDA ARS; credit b: modification of work by USDA Forest Service, Department of Plant Pathology Archive North Carolina State University; scale-bar data from Matt Russell)

No one knows exactly when viruses emerged or from where they came, since viruses do not leave historical footprints such as fossils. Modern viruses are thought to be a mosaic of bits and pieces of nucleic acids picked up from various sources along their respective evolutionary paths. Viruses are acellular, parasitic entities that are not classified within any kingdom. Unlike most living organisms, viruses are not cells and cannot divide. Instead, they infect a host cell and use the host's replication processes to produce identical progeny virus particles. Viruses infect organisms as diverse as bacteria, plants, and animals. They exist in a netherworld between a living organism and a nonliving entity. Living things grow, metabolize, and reproduce. Viruses replicate, but to do so, they are entirely dependent on their host cells. They do not metabolize or grow, but are assembled in their mature form.

VIRAL EVOLUTION, MORPHOLOGY, AND CLASSIFICATION

Viral Evolution, Morphology, and Classification

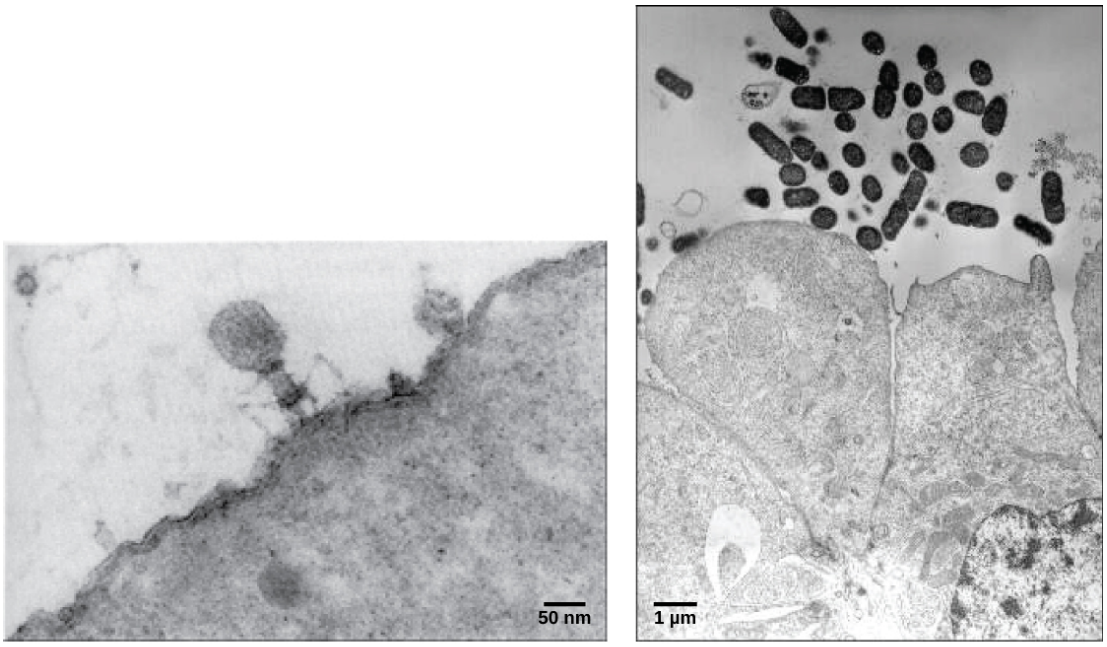
Summary

Viruses are diverse entities. They vary in their structure, their replication methods, and in their target hosts. Nearly all forms of life—from bacteria and archaea to eukaryotes such as plants, animals, and fungi—have viruses that infect them. While most biological diversity can be understood through evolutionary history, such as how species have adapted to conditions and environments, much about virus origins and evolution remains unknown.

Discovery and Detection

Viruses were first discovered after the development of a porcelain filter, called the Chamberland-Pasteur filter, which could remove all bacteria visible in the microscope from any liquid sample. In 1886, Adolph Meyer demonstrated that a disease of tobacco plants, tobacco mosaic disease, could be transferred from a diseased plant to a healthy one via liquid plant extracts. In 1892, Dmitri Ivanowski showed that this disease could be transmitted in this way even after the Chamberland-Pasteur filter had removed all viable bacteria from the extract. Still, it was many years before it was proven that these “filterable” infectious agents were not simply very small bacteria but were a new type of very small, disease-causing particle.

Virions, single virus particles, are very small, about 20–250 nanometers in diameter. These individual virus particles are the infectious form of a virus outside the host cell. Unlike bacteria (which are about 100-times larger), we cannot see viruses with a light microscope, with the exception of some large virions of the poxvirus family. It was not until the development of the electron microscope in the late 1930s that scientists got their first good view of the structure of the tobacco mosaic virus (TMV) ([link]) and other viruses (Figure). The surface structure of virions can be observed by both scanning and transmission electron microscopy, whereas the internal structures of the virus can only be observed in images from a transmission electron microscope. The use of these technologies has allowed for the discovery of many viruses of all types of living organisms. They were initially grouped by shared morphology. Later, groups of viruses were classified by the type of nucleic acid they contained, DNA or RNA, and whether their nucleic acid was single- or double-stranded. More recently, molecular analysis of viral replicative cycles has further refined their classification.



(a) (b)

In these transmission electron micrographs, (a) a virus is dwarfed by the bacterial cell it infects, while (b) these *E. coli* cells are dwarfed by cultured colon cells. (credit a: modification of work by U.S. Dept. of Energy, Office of Science, LBL, PBD; credit b: modification of work by J.P. Nataro and S. Sears, unpub. data, CDC; scale-bar data from Matt Russell)

Evolution of Viruses

Although biologists have accumulated a significant amount of knowledge about how present-day viruses evolve, much less is known about how viruses originated in the first place. When exploring the evolutionary history of most organisms, scientists can look at fossil records and similar historic evidence. However, viruses do not fossilize, so researchers must conjecture by investigating how today's viruses evolve and by using biochemical and genetic information to create speculative virus histories.

While most findings agree that viruses don't have a single common ancestor, scholars have yet to find a single hypothesis about virus origins that is fully accepted in the field. One such hypothesis, called devolution or the regressive hypothesis, proposes to explain the origin of viruses by suggesting that viruses evolved from free-living cells. However, many components of how this process might have occurred are a mystery. A second hypothesis (called escapist or the progressive hypothesis) accounts for viruses having either an RNA or a DNA genome and suggests that viruses originated from RNA and DNA molecules that escaped from a host cell. A third hypothesis posits a system of self-replication similar to that of other self-replicating molecules, likely evolving alongside the cells they rely on as hosts; studies of some plant pathogens support this hypothesis.

As technology advances, scientists may develop and refine further hypotheses to explain the origin of viruses. The emerging field called virus molecular systematics attempts to do just that through comparisons of sequenced genetic material. These researchers hope to one day better understand the origin of viruses, a discovery that could lead to advances in the treatments for the ailments they produce.

Viral Morphology

Viruses are acellular, meaning they are biological entities that do not have a cellular structure. They therefore lack most of the components of cells, such as organelles, ribosomes, and the plasma membrane. A virion consists of a nucleic acid core, an outer protein coating or capsid, and sometimes an outer envelope made of protein and phospholipid membranes derived from the host cell. Viruses may also contain additional proteins, such as enzymes. The most obvious difference between members of viral families is their morphology, which is quite

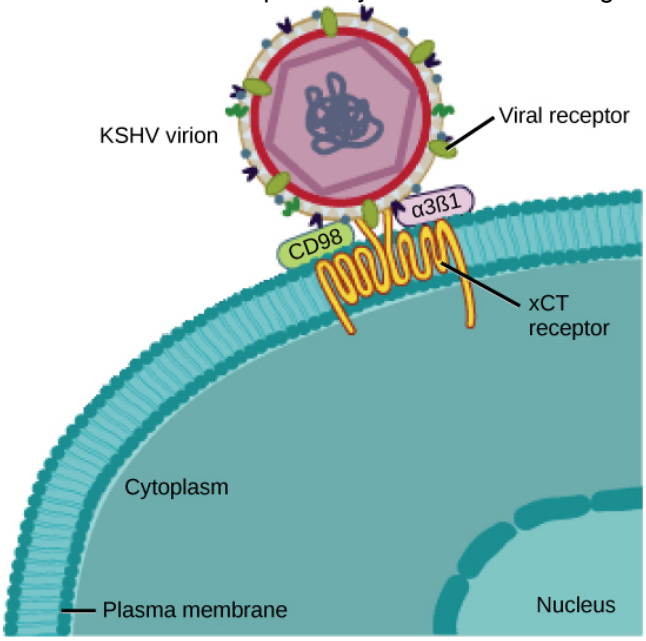
diverse. An interesting feature of viral complexity is that the complexity of the host does not correlate with the complexity of the virion. Some of the most complex virion structures are observed in bacteriophages, viruses that infect the simplest living organisms, bacteria.

Morphology

Viruses come in many shapes and sizes, but these are consistent and distinct for each viral family. All virions have a nucleic acid genome covered by a protective layer of proteins, called a capsid. The capsid is made up of protein subunits calledcapsomeres. Some viral capsids are simple polyhedral “spheres,” whereas others are quite complex in structure.

In general, the shapes of viruses are classified into four groups: filamentous, isometric (or icosahedral), enveloped, and head and tail. Filamentous viruses are long and cylindrical. Many plant viruses are filamentous, including TMV. Isometric viruses have shapes that are roughly spherical, such as poliovirus or herpesviruses. Enveloped viruses have membranes surrounding capsids. Animal viruses, such as HIV, are frequently enveloped. Head and tail viruses infect bacteria and have a head that is similar to icosahedral viruses and a tail shape like filamentous viruses.

Many viruses use some sort of glycoprotein to attach to their host cells via molecules on the cell called viral receptors (Figure). For these viruses, attachment is a requirement for later penetration of the cell membrane, so they can complete their replication inside the cell. The receptors that viruses use are molecules that are normally found on cell surfaces and have their own physiological functions. Viruses have simply evolved to make use of these molecules for their own replication. For example, HIV uses the CD4 molecule on T lymphocytes as one of its receptors. CD4 is a type of molecule called a cell adhesion molecule, which functions to keep different types of immune cells in close proximity to each other during the generation of a T lymphocyte immune response.



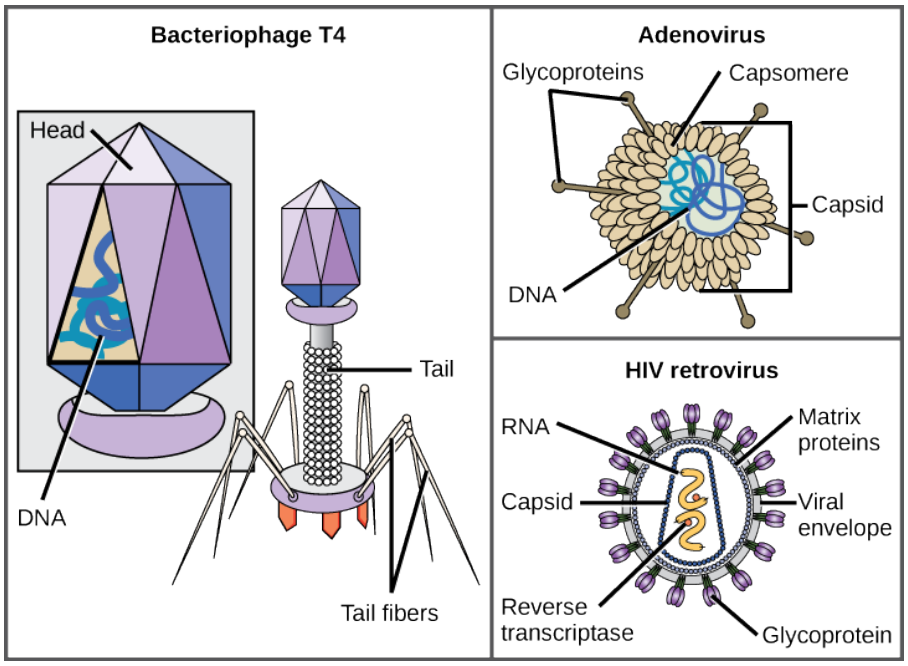
The KSHV virus binds the xCT receptor on the surface of human cells. xCT receptors protect cells against stress. Stressed cells express more xCT receptors than non-stressed cells. The KSHV virion causes cells to become stressed, thereby increasing expression of the receptor to which it binds. (credit: modification of work by NIAID, NIH)

Among the most complex virions known, the T4 bacteriophage, which infects the *Escherichia coli* bacterium, has a tail structure that the virus uses to attach to host cells and a head structure that houses its DNA.

Adenovirus, a non-enveloped animal virus that causes respiratory illnesses in humans, uses glycoprotein spikes protruding from its capsomeres to attach to host cells. Non-enveloped viruses also include those that cause polio (poliovirus), plantar warts (papillomavirus), and hepatitis A (hepatitis A virus).

Enveloped virions like HIV, the causative agent in AIDS, consist of nucleic acid (RNA in the case of HIV) and capsid proteins surrounded by a phospholipid bilayer envelope and its associated proteins. Glycoproteins embedded in the viral envelope are used to attach to host cells. Other envelope proteins are the matrix proteins that stabilize the envelope and often play a role in the assembly of progeny virions. Chicken pox, influenza, and mumps are examples of diseases caused by viruses with envelopes. Because of the fragility of the envelope, non-enveloped viruses are more resistant to changes in temperature, pH, and some disinfectants than enveloped viruses.

Overall, the shape of the virion and the presence or absence of an envelope tell us little about what disease the virus may cause or what species it might infect, but they are still useful means to begin viral classification (Figure).
ART CONNECTION



Viruses can be either complex in shape or relatively simple. This figure shows three relatively complex virions: the bacteriophage T4, with its DNA-containing head group and tail fibers that attach to host cells; adenovirus, which uses spikes from its capsid to bind to host cells; and HIV, which uses glycoproteins embedded in its envelope to bind to host cells. Notice that HIV has proteins called matrix proteins, internal to the envelope, which help stabilize virion shape. (credit “bacteriophage, adenovirus”: modification of work by NCBI, NIH; credit “HIV retrovirus”: modification of work by NIAID, NIH)

Which of the following statements about virus structure is true?

1. All viruses are encased in a viral membrane.
2. The capsomere is made up of small protein subunits called capsids.
3. DNA is the genetic material in all viruses.
4. Glycoproteins help the virus attach to the host cell.

Types of Nucleic Acid

Unlike nearly all living organisms that use DNA as their genetic material, viruses may use either DNA or RNA as theirs. The virus core contains the genome or total genetic content of the virus. Viral genomes tend to be small, containing only those genes that encode proteins that the virus cannot get from the host cell. This genetic material may be single- or double-stranded. It may also be linear or circular. While most viruses contain a single nucleic acid, others have genomes that have several, which are called segments.

In DNA viruses, the viral DNA directs the host cell’s replication proteins to synthesize new copies of the viral genome and to transcribe and translate that genome into viral proteins. DNA viruses cause human diseases, such as chickenpox, hepatitis B, and some venereal diseases, like herpes and genital warts.

RNA viruses contain only RNA as their genetic material. To replicate their genomes in the host cell, the RNA viruses encode enzymes that can replicate RNA into DNA, which cannot be done by the host cell. These RNA polymerase enzymes are more likely to make copying errors than DNA polymerases, and therefore often make mistakes during transcription. For this reason, mutations in RNA viruses occur more frequently than in DNA viruses. This causes them to change and adapt more rapidly to their host. Human diseases caused by RNA viruses include hepatitis C, measles, and rabies.

Virus Classification

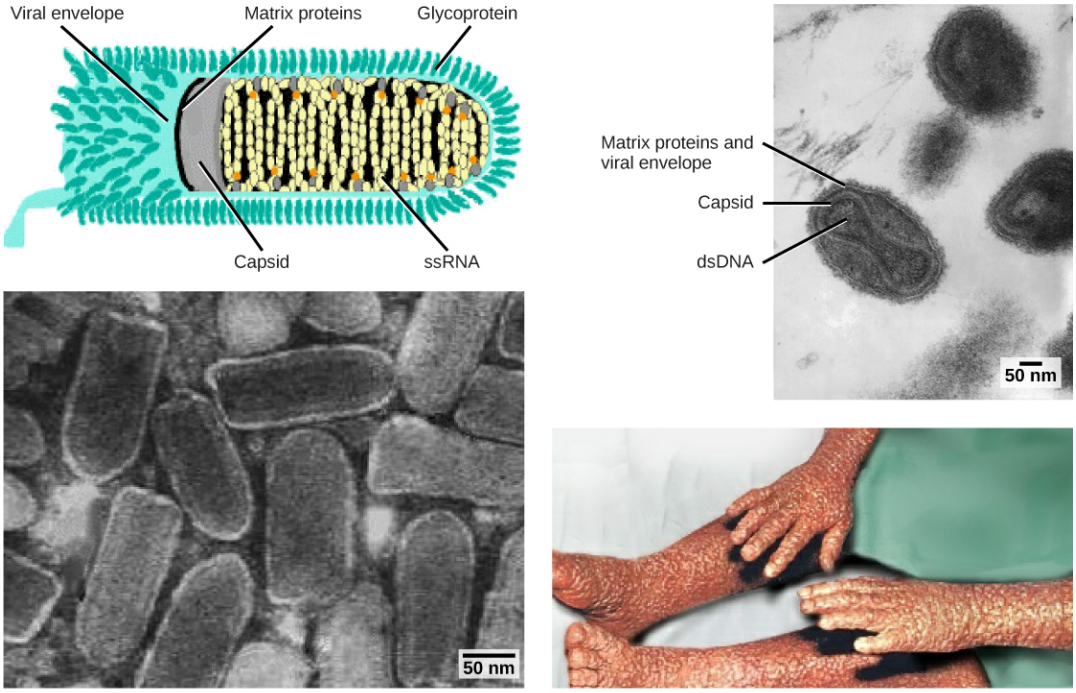
To understand the features shared among different groups of viruses, a classification scheme is necessary. As most viruses are not thought to have evolved from a common ancestor, however, the methods that scientists use to classify living things are not very useful. Biologists have used several classification systems in the past, based on the morphology and genetics of the different viruses. However, these earlier classification methods grouped viruses differently, based on which features of the virus they were using to classify them. The most commonly used classification method today is called the Baltimore classification scheme and is based on how messenger RNA (mRNA) is generated in each particular type of virus.

Past Systems of Classification

Viruses are classified in several ways: by factors such as their core content (Table and Figure), the structure of their capsids, and whether they have an outer envelope. The type of genetic material (DNA or RNA) and its structure (single- or double-stranded, linear or circular, and segmented or non-segmented) are used to classify the virus core structures.

Virus Classification by Genome Structure and Core

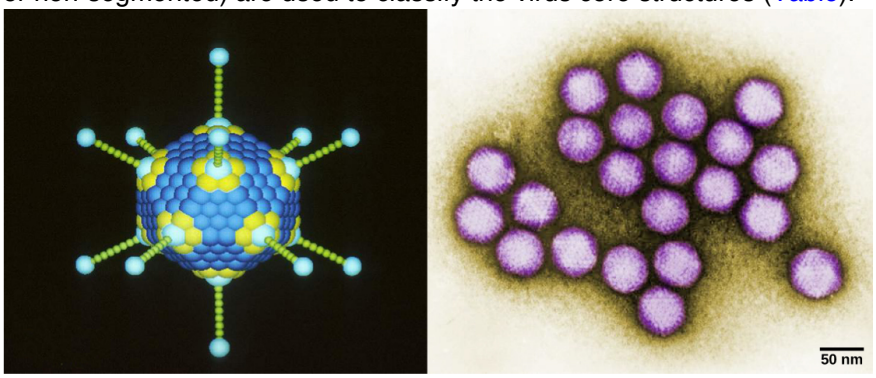
Core Classifications	Examples
• RNA • DNA	• Rabies virus, retroviruses • Herpesviruses, smallpox virus
• Single-stranded • Double-stranded	• Rabies virus, retroviruses • Herpesviruses, smallpox virus
• Linear • Circular	• Rabies virus, retroviruses, herpesviruses, smallpox virus • Papillomaviruses, many bacteriophages
• Non-segmented: genome consists of a single segment of genetic material • Segmented: genome is divided into multiple segments	• Parainfluenza viruses • Influenza viruses



(a) Rabies virus (b) Variola virus

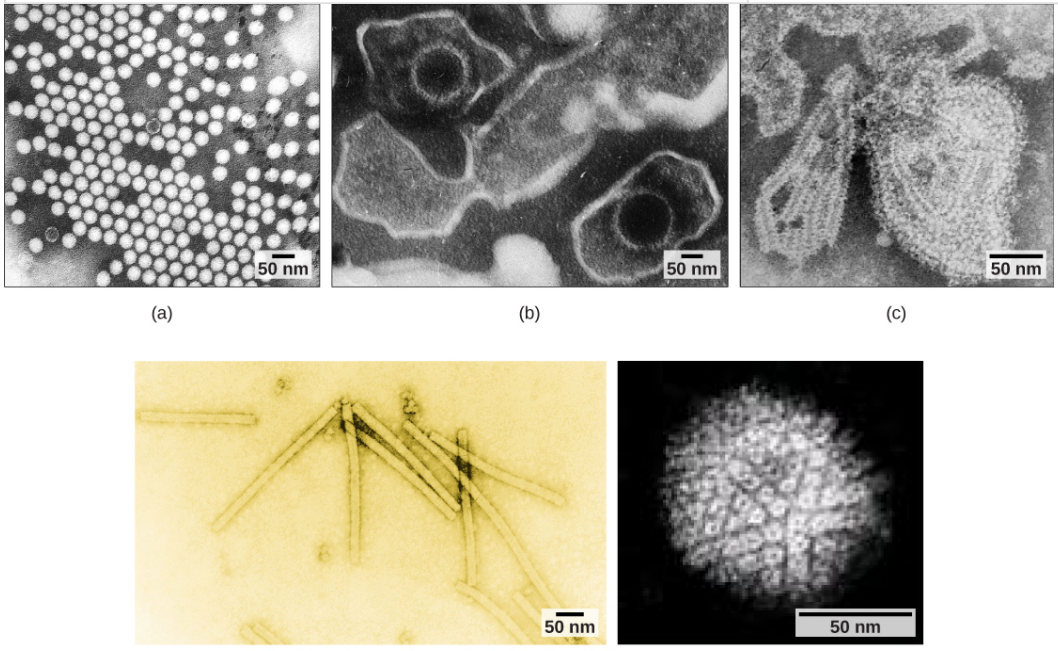
Viruses are classified based on their core genetic material and capsid design. (a) Rabies virus has a single-stranded RNA (ssRNA) core and an enveloped helical capsid, whereas (b) variola virus, the causative agent of smallpox, has a double-stranded DNA (dsDNA) core and a complex capsid. Rabies transmission occurs when saliva from an infected mammal enters a wound. The virus travels through neurons in the peripheral nervous system to the central nervous system where it impairs brain function, and then travels to other tissues. The virus can infect any mammal, and most die within weeks of infection. Smallpox is a human virus transmitted by inhalation of the variola virus, localized in the skin, mouth, and throat, which causes a characteristic rash. Before its eradication in 1979, infection resulted in a 30–35 percent mortality rate. (credit “rabies diagram”: modification of work by CDC; “rabies micrograph”: modification of work by Dr. Fred Murphy, CDC; credit “small pox micrograph”: modification of work by Dr. Fred Murphy, Sylvia Whitfield, CDC; credit “smallpox photo”: modification of work by CDC; scale-bar data from Matt Russell)

Viruses can also be classified by the design of their capsids (Figure and Figure). Capsids are classified as naked icosahedral, enveloped icosahedral, enveloped helical, naked helical, and complex (Figure and Figure). The type of genetic material (DNA or RNA) and its structure (single- or double-stranded, linear or circular, and segmented or non-segmented) are used to classify the virus core structures (Table).



Adenovirus (left) is depicted with a double-stranded DNA genome enclosed in an icosahedral capsid that is 90–100 nm across. The virus, shown clustered in the micrograph (right), is transmitted orally and causes a variety of illnesses in vertebrates, including human eye and respiratory infections. (credit “adenovirus”: modification of work by Dr. Richard Feldmann, National Cancer Institute; credit “micrograph”: modification of work by Dr. G. William Gary, Jr., CDC; scale-bar data from Matt Russell)

Virus Classification by Capsid Structure	
Capsid Classification	Examples
Naked icosahedral	Hepatitis A virus, polioviruses
Enveloped icosahedral	Epstein-Barr virus, herpes simplex virus, rubella virus, yellow fever virus, HIV-1
Enveloped helical	Influenza viruses, mumps virus, measles virus, rabies virus
Naked helical	Tobacco mosaic virus
Complex with many proteins; some have combinations of icosahedral and helical capsid structures	Herpesviruses, smallpox virus, hepatitis B virus, T4 bacteriophage



Transmission electron micrographs of various viruses show their structures. The capsid of the (a) polio virus is naked icosahedral; (b) the Epstein-Barr virus capsid is enveloped icosahedral; (c) the mumps virus capsid is an enveloped helix; (d) the tobacco mosaic virus capsid is naked helical; and (e) the herpesvirus capsid is complex. (credit a: modification of work by Dr. Fred Murphy, Sylvia Whitfield; credit b: modification of work by Liza Gross; credit c: modification of work by Dr. F. A. Murphy, CDC; credit d: modification of work by USDA ARS; credit e: modification of work by Linda Stannard, Department of Medical Microbiology, University of Cape Town, South Africa, NASA; scale-bar data from Matt Russell)

Baltimore Classification

The most commonly used system of virus classification was developed by Nobel Prize-winning biologist David Baltimore in the early 1970s. In addition to the differences in morphology and genetics mentioned above, the Baltimore classification scheme groups viruses according to how the mRNA is produced during the replicative cycle of the virus.

Group I viruses contain double-stranded DNA (dsDNA) as their genome. Their mRNA is produced by transcription in much the same way as with cellular DNA. Group II viruses have single-stranded DNA (ssDNA) as their genome. They convert their single-stranded genomes into a dsDNA intermediate before transcription to mRNA can occur. Group III viruses use dsRNA as their genome. The strands separate, and one of them is used as a

template for the generation of mRNA using the RNA-dependent RNA polymerase encoded by the virus. Group IV viruses have ssRNA as their genome with a positive polarity. Positive polaritymeans that the genomic RNA can serve directly as mRNA. Intermediates of dsRNA, called replicative intermediates, are made in the process of copying the genomic RNA. Multiple, full-length RNA strands of negative polarity (complimentary to the positive-stranded genomic RNA) are formed from these intermediates, which may then serve as templates for the production of RNA with positive polarity, including both full-length genomic RNA and shorter viral mRNAs. Group V viruses contain ssRNA genomes with a negative polarity, meaning that their sequence is complementary to the mRNA. As with Group IV viruses, dsRNA intermediates are used to make copies of the genome and produce mRNA. In this case, the negative-stranded genome can be converted directly to mRNA. Additionally, full-length positive RNA strands are made to serve as templates for the production of the negative-stranded genome. Group VI viruses have diploid (two copies) ssRNA genomes that must be converted, using the enzyme reverse transcriptase, to dsDNA; the dsDNA is then transported to the nucleus of the host cell and inserted into the host genome. Then, mRNA can be produced by transcription of the viral DNA that was integrated into the host genome. Group VII viruses have partial dsDNA genomes and make ssRNA intermediates that act as mRNA, but are also converted back into dsDNA genomes by reverse transcriptase, necessary for genome replication. The characteristics of each group in the Baltimore classification are summarized in Table with examples of each group.

Baltimore Classification			
Group	Characteristics	Mode of mRNA Production	Example
I	Double-stranded DNA	mRNA is transcribed directly from the DNA template	Herpes simplex (herpesvirus)
II	Single-stranded DNA	DNA is converted to double-stranded form before RNA is transcribed	Canine parvovirus (parvovirus)
III	Double-stranded RNA	mRNA is transcribed from the RNA genome	Childhood gastroenteritis (rotavirus)
IV	Single stranded RNA (+)	Genome functions as mRNA	Common cold (pircornavirus)
V	Single stranded RNA (-)	mRNA is transcribed from the RNA genome	Rabies (rhabdovirus)
VI	Single stranded RNA viruses with reverse transcriptase	Reverse transcriptase makes DNA from the RNA genome; DNA is then incorporated in the host genome; mRNA is transcribed from the incorporated DNA	Human immunodeficiency virus (HIV)
VII	Double stranded DNA viruses with reverse transcriptase	The viral genome is double-stranded DNA, but viral DNA is replicated through an RNA intermediate; the RNA may serve directly as mRNA or as a template to make mRNA	Hepatitis B virus (hepadnavirus)

Section Summary

Viruses are tiny, acellular entities that can usually only be seen with an electron microscope. Their genomes contain either DNA or RNA—never both—and they replicate using the replication proteins of a host cell. Viruses are diverse, infecting archaea, bacteria, fungi, plants, and animals. Viruses consist of a nucleic acid core surrounded by a protein capsid with or without an outer lipid envelope. The capsid shape, presence of an envelope, and core composition dictate some elements of the classification of viruses. The most commonly used classification method, the Baltimore classification, categorizes viruses based on how they produce their mRNA.

Art Connections

Figure Which of the following statements about virus structure is true?

1. All viruses are encased in a viral membrane.
2. The capsomere is made up of small protein subunits called capsids.
3. DNA is the genetic material in all viruses.
4. Glycoproteins help the virus attach to the host cell.

Review Questions

Which statement is true?

1. A virion contains DNA and RNA.
2. Viruses are acellular.
3. Viruses replicate outside of the cell.
4. Most viruses are easily visualized with a light microscope.

The viral _____ plays a role in attaching a virion to the host cell.

1. core
2. capsid
3. envelope
4. both b and c

Viruses_____.

1. all have a round shape
2. cannot have a long shape
3. do not maintain any shape
4. vary in shape

Free Response

The first electron micrograph of a virus (tobacco mosaic virus) was produced in 1939. Before that time, how did scientists know that viruses existed if they could not see them? (Hint: Early scientists called viruses “filterable agents.”)

Glossary

acellular lacking cells

capsid protein coating of the viral core

capsomere protein subunit that makes up the capsid

envelope lipid bilayer that envelopes some viruses

group I virus virus with a dsDNA genome

group II virus virus with a ssDNA genome

group III virus virus with a dsRNA genome

group IV virus virus with a ssRNA genome with positive polarity

group V virus virus with a ssRNA genome with negative polarity

group VI virus virus with a ssRNA genomes converted into dsDNA by reverse transcriptase

group VII virus virus with a single-stranded mRNA converted into dsDNA for genome replication

matrix protein envelope protein that stabilizes the envelope and often plays a role in the assembly of progeny virions

negative polarity ssRNA viruses with genomes complimentary to their mRNA

positive polarity ssRNA virus with a genome that contains the same base sequences and codons found in their mRNA

replicative intermediate dsRNA intermediate made in the process of copying genomic RNA

reverse transcriptase enzyme found in Baltimore groups VI and VII that converts single-stranded RNA into double-stranded DNA

viral receptor glycoprotein used to attach a virus to host cells via molecules on the cell

virion individual virus particle outside a host cell

virus core contains the virus genome

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: Open Stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License**: *CC BY: Attribution*

VIRUS INFECTIONS AND HOSTS

Virus Infections and Hosts

Summary

Viruses can be seen as obligate, intracellular parasites. A virus must attach to a living cell, be taken inside, manufacture its proteins and copy its genome, and find a way to escape the cell so that the virus can infect other cells. Viruses can infect only certain species of hosts and only certain cells within that host. Cells that a virus may use to replicate are called permissive. For most viruses, the molecular basis for this specificity is that a particular surface molecule known as the viral receptor must be found on the host cell surface for the virus to attach. Also, metabolic and host cell immune response differences seen in different cell types based on differential gene expression are a likely factor in which cells a virus may target for replication. The permissive cell must make the substances that the virus needs or the virus will not be able to replicate there.

Steps of Virus Infections

A virus must use cell processes to replicate. The viral replication cycle can produce dramatic biochemical and structural changes in the host cell, which may cause cell damage. These changes, called cytopathic (causing cell damage) effects, can change cell functions or even destroy the cell. Some infected cells, such as those infected by the common cold virus known as rhinovirus, die through lysis (bursting) or apoptosis (programmed cell death or “cell suicide”), releasing all progeny virions at once. The symptoms of viral diseases result from the immune response to the virus, which attempts to control and eliminate the virus from the body, and from cell damage caused by the virus. Many animal viruses, such as HIV (human immunodeficiency virus), leave the infected cells of the immune system by a process known as budding, where virions leave the cell individually. During the budding process, the cell does not undergo lysis and is not immediately killed. However, the damage to the cells

that the virus infects may make it impossible for the cells to function normally, even though the cells remain alive for a period of time. Most productive viral infections follow similar steps in the virus replication cycle: attachment, penetration, uncoating, replication, assembly, and release (Figure).

Attachment

A virus attaches to a specific receptor site on the host cell membrane through attachment proteins in the capsid or via glycoproteins embedded in the viral envelope. The specificity of this interaction determines the host—and the cells within the host—that can be infected by a particular virus. This can be illustrated by thinking of several keys and several locks, where each key will fit only one specific lock.

Link to Learning

This video explains how influenza attacks the body.

Entry

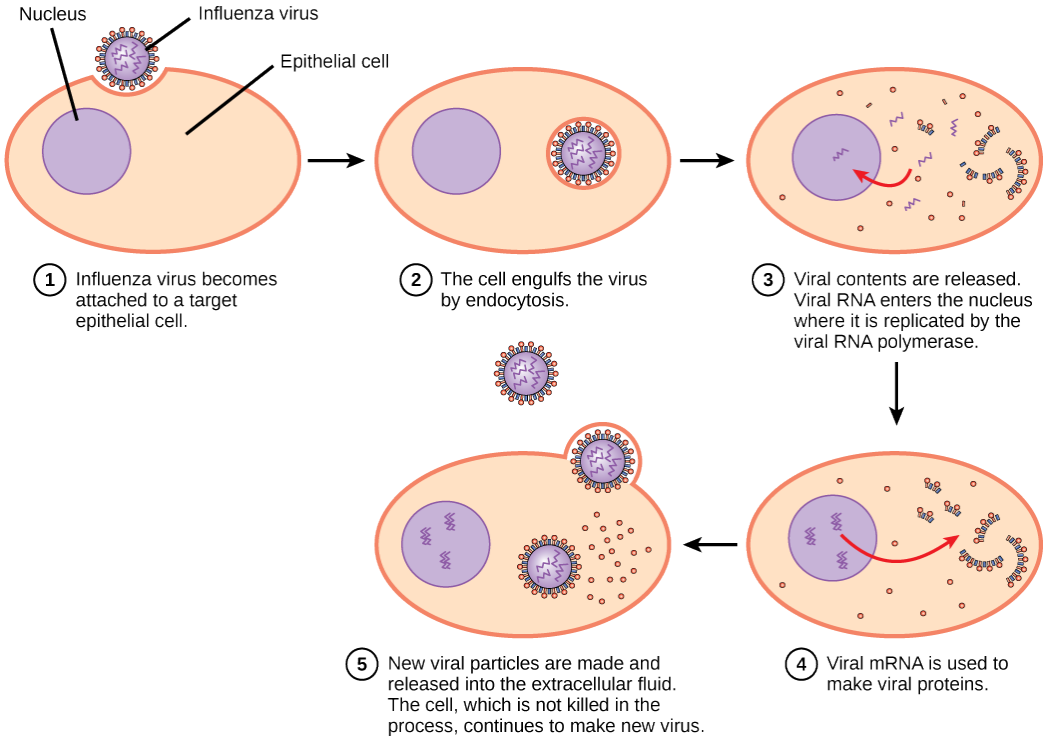
The nucleic acid of bacteriophages enters the host cell naked, leaving the capsid outside the cell. Plant and animal viruses can enter through endocytosis, in which the cell membrane surrounds and engulfs the entire virus. Some enveloped viruses enter the cell when the viral envelope fuses directly with the cell membrane. Once inside the cell, the viral capsid is degraded, and the viral nucleic acid is released, which then becomes available for replication and transcription.

Replication and Assembly

The replication mechanism depends on the viral genome. DNA viruses usually use host cell proteins and enzymes to make additional DNA that is transcribed to messenger RNA (mRNA), which is then used to direct protein synthesis. RNA viruses usually use the RNA core as a template for synthesis of viral genomic RNA and mRNA. The viral mRNA directs the host cell to synthesize viral enzymes and capsid proteins, and assemble new virions. Of course, there are exceptions to this pattern. If a host cell does not provide the enzymes necessary for viral replication, viral genes supply the information to direct synthesis of the missing proteins. Retroviruses, such as HIV, have an RNA genome that must be reverse transcribed into DNA, which then is incorporated into the host cell genome. They are within group VI of the Baltimore classification scheme. To convert RNA into DNA, retroviruses must contain genes that encode the virus-specific enzyme reverse transcriptase that transcribes an RNA template to DNA. Reverse transcription never occurs in uninfected host cells—the needed enzyme reverse transcriptase is only derived from the expression of viral genes within the infected host cells. The fact that HIV produces some of its own enzymes not found in the host has allowed researchers to develop drugs that inhibit these enzymes. These drugs, including the reverse transcriptase inhibitorAZT, inhibit HIV replication by reducing the activity of the enzyme without affecting the host's metabolism. This approach has led to the development of a variety of drugs used to treat HIV and has been effective at reducing the number of infectious virions (copies of viral RNA) in the blood to non-detectable levels in many HIV-infected individuals.

Egress

The last stage of viral replication is the release of the new virions produced in the host organism, where they are able to infect adjacent cells and repeat the replication cycle. As you've learned, some viruses are released when the host cell dies, and other viruses can leave infected cells by budding through the membrane without directly killing the cell.
ART CONNECTION



In influenza virus infection, glycoproteins attach to a host epithelial cell. As a result, the virus is engulfed. RNA and proteins are made and assembled into new virions.

Influenza virus is packaged in a viral envelope that fuses with the plasma membrane. This way, the virus can exit the host cell without killing it. What advantage does the virus gain by keeping the host cell alive?

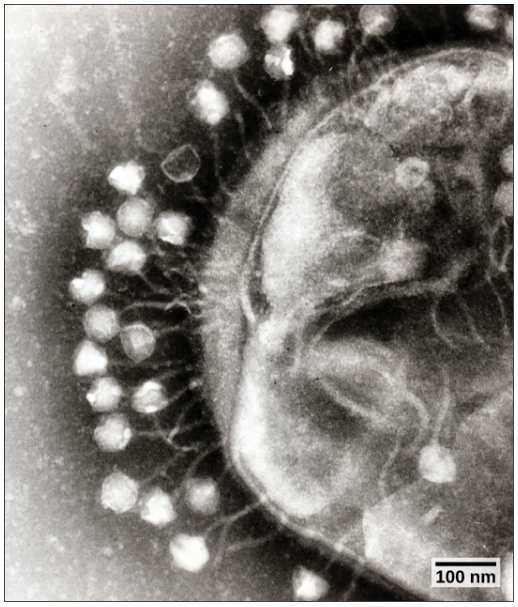
Link to Learning

Watch a video on viruses, identifying structures, modes of transmission, replication, and more.

Different Hosts and Their Viruses

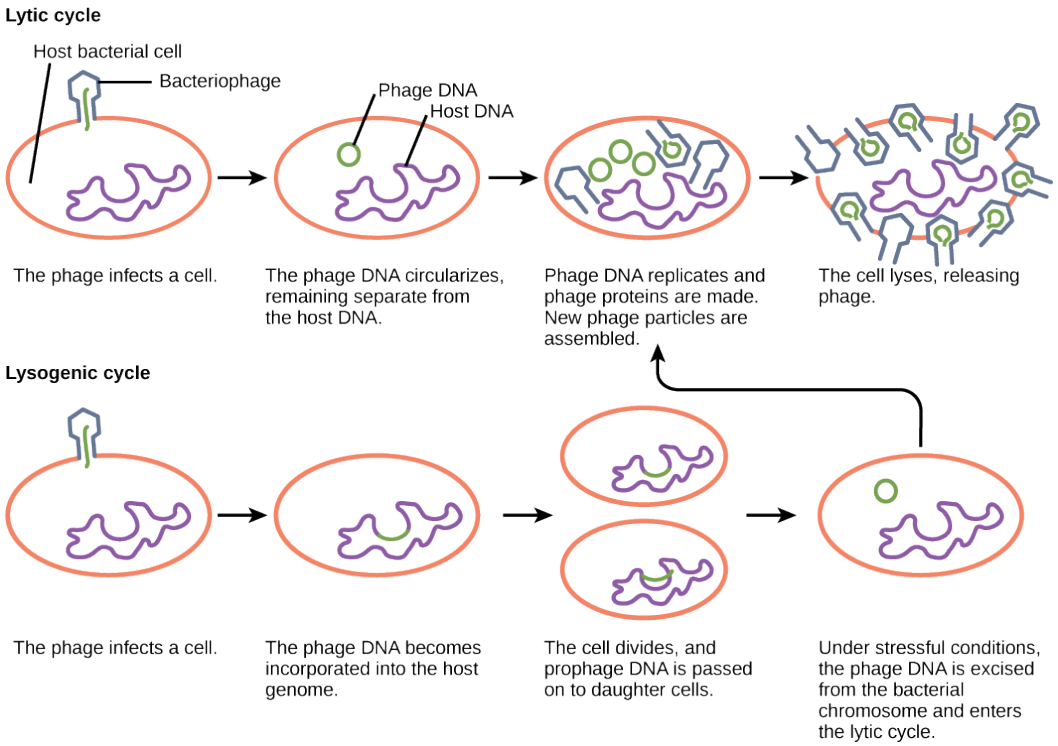
As you've learned, viruses are often very specific as to which hosts and which cells within the host they will infect. This feature of a virus makes it specific to one or a few species of life on Earth. On the other hand, so many different types of viruses exist on Earth that nearly every living organism has its own set of viruses that tries to infect its cells. Even the smallest and simplest of cells, prokaryotic bacteria, may be attacked by specific types of viruses.

Bacteriophages



This transmission electron micrograph shows bacteriophages attached to a bacterial cell. (credit: modification of work by Dr. Graham Beards; scale-bar data from Matt Russell)

Bacteriophages are viruses that infect bacteria (Figure). When infection of a cell by a bacteriophage results in the production of new virions, the infection is said to be productive. If the virions are released by bursting the cell, the virus replicates by means of a lytic cycle (Figure). An example of a lytic bacteriophage is T4, which infects *Escherichia coli* found in the human intestinal tract. Sometimes, however, a virus can remain within the cell without being released. For example, when a temperate bacteriophage infects a bacterial cell, it replicates by means of a lysogenic cycle (Figure), and the viral genome is incorporated into the genome of the host cell. When the phage DNA is incorporated into the host cell genome, it is called a prophage. An example of a lysogenic bacteriophage is the λ (lambda) virus, which also infects the *E. coli* bacterium. Viruses that infect plant or animal cells may also undergo infections where they are not producing virions for long periods. An example is the animal herpesviruses, including herpes simplex viruses, the cause of oral and genital herpes in humans. In a process called latency, these viruses can exist in nervous tissue for long periods of time without producing new virions, only to leave latency periodically and cause lesions in the skin where the virus replicates. Even though there are similarities between lysogeny and latency, the term lysogenic cycle is usually reserved to describe bacteriophages. Latency will be described in more detail below.
ART CONNECTION



A temperate bacteriophage has both lytic and lysogenic cycles. In the lytic cycle, the phage replicates and lyses the host cell. In the lysogenic cycle, phage DNA is incorporated into the host genome, where it is passed on to subsequent generations. Environmental stressors such as starvation or exposure to toxic chemicals may cause the prophage to excise and enter the lytic cycle.

Which of the following statements is false?

1. In the lytic cycle, new phage are produced and released into the environment.
2. In the lysogenic cycle, phage DNA is incorporated into the host genome.
3. An environmental stressor can cause the phage to initiate the lysogenic cycle.
4. Cell lysis only occurs in the lytic cycle.

Animal Viruses

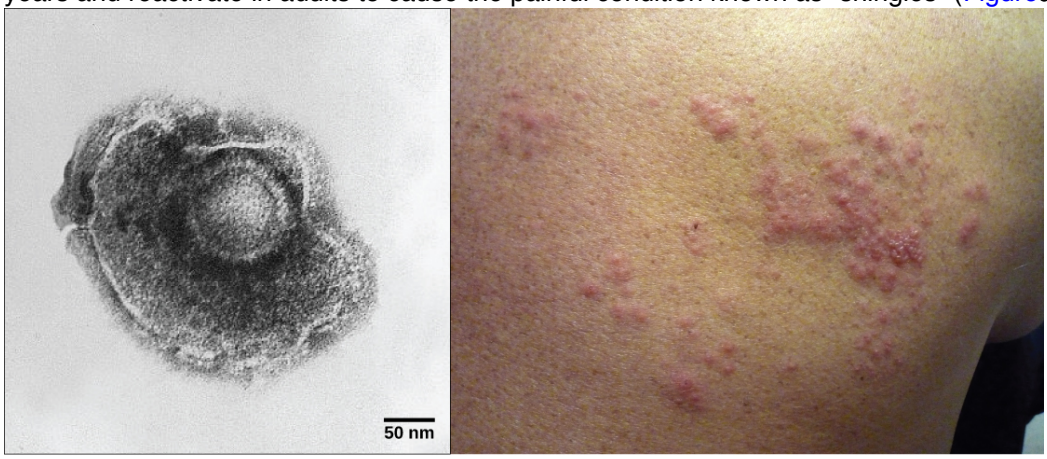
Animal viruses, unlike the viruses of plants and bacteria, do not have to penetrate a cell wall to gain access to the host cell. Non-enveloped or “naked” animal viruses may enter cells in two different ways. As a protein in the viral capsid binds to its receptor on the host cell, the virus may be taken inside the cell via a vesicle during the normal cell process of receptor-mediated endocytosis. An alternative method of cell penetration used by non-enveloped viruses is for capsid proteins to undergo shape changes after binding to the receptor, creating channels in the host cell membrane. The viral genome is then “injected” into the host cell through these channels in a manner analogous to that used by many bacteriophages. Enveloped viruses also have two ways of entering cells after binding to their receptors: receptor-mediated endocytosis, or fusion. Many enveloped viruses enter the cell by receptor-mediated endocytosis in a fashion similar to some non-enveloped viruses. On the other hand, fusion only occurs with enveloped virions. These viruses, which include HIV among others, use special fusion proteins in their envelopes to cause the envelope to fuse with the plasma membrane of the cell, thus releasing the genome and capsid of the virus into the cell cytoplasm.

After making their proteins and copying their genomes, animal viruses complete the assembly of new virions and exit the cell. As we have already discussed using the example of HIV, enveloped animal viruses may bud from the cell membrane as they assemble themselves, taking a piece of the cell’s plasma membrane in the process. On the other hand, non-enveloped viral progeny, such as rhinoviruses, accumulate in infected cells until there is a signal for lysis or apoptosis, and all virions are released together.

As you will learn in the next module, animal viruses are associated with a variety of human diseases. Some of them follow the classic pattern of acute disease, where symptoms get increasingly worse for a short period followed by the elimination of the virus from the body by the immune system and eventual recovery from the infection. Examples of acute viral diseases are the common cold and influenza. Other viruses cause long-term chronic infections, such as the virus causing hepatitis C, whereas others, like herpes simplex virus, only cause intermittent symptoms. Still other viruses, such as human herpesviruses 6 and 7, which in some cases can cause the minor childhood disease roseola, often successfully cause productive infections without causing any symptoms at all in the host, and thus we say these patients have an asymptomatic infection.

In hepatitis C infections, the virus grows and reproduces in liver cells, causing low levels of liver damage. The damage is so low that infected individuals are often unaware that they are infected, and many infections are detected only by routine blood work on patients with risk factors such as intravenous drug use. On the other hand, since many of the symptoms of viral diseases are caused by immune responses, a lack of symptoms is an indication of a weak immune response to the virus. This allows for the virus to escape elimination by the immune system and persist in individuals for years, all the while producing low levels of progeny virions in what is known as a chronic viral disease. Chronic infection of the liver by this virus leads to a much greater chance of developing liver cancer, sometimes as much as 30 years after the initial infection.

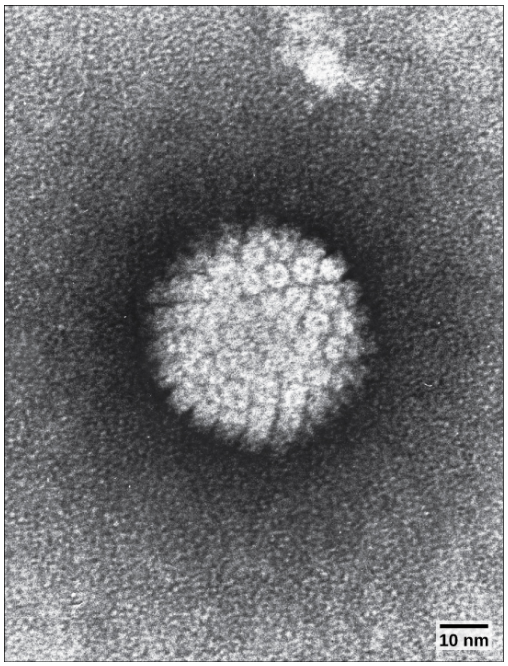
As already discussed, herpes simplex virus can remain in a state of latency in nervous tissue for months, even years. As the virus “hides” in the tissue and makes few if any viral proteins, there is nothing for the immune response to act against, and immunity to the virus slowly declines. Under certain conditions, including various types of physical and psychological stress, the latent herpes simplex virus may be reactivated and undergo a lytic replication cycle in the skin, causing the lesions associated with the disease. Once virions are produced in the skin and viral proteins are synthesized, the immune response is again stimulated and resolves the skin lesions in a few days by destroying viruses in the skin. As a result of this type of replicative cycle, appearances of cold sores and genital herpes outbreaks only occur intermittently, even though the viruses remain in the nervous tissue for life. Latent infections are common with other herpesviruses as well, including the varicella-zoster virus that causes chickenpox. After having a chickenpox infection in childhood, the varicella-zoster virus can remain latent for many years and reactivate in adults to cause the painful condition known as “shingles” (Figure**ab**).



(a) (b)

(a) Varicella-zoster, the virus that causes chickenpox, has an enveloped icosahedral capsid visible in this transmission electron micrograph. Its double-stranded DNA genome becomes incorporated in the host DNA and can reactivate after latency in the form of (b) shingles, often exhibiting a rash. (credit a: modification of work by Dr. Erskine Palmer, B. G. Martin, CDC; credit b: modification of work by “rosmary”/Flickr; scale-bar data from Matt Russell)

Some animal-infecting viruses, including the hepatitis C virus discussed above, are known as oncogenic viruses: They have the ability to cause cancer. These viruses interfere with the normal regulation of the host cell cycle either by either introducing genes that stimulate unregulated cell growth (oncogenes) or by interfering with the expression of genes that inhibit cell growth. Oncogenic viruses can be either DNA or RNA viruses. Cancers known to be associated with viral infections include cervical cancer caused by human papillomavirus (HPV) (Figure), liver cancer caused by hepatitis B virus, T-cell leukemia, and several types of lymphoma.



HPV, or human papillomavirus, has a naked icosahedral capsid visible in this transmission electron micrograph and a double-stranded DNA genome that is incorporated into the host DNA. The virus, which is sexually transmitted, is oncogenic and can lead to cervical cancer. (credit: modification of work by NCI, NIH; scale-bar data from Matt Russell)

Link to Learning

Visit the interactive animations showing the various stages of the replicative cycles of animal viruses and click on the flash animation links.

Plant Viruses

Plant viruses, like other viruses, contain a core of either DNA or RNA. You have already learned about one of these, the tobacco mosaic virus. As plant viruses have a cell wall to protect their cells, these viruses do not use receptor-mediated endocytosis to enter host cells as is seen with animal viruses. For many plant viruses to be transferred from plant to plant, damage to some of the plants' cells must occur to allow the virus to enter a new host. This damage is often caused by weather, insects, animals, fire, or human activities like farming or landscaping. Additionally, plant offspring may inherit viral diseases from parent plants. Plant viruses can be transmitted by a variety of vectors, through contact with an infected plant's sap, by living organisms such as insects and nematodes, and through pollen. When plants viruses are transferred between different plants, this is known ashorizontal transmission, and when they are inherited from a parent, this is called vertical transmission.

Symptoms of viral diseases vary according to the virus and its host (Table). One common symptom is hyperplasia, the abnormal proliferation of cells that causes the appearance of plant tumors known as galls. Other viruses induce hypoplasia, or decreased cell growth, in the leaves of plants, causing thin, yellow areas to appear. Still other viruses affect the plant by directly killing plant cells, a process known as cell necrosis. Other symptoms of plant viruses include malformed leaves, black streaks on the stems of the plants, altered growth of stems, leaves, or fruits, and ring spots, which are circular or linear areas of discoloration found in a leaf.

Some Common Symptoms of Plant Viral Diseases	
Symptom	Appears as
Hyperplasia	Galls (tumors)

Some Common Symptoms of Plant Viral Diseases	
Symptom	Appears as
Hypoplasia	Thinned, yellow splotches on leaves
Cell necrosis	Dead, blackened stems, leaves, or fruit
Abnormal growth patterns	Malformed stems, leaves, or fruit
Discoloration	Yellow, red, or black lines, or rings in stems, leaves, or fruit

Plant viruses can seriously disrupt crop growth and development, significantly affecting our food supply. They are responsible for poor crop quality and quantity globally, and can bring about huge economic losses annually. Others viruses may damage plants used in landscaping. Some viruses that infect agricultural food plants include the name of the plant they infect, such as tomato spotted wilt virus, bean common mosaic virus, and cucumber mosaic virus. In plants used for landscaping, two of the most common viruses are peony ring spot and rose mosaic virus. There are far too many plant viruses to discuss each in detail, but symptoms of bean common mosaic virus result in lowered bean production and stunted, unproductive plants. In the ornamental rose, the rose mosaic disease causes wavy yellow lines and colored splotches on the leaves of the plant.

Section Summary

Viral replication within a living cell always produces changes in the cell, sometimes resulting in cell death and sometimes slowly killing the infected cells. There are six basic stages in the virus replication cycle: attachment, penetration, uncoating, replication, assembly, and release. A viral infection may be productive, resulting in new virions, or nonproductive, which means that the virus remains inside the cell without producing new virions. Bacteriophages are viruses that infect bacteria. They have two different modes of replication: the lytic cycle, where the virus replicates and bursts out of the bacteria, and the lysogenic cycle, which involves the incorporation of the viral genome into the bacterial host genome. Animal viruses cause a variety of infections, with some causing chronic symptoms (hepatitis C), some intermittent symptoms (latent viruses such a herpes simplex virus 1), and others that cause very few symptoms, if any (human herpesviruses 6 and 7). Oncogenic viruses in animals have the ability to cause cancer by interfering with the regulation of the host cell cycle. Viruses of plants are responsible for significant economic damage in both agriculture and plants used for ornamentation.

Art Connections

Figure Influenza virus is packaged in a viral envelope that fuses with the plasma membrane. This way, the virus can exit the host cell without killing it. What advantage does the virus gain by keeping the host cell alive?

Figure Which of the following statements is false?

1. In the lytic cycle, new phage are produced and released into the environment.
2. In the lysogenic cycle, phage DNA is incorporated into the host genome.
3. An environmental stressor can cause the phage to initiate the lysogenic cycle.
4. Cell lysis only occurs in the lytic cycle.

Review Questions

Which statement is *not* true of viral replication?

1. A lysogenic cycle kills the host cell.
2. There are six basic steps in the viral replication cycle.
3. Viral replication does not affect host cell function.

4. Newly released virions can infect adjacent cells.

Which statement is true of viral replication?

1. In the process of apoptosis, the cell survives.
2. During attachment, the virus attaches at specific sites on the cell surface.
3. The viral capsid helps the host cell produce more copies of the viral genome.
4. mRNA works outside of the host cell to produce enzymes and proteins.

Which statement is true of reverse transcriptase?

1. It is a nucleic acid.
2. It infects cells.
3. It transcribes RNA to make DNA.
4. It is a lipid.

Oncogenic virus cores can be_____.

1. RNA
2. DNA
3. neither RNA nor DNA
4. either RNA or DNA

Which is true of DNA viruses?

1. They use the host cell's machinery to produce new copies of their genome.
2. They all have envelopes.
3. They are the only kind of viruses that can cause cancer.
4. They are not important plant pathogens.

A bacteriophage can infect _____.

1. the lungs
2. viruses
3. prions
4. bacteria

Free Response

Why can't dogs catch the measles?

One of the first and most important targets for drugs to fight infection with HIV (a retrovirus) is the reverse transcriptase enzyme. Why?

In this section, you were introduced to different types of viruses and viral diseases. Briefly discuss the most interesting or surprising thing you learned about viruses.

Although plant viruses cannot infect humans, what are some of the ways in which they affect humans?

Glossary

acute disease disease where the symptoms rise and fall within a short period of time

asymptomatic disease disease where there are no symptoms and the individual is unaware of being infected unless lab tests are performed

AZT anti-HIV drug that inhibits the viral enzyme reverse transcriptase

bacteriophage virus that infects bacteria

budding method of exit from the cell used in certain animal viruses, where virions leave the cell individually by capturing a piece of the host plasma membrane

cell necrosis cell death

chronic infection describes when the virus persists in the body for a long period of time

cytopathic causing cell damage

fusion method of entry by some enveloped viruses, where the viral envelope fuses with the plasma membrane of the host cell

gall appearance of a plant tumor

horizontal transmission transmission of a disease from parent to offspring

hyperplasia abnormally high cell growth and division

hypoplasia abnormally low cell growth and division

intermittent symptom symptom that occurs periodically

latency virus that remains in the body for a long period of time but only causes intermittent symptoms

lysis bursting of a cell

lytic cycle type of virus replication in which virions are released through lysis, or bursting, of the cell

lysogenic cycle type of virus replication in which the viral genome is incorporated into the genome of the host cell

oncogenic virus virus that has the ability to cause cancer

permissive cell type that is able to support productive replication of a virus

productive viral infection that leads to the production of new virions

prophage phage DNA that is incorporated into the host cell genome

vertical transmission transmission of disease between unrelated individuals

Licensing & Attributions

CC licensed content, Shared previously

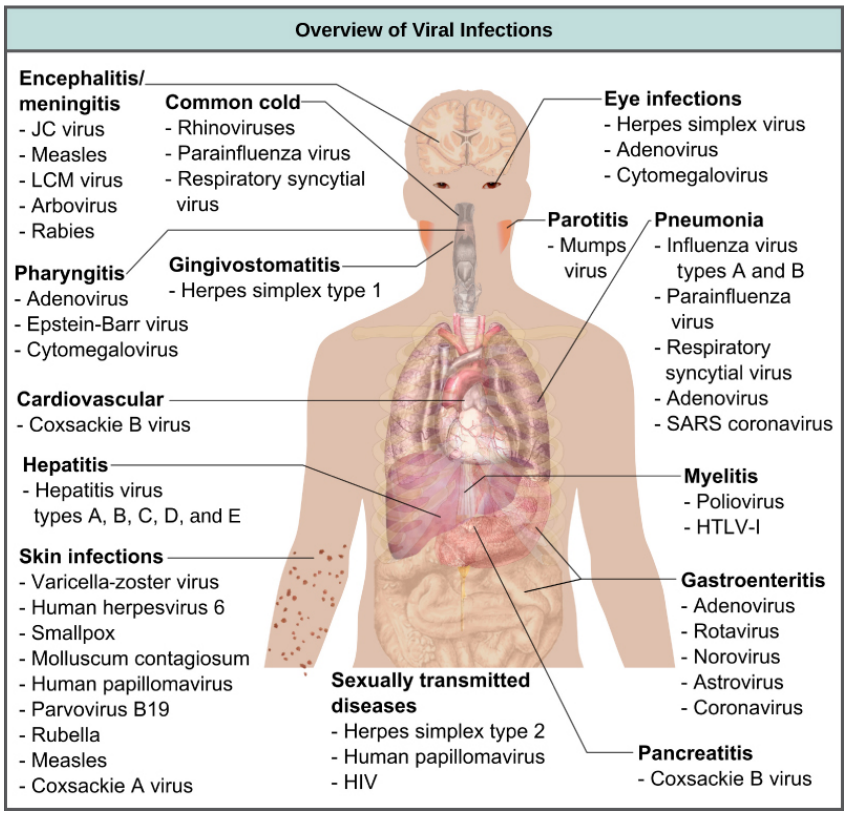
- Biology. **Authored by**: OpenStax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License**: *CC BY: Attribution*

PREVENTION AND TREATMENT OF VIRAL INFECTIONS

Prevention and Treatment of Viral Infections

Summary

Viruses cause a variety of diseases in animals, including humans, ranging from the common cold to potentially fatal illnesses like meningitis (Figure). These diseases can be treated by antiviral drugs or by vaccines, but some viruses, such as HIV, are capable of both avoiding the immune response and mutating to become resistant to antiviral drugs.



Viruses can cause dozens of ailments in humans, ranging from mild illnesses to serious diseases. (credit: modification of work by Mikael Häggström)

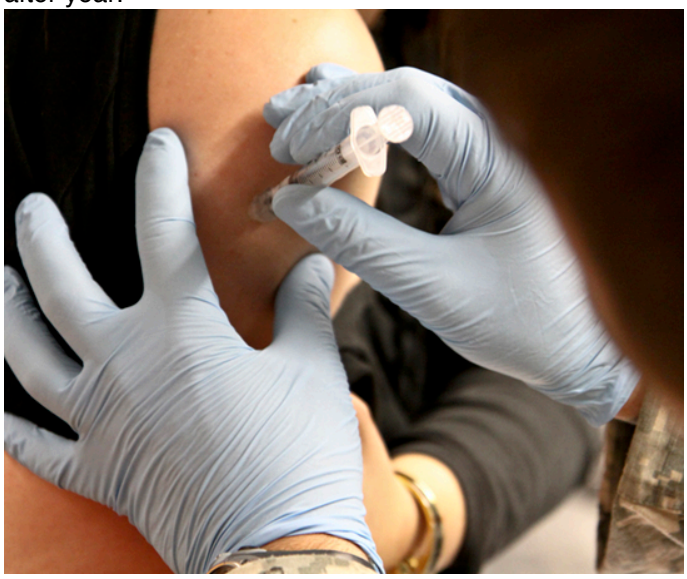
Vaccines for Prevention

While we do have limited numbers of effective antiviral drugs, such as those used to treat HIV and influenza, the primary method of controlling viral disease is by vaccination, which is intended to prevent outbreaks by building immunity to a virus or virus family (Figure). Vaccines may be prepared using live viruses, killed viruses, or molecular subunits of the virus. The killed viral vaccines and subunit viruses are both incapable of causing disease.

Live viral vaccines are designed in the laboratory to cause few symptoms in recipients while giving them protective immunity against future infections. Polio was one disease that represented a milestone in the use of vaccines. Mass immunization campaigns in the 1950s (killed vaccine) and 1960s (live vaccine) significantly reduced the incidence of the disease, which caused muscle paralysis in children and generated a great amount of fear in the general population when regional epidemics occurred. The success of the polio vaccine paved the way for the routine dispensation of childhood vaccines against measles, mumps, rubella, chickenpox, and other diseases.

The danger of using live vaccines, which are usually more effective than killed vaccines, is the low but significant danger that these viruses will revert to their disease-causing form by back mutations. Live vaccines are usually made by attenuating(weakening) the “wild-type” (disease-causing) virus by growing it in the laboratory in tissues or at temperatures different from what the virus is accustomed to in the host. Adaptations to these new cells or temperatures induce mutations in the genomes of the virus, allowing it to grow better in the laboratory while inhibiting its ability to cause disease when reintroduced into conditions found in the host. These attenuated viruses thus still cause infection, but they do not grow very well, allowing the immune response to develop in time to prevent major disease. Back mutations occur when the vaccine undergoes mutations in the host such that it readapts to the host and can again cause disease, which can then be spread to other humans in an epidemic. This type of scenario happened as recently as 2007 in Nigeria where mutations in a polio vaccine led to an epidemic of polio in that country.

Some vaccines are in continuous development because certain viruses, such as influenza and HIV, have a high mutation rate compared to other viruses and normal host cells. With influenza, mutations in the surface molecules of the virus help the organism evade the protective immunity that may have been obtained in a previous influenza season, making it necessary for individuals to get vaccinated every year. Other viruses, such as those that cause the childhood diseases measles, mumps, and rubella, mutate so infrequently that the same vaccine is used year after year.



Vaccinations are designed to boost immunity to a virus to prevent infection. (credit: USACE Europe District)

Link to Learning

Watch this NOVA video to learn how microbiologists are attempting to replicate the deadly 1918 Spanish influenza virus so they can understand more about virology.

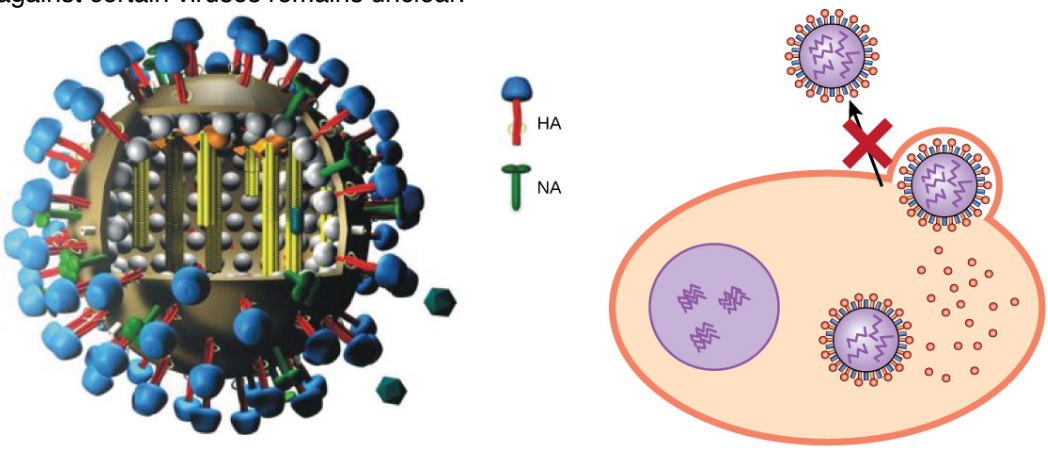
Vaccines and Anti-viral Drugs for Treatment

In some cases, vaccines can be used to treat an active viral infection. The concept behind this is that by giving the vaccine, immunity is boosted without adding more disease-causing virus. In the case of rabies, a fatal neurological disease transmitted via the saliva of rabies virus-infected animals, the progression of the disease from the time of the animal bite to the time it enters the central nervous system may be 2 weeks or longer. This is enough time to vaccinate an individual who suspects that they have been bitten by a rabid animal, and their boosted immune response is sufficient to prevent the virus from entering nervous tissue. Thus, the potentially fatal neurological consequences of the disease are averted, and the individual only has to recover from the infected bite. This approach is also being used for the treatment of Ebola, one of the fastest and most deadly viruses on earth. Transmitted by bats and great apes, this disease can cause death in 70–90 percent of infected humans within 2 weeks. Using newly developed vaccines that boost the immune response in this way, there is hope that affected individuals will be better able to control the virus, potentially saving a greater percentage of infected persons from a rapid and very painful death.

Another way of treating viral infections is the use of antiviral drugs. These drugs often have limited success in curing viral disease, but in many cases, they have been used to control and reduce symptoms for a wide variety of viral diseases. For most viruses, these drugs can inhibit the virus by blocking the actions of one or more of its proteins. It is important that the targeted proteins be encoded by viral genes and that these molecules are not present in a healthy host cell. In this way, viral growth is inhibited without damaging the host. There are large numbers of antiviral drugs available to treat infections, some specific for a particular virus and others that can affect multiple viruses.

Antivirals have been developed to treat genital herpes (herpes simplex II) and influenza. For genital herpes, drugs such as acyclovir can reduce the number and duration of episodes of active viral disease, during which patients develop viral lesions in their skin cells. As the virus remains latent in nervous tissue of the body for life, this drug is

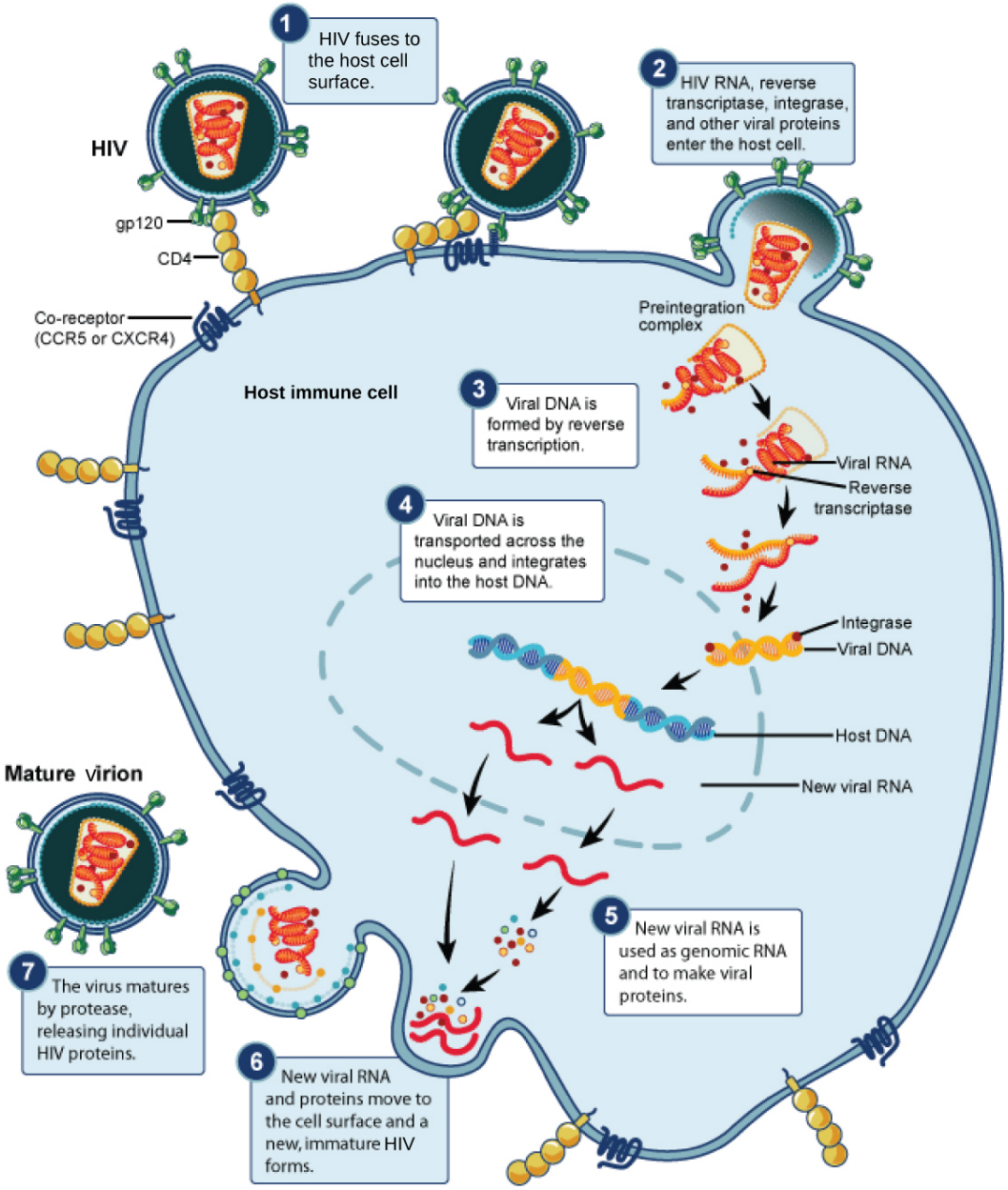
not curative but can make the symptoms of the disease more manageable. For influenza, drugs like Tamiflu (oseltamivir) (Figure) can reduce the duration of “flu” symptoms by 1 or 2 days, but the drug does not prevent symptoms entirely. Tamiflu works by inhibiting an enzyme (viral neuraminidase) that allows new virions to leave their infected cells. Thus, Tamiflu inhibits the spread of virus from infected to uninfected cells. Other antiviral drugs, such as Ribavirin, have been used to treat a variety of viral infections, although its mechanism of action against certain viruses remains unclear.



(a) Tamiflu inhibits a viral enzyme called neuraminidase (NA) found in the influenza viral envelope. (b) Neuraminidase cleaves the connection between viral hemagglutinin (HA), also found in the viral envelope, and glycoproteins on the host cell surface. Inhibition of neuraminidase prevents the virus from detaching from the host cell, thereby blocking further infection. (credit a: modification of work by M. Eickmann)

By far, the most successful use of antivirals has been in the treatment of the retrovirus HIV, which causes a disease that, if untreated, is usually fatal within 10–12 years after infection. Anti-HIV drugs have been able to control viral replication to the point that individuals receiving these drugs survive for a significantly longer time than the untreated.

Anti-HIV drugs inhibit viral replication at many different phases of the HIV replicative cycle (Figure). Drugs have been developed that inhibit the fusion of the HIV viral envelope with the plasma membrane of the host cell (fusion inhibitors), the conversion of its RNA genome into double-stranded DNA (reverse transcriptase inhibitors), the integration of the viral DNA into the host genome (integrase inhibitors), and the processing of viral proteins (protease inhibitors).



HIV, an enveloped, icosahedral virus, attaches to the CD4 receptor of an immune cell and fuses with the cell membrane. Viral contents are released into the cell, where viral enzymes convert the single-stranded RNA genome into DNA and incorporate it into the host genome. (credit: NIAID, NIH)

When any of these drugs are used individually, the high mutation rate of the virus allows it to easily and rapidly develop resistance to the drug, limiting the drug's effectiveness. The breakthrough in the treatment of HIV was the development of HAART, highly active anti-retroviral therapy, which involves a mixture of different drugs, sometimes called a drug "cocktail." By attacking the virus at different stages of its replicative cycle, it is much more difficult for the virus to develop resistance to multiple drugs at the same time. Still, even with the use of combination HAART therapy, there is concern that, over time, the virus will develop resistance to this therapy. Thus, new anti-HIV drugs are constantly being developed with the hope of continuing the battle against this highly fatal virus.

EVERYDAY CONNECTION

Applied VirologyThe study of viruses has led to the development of a variety of new ways to treat non-viral diseases. Viruses have been used in gene therapy. Gene therapy is used to treat genetic diseases such as severe combined immunodeficiency (SCID), a heritable, recessive disease in which children are born with severely compromised immune systems. One common type of SCID is due to the lack of an enzyme, adenosine deaminase (ADA), which breaks down purine bases. To treat this disease by gene therapy, bone marrow cells are

taken from a SCID patient and the ADA gene is inserted. This is where viruses come in, and their use relies on their ability to penetrate living cells and bring genes in with them. Viruses such as adenovirus, an upper respiratory human virus, are modified by the addition of the ADA gene, and the virus then transports this gene into the cell. The modified cells, now capable of making ADA, are then given back to the patients in the hope of curing them. Gene therapy using viruses as carrier of genes (viral vectors), although still experimental, holds promise for the treatment of many genetic diseases. Still, many technological problems need to be solved for this approach to be a viable method for treating genetic disease.

Another medical use for viruses relies on their specificity and ability to kill the cells they infect.Oncolytic viruses are engineered in the laboratory specifically to attack and kill cancer cells. A genetically modified adenovirus known as H101 has been used since 2005 in clinical trials in China to treat head and neck cancers. The results have been promising, with a greater short-term response rate to the combination of chemotherapy and viral therapy than to chemotherapy treatment alone. This ongoing research may herald the beginning of a new age of cancer therapy, where viruses are engineered to find and specifically kill cancer cells, regardless of where in the body they may have spread.

A third use of viruses in medicine relies on their specificity and involves using bacteriophages in the treatment of bacterial infections. Bacterial diseases have been treated with antibiotics since the 1940s. However, over time, many bacteria have developed resistance to antibiotics. A good example is methicillin-resistant *Staphylococcus aureus* (MRSA, pronounced “mersa”), an infection commonly acquired in hospitals. This bacterium is resistant to a variety of antibiotics, making it difficult to treat. The use of bacteriophages specific for such bacteria would bypass their resistance to antibiotics and specifically kill them. Although phage therapy is in use in the Republic of Georgia to treat antibiotic-resistant bacteria, its use to treat human diseases has not been approved in most countries. However, the safety of the treatment was confirmed in the United States when the U.S. Food and Drug Administration approved spraying meats with bacteriophages to destroy the food pathogen*Listeria.* As more and more antibiotic-resistant strains of bacteria evolve, the use of bacteriophages might be a potential solution to the problem, and the development of phage therapy is of much interest to researchers worldwide.

Section Summary

Viruses cause a variety of diseases in humans. Many of these diseases can be prevented by the use of viral vaccines, which stimulate protective immunity against the virus without causing major disease. Viral vaccines may also be used in active viral infections, boosting the ability of the immune system to control or destroy the virus. A series of antiviral drugs that target enzymes and other protein products of viral genes have been developed and used with mixed success. Combinations of anti-HIV drugs have been used to effectively control the virus, extending the lifespans of infected individuals. Viruses have many uses in medicines, such as in the treatment of genetic disorders, cancer, and bacterial infections.

Review Questions

Which of the following is NOT used to treat active viral disease?

1. vaccines
2. antiviral drugs
3. antibiotics
4. phage therapy

Vaccines_____.

1. are similar to viroids
2. are only needed once
3. kill viruses
4. stimulate an immune response

Free Response

Why is immunization after being bitten by a rabid animal so effective and why aren't people vaccinated for rabies like dogs and cats are?

Glossary

attenuation weakening of a virus during vaccine development

back mutation when a live virus vaccine reverts back to it disease-causing phenotype

gene therapy treatment of genetic disease by adding genes, using viruses to carry the new genes inside the cell

oncolytic virus virus engineered to specifically infect and kill cancer cells

phage therapy treatment of bacterial diseases using bacteriophages specific to a particular bacterium

vaccine weakened solution of virus components, viruses, or other agents that produce an immune response

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by:** Open Stax. **License:** *CC BY: Attribution*

OTHER ACELLULAR ENTITIES: PRIONS AND VIROIDS

Other Acellular Entities: Prions and Viroids

Summary

Prions and viroids are pathogens (agents with the ability to cause disease) that have simpler structures than viruses but, in the case of prions, still can produce deadly diseases.

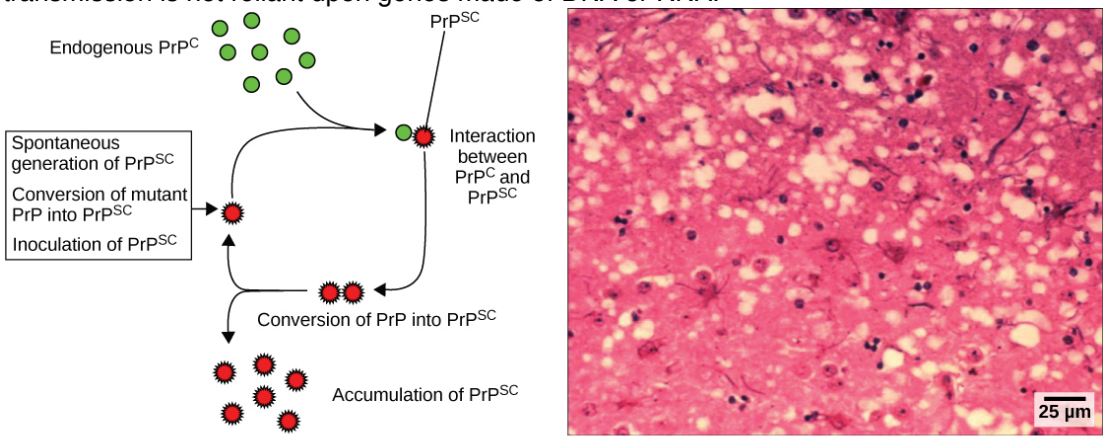
Prions

Prions, so-called because they are proteinaceous, are infectious particles—smaller than viruses—that contain no nucleic acids (neither DNA nor RNA). Historically, the idea of an infectious agent that did not use nucleic acids was considered impossible, but pioneering work by Nobel Prize-winning biologist Stanley Prusiner has convinced the majority of biologists that such agents do indeed exist.

Fatal neurodegenerative diseases, such as kuru in humans and bovine spongiform encephalopathy (BSE) in cattle (commonly known as "mad cow disease") were shown to be transmitted by prions. The disease was spread by the consumption of meat, nervous tissue, or internal organs between members of the same species. Kuru, native to humans in Papua New Guinea, was spread from human to human via ritualistic cannibalism. BSE, originally detected in the United Kingdom, was spread between cattle by the practice of including cattle nervous tissue in feed for other cattle. Individuals with kuru and BSE show symptoms of loss of motor control and unusual behaviors, such as uncontrolled bursts of laughter with kuru, followed by death. Kuru was controlled by inducing the population to abandon its ritualistic cannibalism.

On the other hand, BSE was initially thought to only affect cattle. Cattle dying of the disease were shown to have developed lesions or “holes” in the brain, causing the brain tissue to resemble a sponge. Later on in the outbreak, however, it was shown that a similar encephalopathy in humans known as variant Creutzfeldt-Jakob disease (CJD) could be acquired from eating beef from animals with BSE, sparking bans by various countries on the importation of British beef and causing considerable economic damage to the British beef industry (Figure). BSE still exists in various areas, and although a rare disease, individuals that acquire CJD are difficult to treat. The disease can be spread from human to human by blood, so many countries have banned blood donation from regions associated with BSE.

The cause of spongiform encephalopathies, such as kuru and BSE, is an infectious structural variant of a normal cellular protein called PrP (prion protein). It is this variant that constitutes the prion particle. PrP exists in two forms, PrP^{C}, the normal form of the protein, and PrP^{SC}, the infectious form. Once introduced into the body, the PrP^{SC} contained within the prion binds to PrP^{C} and converts it to PrP^{SC}. This leads to an exponential increase of the PrP^{SC} protein, which aggregates. PrP^{SC} is folded abnormally, and the resulting conformation (shape) is directly responsible for the lesions seen in the brains of infected cattle. Thus, although not without some detractors among scientists, the prion seems likely to be an entirely new form of infectious agent, the first one found whose transmission is not reliant upon genes made of DNA or RNA.



(a) (b)

(a) Endogenous normal prion protein (PrP^{C}) is converted into the disease-causing form (PrP^{SC}) when it encounters this variant form of the protein. PrP^{SC} may arise spontaneously in brain tissue, especially if a mutant form of the protein is present, or it may occur via the spread of misfolded prions consumed in food into brain tissue. (b) This prion-infected brain tissue, visualized using light microscopy, shows the vacuoles that give it a spongy texture, typical of transmissible spongiform encephalopathies. (credit b: modification of work by Dr. Al Jenny, USDA APHIS; scale-bar data from Matt Russell)

Viroids

Viroids are plant pathogens: small, single-stranded, circular RNA particles that are much simpler than a virus. They do not have a capsid or outer envelope, but like viruses can reproduce only within a host cell. Viroids do not, however, manufacture any proteins, and they only produce a single, specific RNA molecule. Human diseases caused by viroids have yet to be identified.

Viroids are known to infect plants (Figure) and are responsible for crop failures and the loss of millions of dollars in agricultural revenue each year. Some of the plants they infect include potatoes, cucumbers, tomatoes, chrysanthemums, avocados, and coconut palms.

These potatoes have been infected by the potato spindle tuber viroid (PSTV), which is typically spread when infected knives are used to cut healthy potatoes, which are then planted. (credit: Pamela Roberts, University of Florida Institute of Food and Agricultural Sciences, USDA ARS)

CAREER CONNECTION

VirologistVirology is the study of viruses, and a virologist is an individual trained in this discipline. Training in virology can lead to many different career paths. Virologists are actively involved in academic research and teaching in colleges and medical schools. Some virologists treat patients or are involved in the generation and production of vaccines. They might participate in epidemiologic studies (Figure) or become science writers, to name just a few possible careers.


This virologist is engaged in fieldwork, sampling eggs from this nest for avian influenza. (credit: Don Becker, USGS EROS, U.S. Fish and Wildlife Service)

If you think you may be interested in a career in virology, find a mentor in the field. Many large medical centers have departments of virology, and smaller hospitals usually have virology labs within their microbiology departments. Volunteer in a virology lab for a semester or work in one over the summer. Discussing the profession and getting a first-hand look at the work will help you decide whether a career in virology is right for you. The American Society of Virology's website is a good resource for information regarding training and careers in virology.

Section Summary

Prions are infectious agents that consist of protein, but no DNA or RNA, and seem to produce their deadly effects by duplicating their shapes and accumulating in tissues. They are thought to contribute to several progressive brain disorders, including mad cow disease and Creutzfeldt-Jakob disease. Viroids are single-stranded RNA pathogens that infect plants. Their presence can have a severe impact on the agriculture industry.

Review Questions

Which of the following is not associated with prions?

1. replicating shapes
2. mad cow disease
3. DNA
4. toxic proteins

Which statement is true of viroids?

1. They are single-stranded RNA particles.
2. They reproduce only outside of the cell.
3. They produce proteins.
4. They affect both plants and animals.

Free Response

Prions are responsible for variant Creutzfeldt-Jakob Disease, which has resulted in over 100 human deaths in Great Britain during the last 10 years. How do humans obtain this disease?

How are viroids like viruses?

Glossary

pathogen agent with the ability to cause disease

prion infectious particle that consists of proteins that replicate without DNA or RNA

PrP^c normal prion protein

PrP^{sc} infectious form of a prion protein

viroid plant pathogen that produces only a single, specific RNA

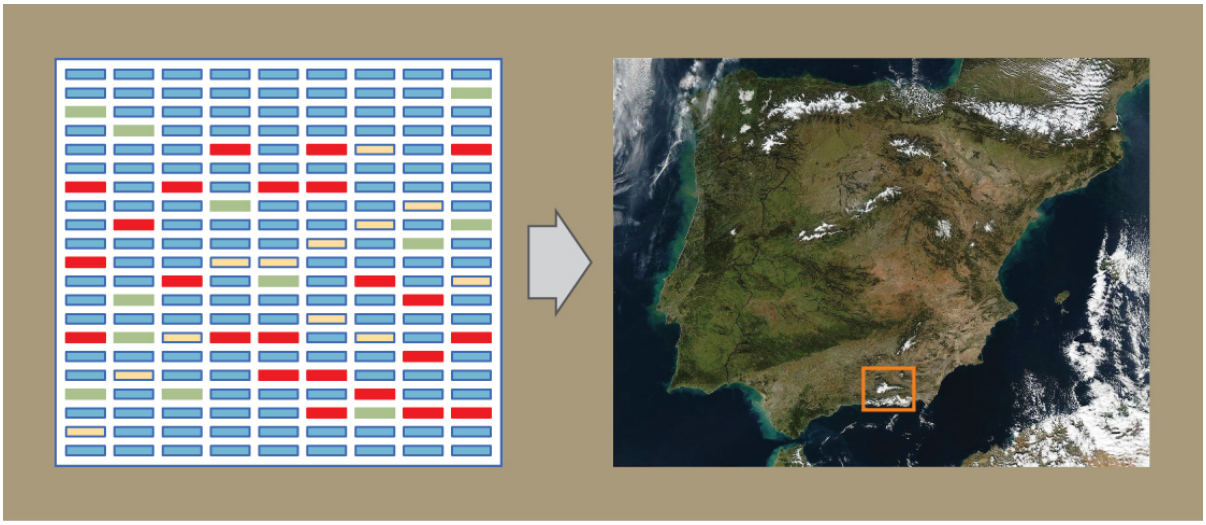
Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: Open Stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biologytt. **License**: *CC BY: Attribution*

BIOTECHNOLOGY AND GENOMICS

INTRODUCTION: BIOTECHNOLOGY



In genomics, the DNA of different organisms is compared, enabling scientists to create maps with which to navigate the DNA of different organisms. (credit “map”: modification of photo by NASA)

The study of nucleic acids began with the discovery of DNA, progressed to the study of genes and small fragments, and has now exploded to the field of genomics. Genomics is the study of entire genomes, including the complete set of genes, their nucleotide sequence and organization, and their interactions within a species and with other species. The advances in genomics have been made possible by DNA sequencing technology. Just as information technology has led to Google maps that enable people to get detailed information about locations around the globe, genomic information is used to create similar maps of the DNA of different organisms. These findings have helped anthropologists to better understand human migration and have aided the field of medicine through the mapping of human genetic diseases. The ways in which genomic information can contribute to scientific understanding are varied and quickly growing.

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: Open Stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biology. **License**: *CC BY: Attribution*

MAPPING GENOMES

Objectives:

By the end of this section, you will be able to:

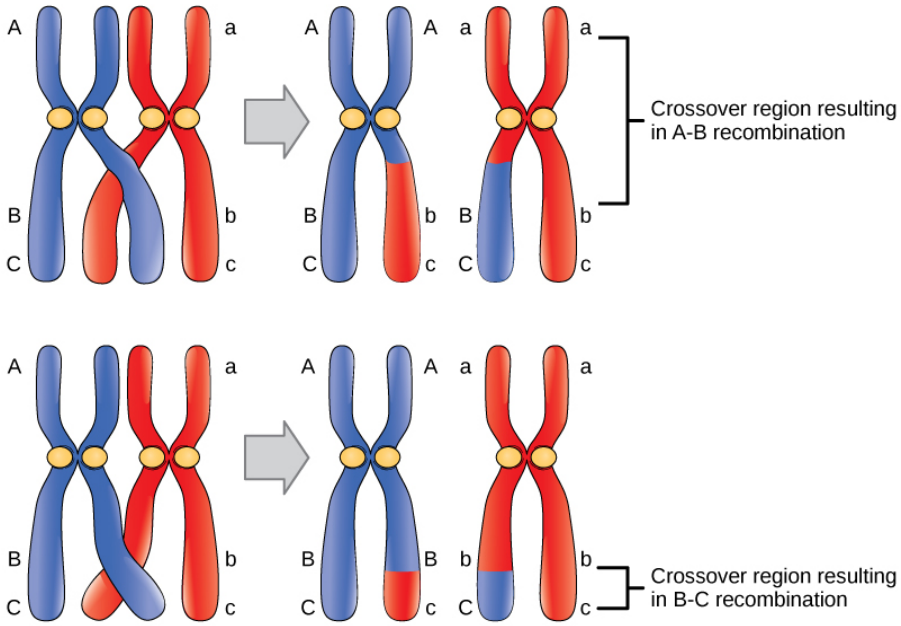
- Define genomics
- Describe genetic and physical maps
- Describe genomic mapping methods

Genomics is the study of entire genomes, including the complete set of genes, their nucleotide sequence and organization, and their interactions within a species and with other species. Genome mapping is the process of finding the locations of genes on each chromosome. The maps created by genome mapping are comparable to the maps that we use to navigate streets. Agenetic map is an illustration that lists genes and their location on a chromosome. Genetic maps provide the big picture (similar to a map of interstate highways) and use genetic markers (similar to landmarks). A genetic marker is a gene or sequence on a chromosome that co-segregates (shows genetic linkage) with a specific trait. Early geneticists called this linkage analysis. Physical maps present the intimate details of smaller regions of the chromosomes (similar to a detailed road map). A physical map is a representation of the physical distance, in nucleotides, between genes or genetic markers. Both genetic linkage maps and physical maps are required to build a complete picture of the genome. Having a complete map of the genome makes it easier for researchers to study individual genes. Human genome maps help researchers in their efforts to identify human disease-causing genes related to illnesses like cancer, heart disease, and cystic fibrosis. Genome mapping can be used in a variety of other applications, such as using live microbes to clean up pollutants or even prevent pollution. Research involving plant genome mapping may lead to producing higher crop yields or developing plants that better adapt to climate change.

Genetic Maps

The study of genetic maps begins with linkage analysis, a procedure that analyzes the recombination frequency between genes to determine if they are linked or show independent assortment. The term *linkage* was used before the discovery of DNA. Early geneticists relied on the observation of phenotypic changes to understand the genotype of an organism. Shortly after Gregor Mendel (the father of modern genetics) proposed that traits were determined by what are now known as genes, other researchers observed that different traits were often inherited together, and thereby deduced that the genes were physically linked by being located on the same chromosome. The mapping of genes relative to each other based on linkage analysis led to the development of the first genetic maps.

Observations that certain traits were always linked and certain others were not linked came from studying the offspring of crosses between parents with different traits. For example, in experiments performed on the garden pea, it was discovered that the color of the flower and shape of the plant's pollen were linked traits, and therefore the genes encoding these traits were in close proximity on the same chromosome. The exchange of DNA between homologous pairs of chromosomes is called genetic recombination, which occurs by the crossing over of DNA between homologous strands of DNA, such as nonsister chromatids. Linkage analysis involves studying the recombination frequency between any two genes. The greater the distance between two genes, the higher the chance that a recombination event will occur between them, and the higher the recombination frequency between them. Two possibilities for recombination between two nonsister chromatids during meiosis are shown in Figure. If the recombination frequency between two genes is less than 50 percent, they are said to be linked.



Crossover may occur at different locations on the chromosome. Recombination between genes *A* and *B* is more frequent than recombination between genes *B* and *C* because genes *A* and *B* are farther apart; a crossover is therefore more likely to occur between them.

The generation of genetic maps requires markers, just as a road map requires landmarks (such as rivers and mountains). Early genetic maps were based on the use of known genes as markers. More sophisticated markers, including those based on non-coding DNA, are now used to compare the genomes of individuals in a population. Although individuals of a given species are genetically similar, they are not identical; every individual has a unique set of traits. These minor differences in the genome between individuals in a population are useful for the purposes of genetic mapping. In general, a good genetic marker is a region on the chromosome that shows variability or polymorphism (multiple forms) in the population.

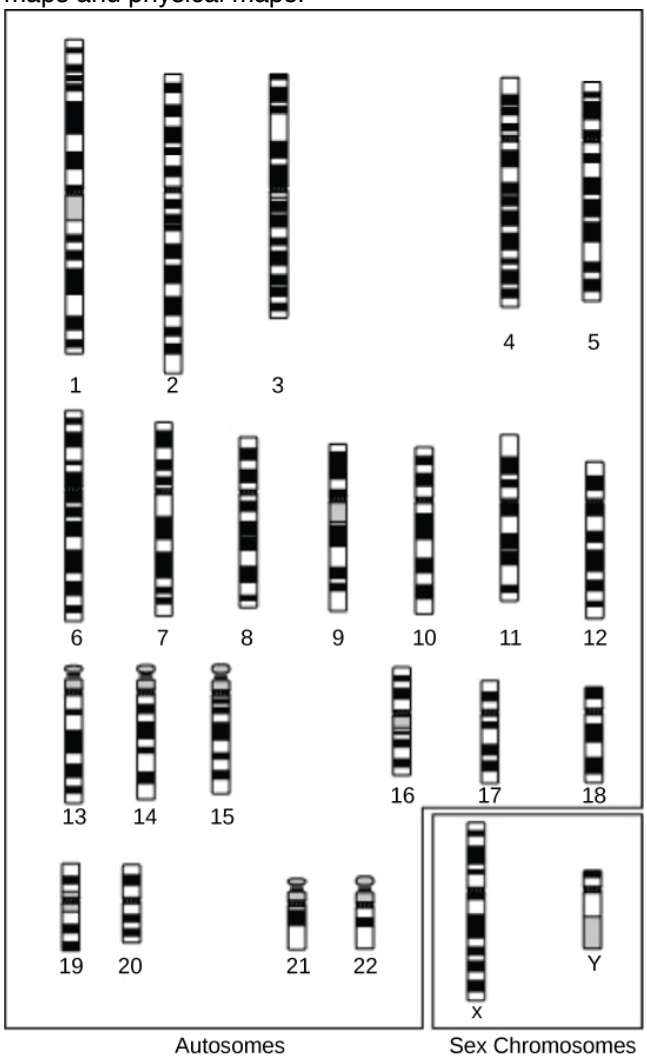
Some genetic markers used in generating genetic maps are restriction fragment length polymorphisms (RFLP), variable number of tandem repeats (VNTRs), microsatellite polymorphisms, and the single nucleotide polymorphisms (SNPs). RFLPs (sometimes pronounced “rif-lips”) are detected when the DNA of an individual is cut with a restriction endonuclease that recognizes specific sequences in the DNA to generate a series of DNA fragments, which are then analyzed by gel electrophoresis. The DNA of every individual will give rise to a unique pattern of bands when cut with a particular set of restriction endonucleases; this is sometimes referred to as an individual’s DNA “fingerprint.” Certain regions of the chromosome that are subject to polymorphism will lead to the generation of the unique banding pattern. VNTRs are repeated sets of nucleotides present in the non-coding regions of DNA. Non-coding, or “junk,” DNA has no known biological function; however, research shows that much of this DNA is actually transcribed. While its function is uncertain, it is certainly active, and it may be involved in the regulation of coding genes. The number of repeats may vary in individual organisms of a population. Microsatellite polymorphisms are similar to VNTRs, but the repeat unit is very small. SNPs are variations in a single nucleotide.

Because genetic maps rely completely on the natural process of recombination, mapping is affected by natural increases or decreases in the level of recombination in any given area of the genome. Some parts of the genome are recombination hotspots, whereas others do not show a propensity for recombination. For this reason, it is important to look at mapping information developed by multiple methods.

Physical Maps

A physical map provides detail of the actual physical distance between genetic markers, as well as the number of nucleotides. There are three methods used to create a physical map: cytogenetic mapping, radiation hybrid mapping, and sequence mapping.Cytogenetic mapping uses information obtained by microscopic analysis of stained sections of the chromosome (Figure). It is possible to determine the approximate distance between genetic markers using cytogenetic mapping, but not the exact distance (number of base pairs). Radiation hybrid

mapping uses radiation, such as x-rays, to break the DNA into fragments. The amount of radiation can be adjusted to create smaller or larger fragments. This technique overcomes the limitation of genetic mapping and is not affected by increased or decreased recombination frequency. Sequence mapping resulted from DNA sequencing technology that allowed for the creation of detailed physical maps with distances measured in terms of the number of base pairs. The creation of genomic libraries and complementary DNA (cDNA) libraries (collections of cloned sequences or all DNA from a genome) has sped up the process of physical mapping. A genetic site used to generate a physical map with sequencing technology (a sequence-tagged site, or STS) is a unique sequence in the genome with a known exact chromosomal location. An expressed sequence tag (EST) and a single sequence length polymorphism (SSLP) are common STSs. An EST is a short STS that is identified with cDNA libraries, while SSLPs are obtained from known genetic markers and provide a link between genetic maps and physical maps.



A cytogenetic map shows the appearance of a chromosome after it is stained and examined under a microscope. (credit: National Human Genome Research Institute)

Integration of Genetic and Physical Maps

Genetic maps provide the outline and physical maps provide the details. It is easy to understand why both types of genome mapping techniques are important to show the big picture. Information obtained from each technique is used in combination to study the genome. Genomic mapping is being used with different model organisms that are used for research. Genome mapping is still an ongoing process, and as more advanced techniques are developed, more advances are expected. Genome mapping is similar to completing a complicated puzzle using every piece of available data. Mapping information generated in laboratories all over the world is entered into central databases, such as GenBank at the National Center for Biotechnology Information (NCBI). Efforts are

being made to make the information more easily accessible to researchers and the general public. Just as we use global positioning systems instead of paper maps to navigate through roadways, NCBI has created a genome viewer tool to simplify the data-mining process.
SCIENTIFIC METHOD CONNECTION

How to Use a Genome Map Viewer

Problem statement: Do the human, macaque, and mouse genomes contain common DNA sequences?

Develop a hypothesis.

To test the hypothesis, click this link.

In Search box on the left panel, type any gene name or phenotypic characteristic, such as iris pigmentation (eye color). Select the species you want to study, and then press Enter. The genome map viewer will indicate which chromosome encodes the gene in your search. Click each hit in the genome viewer for more detailed information. This type of search is the most basic use of the genome viewer; it can also be used to compare sequences between species, as well as many other complicated tasks.

Is the hypothesis correct? Why or why not?

Link to Learning

Online Mendelian Inheritance in Man (OMIM) is a searchable online catalog of human genes and genetic disorders. This website shows genome mapping information, and also details the history and research of each trait and disorder. Click this link to search for traits (such as handedness) and genetic disorders (such as diabetes).

Section Summary

Genome mapping is similar to solving a big, complicated puzzle with pieces of information coming from laboratories all over the world. Genetic maps provide an outline for the location of genes within a genome, and they estimate the distance between genes and genetic markers on the basis of recombination frequencies during meiosis. Physical maps provide detailed information about the physical distance between the genes. The most detailed information is available through sequence mapping. Information from all mapping and sequencing sources is combined to study an entire genome.

Review Questions

ESTs are _____.

1. generated after a cDNA library is made
2. unique sequences in the genome
3. useful for mapping using sequence information
4. all of the above

Linkage analysis _____.

1. is used to create a physical map
2. is based on the natural recombination process
3. requires radiation hybrid mapping
4. involves breaking and re-joining of DNA artificially

Genetic recombination occurs by which process?

1. independent assortment
2. crossing over

3. chromosome segregation
4. sister chromatids

Individual genetic maps in a given species are:

1. genetically similar
2. genetically identical
3. genetically dissimilar
4. not useful in species analysis

Information obtained by microscopic analysis of stained chromosomes is used in:

1. radiation hybrid mapping
2. sequence mapping
3. RFLP mapping
4. cytogenetic mapping

Free Response

Why is so much effort being poured into genome mapping applications?

How could a genetic map of the human genome help find a cure for cancer?

Glossary

cytogenetic mapping technique that uses a microscope to create a map from stained chromosomes

expressed sequence tag (EST) short STS that is identified with cDNA

genetic map outline of genes and their location on a chromosome

genetic marker gene or sequence on a chromosome with a known location that is associated with a specific trait

genetic recombination exchange of DNA between homologous pairs of chromosomes

genome mapping process of finding the location of genes on each chromosome

cDNA library collection of cloned cDNA sequences

genomic library collection of cloned DNA which represents all of the sequences and fragments from a genome

genomics study of entire genomes including the complete set of genes, their nucleotide sequence and organization, and their interactions within a species and with other species

linkage analysis procedure that analyzes the recombination of genes to determine if they are linked

microsatellite polymorphism variation between individuals in the sequence and number of repeats of microsatellite DNA

physical map representation of the physical distance between genes or genetic markers

radiation hybrid mapping information obtained by fragmenting the chromosome with x-rays

restriction fragment length polymorphism (RFLP) variation between individuals in the length of DNA fragments generated by restriction endonucleases

sequence mapping mapping information obtained after DNA sequencing

single nucleotide polymorphism (SNP) variation between individuals in a single nucleotide

variable number of tandem repeats (VNTRs) variation in the number of tandem repeats between individuals in the population

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: Open Stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biologyt. **License**: *CC BY: Attribution*

WHOLE-GENOME SEQUENCING

Objectives:

By the end of this section, you will be able to:

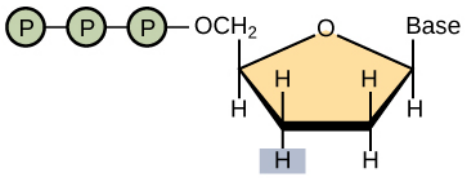
- Describe three types of sequencing
- Define whole-genome sequencing

Although there have been significant advances in the medical sciences in recent years, doctors are still confounded by some diseases, and they are using whole-genome sequencing to get to the bottom of the problem. Whole-genome sequencing is a process that determines the DNA sequence of an entire genome. Whole-genome sequencing is a brute-force approach to problem solving when there is a genetic basis at the core of a disease. Several laboratories now provide services to sequence, analyze, and interpret entire genomes.

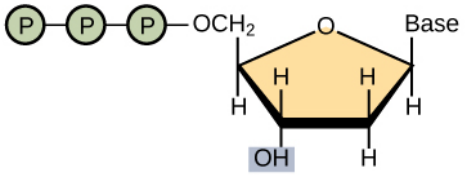
For example, whole-exome sequencing is a lower-cost alternative to whole genome sequencing. In exome sequencing, only the coding, exon-producing regions of the DNA are sequenced. In 2010, whole-exome sequencing was used to save a young boy whose intestines had multiple mysterious abscesses. The child had several colon operations with no relief. Finally, whole-exome sequencing was performed, which revealed a defect in a pathway that controls apoptosis (programmed cell death). A bone-marrow transplant was used to overcome this genetic disorder, leading to a cure for the boy. He was the first person to be successfully treated based on a diagnosis made by whole-exome sequencing. Today, human genome sequencing is more readily available and can be completed in a day or two for about $1000.

Strategies Used in Sequencing Projects

The basic sequencing technique used in all modern day sequencing projects is the chain termination method (also known as the dideoxy method), which was developed by Fred Sanger in the 1970s. The chain termination method involves DNA replication of a single-stranded template with the use of a primer and a regular deoxynucleotide (dNTP), which is a monomer, or a single unit, of DNA. The primer and dNTP are mixed with a small proportion of fluorescently labeled dideoxynucleotides (ddNTPs). The ddNTPs are monomers that are missing a hydroxyl group (–OH) at the site at which another nucleotide usually attaches to form a chain (Figure). Each ddNTP is labeled with a different color of fluorophore. Every time a ddNTP is incorporated in the growing complementary strand, it terminates the process of DNA replication, which results in multiple short strands of replicated DNA that are each terminated at a different point during replication. When the reaction mixture is processed by gel electrophoresis after being separated into single strands, the multiple newly replicated DNA strands form a ladder because of the differing sizes. Because the ddNTPs are fluorescently labeled, each band on the gel reflects the size of the DNA strand and the ddNTP that terminated the reaction. The different colors of the fluorophore-labeled ddNTPs help identify the ddNTP incorporated at that position. Reading the gel on the basis of the color of each band on the ladder produces the sequence of the template strand (Figure).

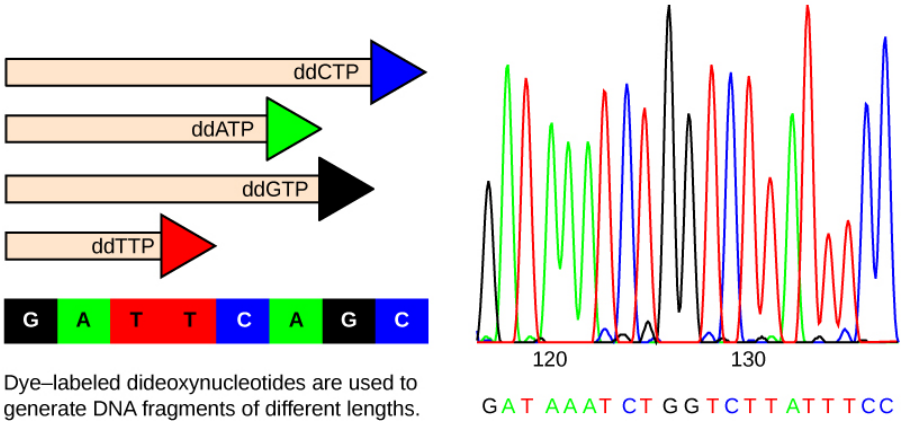


Dideoxynucleotide (ddNTP)



Deoxynucleotide (dNTP)

A dideoxynucleotide is similar in structure to a deoxynucleotide, but is missing the 3′ hydroxyl group (indicated by the box). When a dideoxynucleotide is incorporated into a DNA strand, DNA synthesis stops.



Dye–labeled dideoxynucleotides are used to generate DNA fragments of different lengths.

Frederick Sanger's dideoxy chain termination method is illustrated. Using dideoxynucleotides, the DNA fragment can be terminated at different points. The DNA is separated on the basis of size, and these bands, based on the size of the fragments, can be read.

Early Strategies: Shotgun Sequencing and Pair-Wise End Sequencing

In shotgun sequencing method, several copies of a DNA fragment are cut randomly into many smaller pieces (somewhat like what happens to a round shot cartridge when fired from a shotgun). All of the segments are then sequenced using the chain-sequencing method. Then, with the help of a computer, the fragments are analyzed to see where their sequences overlap. By matching up overlapping sequences at the end of each fragment, the entire DNA sequence can be reformed. A larger sequence that is assembled from overlapping shorter sequences is called a contig. As an analogy, consider that someone has four copies of a landscape photograph that you have never seen before and know nothing about how it should appear. The person then rips up each photograph with their hands, so that different size pieces are present from each copy. The person then mixes all of the pieces together and asks you to reconstruct the photograph. In one of the smaller pieces you see a mountain. In a larger piece, you see that the same mountain is behind a lake. A third fragment shows only the lake, but it reveals that there is a cabin on the shore of the lake. Therefore, from looking at the overlapping information in these three fragments, you know that the picture contains a mountain behind a lake that has a cabin on its shore. This is the principle behind reconstructing entire DNA sequences using shotgun sequencing.

Originally, shotgun sequencing only analyzed one end of each fragment for overlaps. This was sufficient for sequencing small genomes. However, the desire to sequence larger genomes, such as that of a human, led to the development of double-barrel shotgun sequencing, more formally known as pairwise-end sequencing. In pairwise-end sequencing, both ends of each fragment are analyzed for overlap. Pairwise-end sequencing is,

therefore, more cumbersome than shotgun sequencing, but it is easier to reconstruct the sequence because there is more available information.

Next-generation Sequencing

Since 2005, automated sequencing techniques used by laboratories are under the umbrella of next-generation sequencing, which is a group of automated techniques used for rapid DNA sequencing. These automated low-cost sequencers can generate sequences of hundreds of thousands or millions of short fragments (25 to 500 base pairs) in the span of one day. These sequencers use sophisticated software to get through the cumbersome process of putting all the fragments in order.
EVOLUTION CONNECTION

Comparing SequencesA sequence alignment is an arrangement of proteins, DNA, or RNA; it is used to identify regions of similarity between cell types or species, which may indicate conservation of function or structures. Sequence alignments may be used to construct phylogenetic trees. The following website uses a software program called BLAST (basic local alignment search tool).

Under “Basic Blast,” click “Nucleotide Blast.” Input the following sequence into the large “query sequence” box: ATTGCTTCGATTGCA. Below the box, locate the “Species” field and type “human” or “Homo sapiens”. Then click “BLAST” to compare the inputted sequence against known sequences of the human genome. The result is that this sequence occurs in over a hundred places in the human genome. Scroll down below the graphic with the horizontal bars and you will see short description of each of the matching hits. Pick one of the hits near the top of the list and click on “Graphics”. This will bring you to a page that shows where the sequence is found within the entire human genome. You can move the slider that looks like a green flag back and forth to view the sequences immediately around the selected gene. You can then return to your selected sequence by clicking the “ATG” button.

Use of Whole-Genome Sequences of Model Organisms

The first genome to be completely sequenced was of a bacterial virus, the bacteriophage fx174 (5368 base pairs); this was accomplished by Fred Sanger using shotgun sequencing. Several other organelle and viral genomes were later sequenced. The first organism whose genome was sequenced was the bacterium *Haemophilus influenzae*; this was accomplished by Craig Venter in the 1980s. Approximately 74 different laboratories collaborated on the sequencing of the genome of the yeast*Saccharomyces cerevisiae*, which began in 1989 and was completed in 1996, because it was 60 times bigger than any other genome that had been sequenced. By 1997, the genome sequences of two important model organisms were available: the bacterium *Escherichia coli* K12 and the yeast *Saccharomyces cerevisiae*. Genomes of other model organisms, such as the mouse *Mus musculus*, the fruit fly *Drosophila melanogaster*, the nematode *Caenorhabditis. elegans*, and humans *Homo sapiens*are now known. A lot of basic research is performed in model organisms because the information can be applied to genetically similar organisms. A model organism is a species that is studied as a model to understand the biological processes in other species represented by the model organism. Having entire genomes sequenced helps with the research efforts in these model organisms. The process of attaching biological information to gene sequences is called genome annotation. Annotation of gene sequences helps with basic experiments in molecular biology, such as designing PCR primers and RNA targets.

Link to Learning

Click through each step of genome sequencing at this site.

Uses of Genome Sequences

DNA microarrays are methods used to detect gene expression by analyzing an array of DNA fragments that are fixed to a glass slide or a silicon chip to identify active genes and identify sequences. Almost one million genotypic abnormalities can be discovered using microarrays, whereas whole-genome sequencing can provide information about all six billion base pairs in the human genome. Although the study of medical applications of genome

sequencing is interesting, this discipline tends to dwell on abnormal gene function. Knowledge of the entire genome will allow future onset diseases and other genetic disorders to be discovered early, which will allow for more informed decisions to be made about lifestyle, medication, and having children. Genomics is still in its infancy, although someday it may become routine to use whole-genome sequencing to screen every newborn to detect genetic abnormalities.

In addition to disease and medicine, genomics can contribute to the development of novel enzymes that convert biomass to biofuel, which results in higher crop and fuel production, and lower cost to the consumer. This knowledge should allow better methods of control over the microbes that are used in the production of biofuels. Genomics could also improve the methods used to monitor the impact of pollutants on ecosystems and help clean up environmental contaminants. Genomics has allowed for the development of agrochemicals and pharmaceuticals that could benefit medical science and agriculture.

It sounds great to have all the knowledge we can get from whole-genome sequencing; however, humans have a responsibility to use this knowledge wisely. Otherwise, it could be easy to misuse the power of such knowledge, leading to discrimination based on a person's genetics, human genetic engineering, and other ethical concerns. This information could also lead to legal issues regarding health and privacy.

Section Summary

Whole-genome sequencing is the latest available resource to treat genetic diseases. Some doctors are using whole-genome sequencing to save lives. Genomics has many industrial applications including biofuel development, agriculture, pharmaceuticals, and pollution control. The basic principle of all modern-day sequencing strategies involves the chain termination method of sequencing.

Although the human genome sequences provide key insights to medical professionals, researchers use whole-genome sequences of model organisms to better understand the genome of the species. Automation and the decreased cost of whole-genome sequencing may lead to personalized medicine in the future.

Review Questions

The chain termination method of sequencing:

1. uses labeled ddNTPs
2. uses only dideoxynucleotides
3. uses only deoxynucleotides
4. uses labeled dNTPs

Whole-genome sequencing can be used for advances in:

1. the medical field
2. agriculture
3. biofuels
4. all of the above

Sequencing an individual person's genome

1. is currently possible
2. could lead to legal issues regarding discrimination and privacy
3. could help make informed choices about medical treatment
4. all of the above

What is the most challenging issue facing genome sequencing?

1. the inability to develop fast and accurate sequencing techniques
2. the ethics of using information from genomes at the individual level
3. the availability and stability of DNA
4. all of the above

Glossary

chain termination method method of DNA sequencing using labeled dideoxynucleotides to terminate DNA replication; it is also called the dideoxy method or the Sanger method

contig larger sequence of DNA assembled from overlapping shorter sequences

deoxynucleotide individual monomer (single unit) of DNA

dideoxynucleotide individual monomer of DNA that is missing a hydroxyl group (–OH)

DNA microarray method used to detect gene expression by analyzing an array of DNA fragments that are fixed to a glass slide or a silicon chip to identify active genes and identify sequences

genome annotation process of attaching biological information to gene sequences

model organism species that is studied and used as a model to understand the biological processes in other species represented by the model organism

next-generation sequencing group of automated techniques used for rapid DNA sequencing

shotgun sequencing method used to sequence multiple DNA fragments to generate the sequence of a large piece of DNA

whole-genome sequencing process that determines the DNA sequence of an entire genome

Licensing & Attributions

CC licensed content, Original

- Biology. **Authored by**: Open Stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biologyt. **License**: *CC BY: Attribution*

APPLYING GENOMICS

Objectives:

By the end of this section, you will be able to:

- Explain pharmacogenomics
- Define polygenic

The introduction of DNA sequencing and whole genome sequencing projects, particularly the Human Genome project, has expanded the applicability of DNA sequence information. Genomics is now being used in a wide variety of fields, such as metagenomics, pharmacogenomics, and mitochondrial genomics. The most commonly known application of genomics is to understand and find cures for diseases.

Predicting Disease Risk at the Individual Level

Predicting the risk of disease involves screening currently healthy individuals by genome analysis at the individual level. Intervention with lifestyle changes and drugs can be recommended before disease onset. However, this approach is most applicable when the problem resides within a single gene defect. Such defects only account for approximately 5 percent of diseases in developed countries. Most of the common diseases, such as heart disease, are multi-factored or polygenic, which is a phenotypic characteristic that involves two or more genes, and also involve environmental factors such as diet. In April 2010, scientists at Stanford University published the

genome analysis of a healthy individual (Stephen Quake, a scientist at Stanford University, who had his genome sequenced); the analysis predicted his propensity to acquire various diseases. A risk assessment was performed to analyze Quake's percentage of risk for 55 different medical conditions. A rare genetic mutation was found, which showed him to be at risk for sudden heart attack. He was also predicted to have a 23 percent risk of developing prostate cancer and a 1.4 percent risk of developing Alzheimer's. The scientists used databases and several publications to analyze the genomic data. Even though genomic sequencing is becoming more affordable and analytical tools are becoming more reliable, ethical issues surrounding genomic analysis at a population level remain to be addressed.

ART CONNECTION

PCA3

Step 1: *PCA3* mRNA anneals to complementary DNA primers that are attached to beads.

Step 2: The mRNA is amplified using reverse-transcriptase PCR.

Step 3: The mRNA is detected using a chemilumi-nescent probe.

PCA3 is a gene that is expressed in prostate epithelial cells and overexpressed in cancerous cells. A high concentration of *PCA3* in urine is indicative of prostate cancer. The *PCA3* test is considered to be a better indicator of cancer than the more well know PSA test, which measures the level of PSA (prostate-specific antigen) in the blood.

In 2011, the United States Preventative Services Task Force recommended against using the PSA test to screen healthy men for prostate cancer. Their recommendation is based on evidence that screening does not reduce the risk of death from prostate cancer. Prostate cancer often develops very slowly and does not cause problems, while the cancer treatment can have severe side effects. The *PCA3* test is considered to be more accurate, but screening may still result in men who would not have been harmed by the cancer itself suffering side effects from treatment. What do you think? Should all healthy men be screened for prostate cancer using the *PCA3* or PSA test? Should people in general be screened to find out if they have a genetic risk for cancer or other diseases?

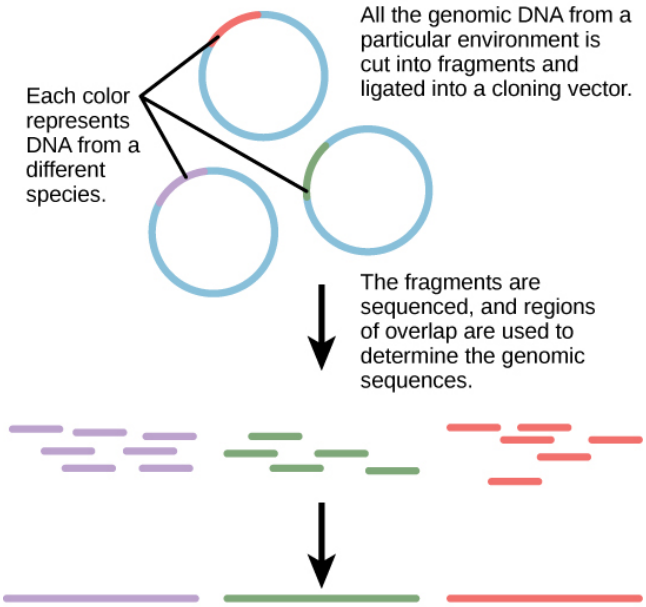
Pharmacogenomics and Toxicogenomics

Pharmacogenomics, also called toxicogenomics, involves evaluating the effectiveness and safety of drugs on the basis of information from an individual's genomic sequence. Genomic responses to drugs can be studied using experimental animals (such as laboratory rats or mice) or live cells in the laboratory before embarking on studies with humans. Studying changes in gene expression could provide information about the transcription profile in the presence of the drug, which can be used as an early indicator of the potential for toxic effects. For example, genes involved in cellular growth and controlled cell death, when disturbed, could lead to the growth of cancerous cells. Genome-wide studies can also help to find new genes involved in drug toxicity. Personal genome sequence information can be used to prescribe medications that will be most effective and least toxic on the basis of the individual patient's genotype. The gene signatures may not be completely accurate, but can be tested further before pathologic symptoms arise.

Microbial Genomics: Metagenomics

Traditionally, microbiology has been taught with the view that microorganisms are best studied under pure culture conditions, which involves isolating a single type of cell and culturing it in the laboratory. Because microorganisms can go through several generations in a matter of hours, their gene expression profiles adapt to the new laboratory environment very quickly. In addition, the vast majority of bacterial species resist being cultured in

isolation. Most microorganisms do not live as isolated entities, but in microbial communities known as biofilms. For all of these reasons, pure culture is not always the best way to study microorganisms. Metagenomics is the study of the collective genomes of multiple species that grow and interact in an environmental niche. Metagenomics can be used to identify new species more rapidly and to analyze the effect of pollutants on the environment (Figure).



Metagenomics involves isolating DNA from multiple species within an environmental niche.

Microbial Genomics: Creation of New Biofuels

Knowledge of the genomics of microorganisms is being used to find better ways to harness biofuels from algae and cyanobacteria. The primary sources of fuel today are coal, oil, wood, and other plant products, such as ethanol. Although plants are renewable resources, there is still a need to find more alternative renewable sources of energy to meet our population's energy demands. The microbial world is one of the largest resources for genes that encode new enzymes and produce new organic compounds, and it remains largely untapped. Microorganisms are used to create products, such as enzymes that are used in research, antibiotics, and other anti-microbial mechanisms. Microbial genomics is helping to develop diagnostic tools, improved vaccines, new disease treatments, and advanced environmental cleanup techniques.

Mitochondrial Genomics

Mitochondria are intracellular organelles that contain their own DNA. Mitochondrial DNA mutates at a rapid rate and is often used to study evolutionary relationships. Another feature that makes studying the mitochondrial genome interesting is that the mitochondrial DNA in most multicellular organisms is passed on from the mother during the process of fertilization. For this reason, mitochondrial genomics is often used to trace genealogy.

Information and clues obtained from DNA samples found at crime scenes have been used as evidence in court cases, and genetic markers have been used in forensic analysis. Genomic analysis has also become useful in this field. In 2001, the first use of genomics in forensics was published. It was a collaborative attempt between academic research institutions and the FBI to solve the mysterious cases of anthrax communicated via the US Postal Service. Using microbial genomics, researchers determined that a specific strain of anthrax was used in all the mailings.

Genomics in Agriculture

Genomics can reduce the trials and failures involved in scientific research to a certain extent, which could improve the quality and quantity of crop yields in agriculture. Linking traits to genes or gene signatures helps to improve crop breeding to generate hybrids with the most desirable qualities. Scientists use genomic data to identify desirable traits, and then transfer those traits to a different organism. Scientists are discovering how genomics can improve the quality and quantity of agricultural production. For example, scientists could use desirable traits to create a useful product or enhance an existing product, such as making a drought-sensitive crop more tolerant of the dry season.

Section Summary

Imagination is the only barrier to the applicability of genomics. Genomics is being applied to most fields of biology; it is being used for personalized medicine, prediction of disease risks at an individual level, the study of drug interactions before the conduct of clinical trials, and the study of microorganisms in the environment as opposed to the laboratory. It is also being applied to developments such as the generation of new biofuels, genealogical assessment using mitochondria, advances in forensic science, and improvements in agriculture.

Art Connections

Figure In 2011, the United States Preventative Services Task Force recommended against using the PSA test to screen healthy men for prostate cancer. Their recommendation is based on evidence that screening does not reduce the risk of death from prostate cancer. Prostate cancer often develops very slowly and does not cause problems, while the cancer treatment can have severe side effects. The *PCA3* test is considered to be more accurate, but screening may still result in men who would not have been harmed by the cancer itself suffering side effects from treatment. What do you think? Should all healthy men be screened for prostate cancer using the *PCA3* or PSA test? Should people in general be screened to find out if they have a genetic risk for cancer or other diseases?

Review Questions

Genomics can be used in agriculture to:

1. generate new hybrid strains
2. improve disease resistance
3. improve yield
4. all of the above

Genomics can be used on a personal level to:

1. decrease transplant rejection
2. Predict genetic diseases that a person may have inherited
3. Determine the risks of genetic diseases for an individual's children
4. All the above

Free Response

Explain why metagenomics is probably the most revolutionary application of genomics.

How can genomics be used to predict disease risk and treatment options?

Glossary

metagenomics study of the collective genomes of multiple species that grow and interact in an environmental niche

pharmacogenomics study of drug interactions with the genome or proteome; also called toxicogenomics

polygenic phenotypic characteristic caused by two or more genes

pure culture growth of a single type of cell in the laboratory

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: Open Stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biologytt. **License**: *CC BY: Attribution*

GENOMICS AND PROTEOMICS

Objectives:

By the end of this section, you will be able to:

- Explain systems biology
- Describe a proteome
- Define protein signature

Proteins are the final products of genes, which help perform the function encoded by the gene. Proteins are composed of amino acids and play important roles in the cell. All enzymes (except ribozymes) are proteins that act as catalysts to affect the rate of reactions. Proteins are also regulatory molecules, and some are hormones. Transport proteins, such as hemoglobin, help transport oxygen to various organs. Antibodies that defend against foreign particles are also proteins. In the diseased state, protein function can be impaired because of changes at the genetic level or because of direct impact on a specific protein.

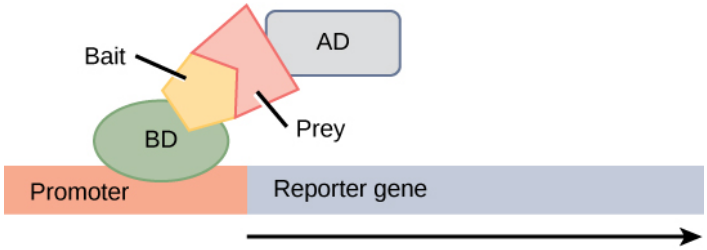
A proteome is the entire set of proteins produced by a cell type. Proteomes can be studied using the knowledge of genomes because genes code for mRNAs, and the mRNAs encode proteins. Although mRNA analysis is a step in the right direction, not all mRNAs are translated into proteins. The study of the function of proteomes is called proteomics. Proteomics complements genomics and is useful when scientists want to test their hypotheses that were based on genes. Even though all cells of a multicellular organism have the same set of genes, the set of proteins produced in different tissues is different and dependent on gene expression. Thus, the genome is constant, but the proteome varies and is dynamic within an organism. In addition, RNAs can be alternately spliced (cut and pasted to create novel combinations and novel proteins) and many proteins are modified after translation by processes such as proteolytic cleavage, phosphorylation, glycosylation, and ubiquitination. There are also protein-protein interactions, which complicate the study of proteomes. Although the genome provides a blueprint, the final architecture depends on several factors that can change the progression of events that generate the proteome.

Metabolomics is related to genomics and proteomics. Metabolomics involves the study of small molecule metabolites found in an organism. The metabolome is the complete set of metabolites that are related to the genetic makeup of an organism. Metabolomics offers an opportunity to compare genetic makeup and physical characteristics, as well as genetic makeup and environmental factors. The goal of metabolome research is to identify, quantify, and catalogue all of the metabolites that are found in the tissues and fluids of living organisms.

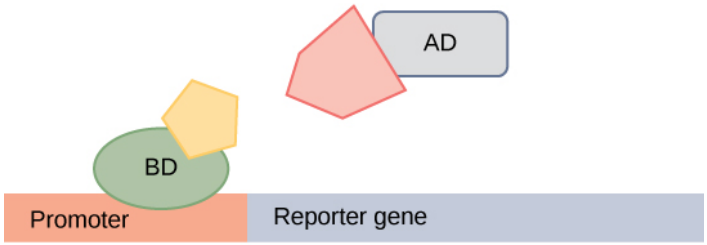
Basic Techniques in Protein Analysis

The ultimate goal of proteomics is to identify or compare the proteins expressed from a given genome under specific conditions, study the interactions between the proteins, and use the information to predict cell behavior or develop drug targets. Just as the genome is analyzed using the basic technique of DNA sequencing, proteomics requires techniques for protein analysis. The basic technique for protein analysis, analogous to DNA sequencing, is mass spectrometry. Mass spectrometry is used to identify and determine the characteristics of a molecule. Advances in spectrometry have allowed researchers to analyze very small samples of protein. X-ray crystallography, for example, enables scientists to determine the three-dimensional structure of a protein crystal at atomic resolution. Another protein imaging technique, nuclear magnetic resonance (NMR), uses the magnetic properties of atoms to determine the three-dimensional structure of proteins in aqueous solution. Protein microarrays have also been used to study interactions between proteins. Large-scale adaptations of the basic two-hybrid screen (Figure) have provided the basis for protein microarrays. Computer software is used to analyze the vast amount of data generated for proteomic analysis.

Genomic- and proteomic-scale analyses are part of systems biology. Systems biology is the study of whole biological systems (genomes and proteomes) based on interactions within the system. The European Bioinformatics Institute and the Human Proteome Organization (HUPO) are developing and establishing effective tools to sort through the enormous pile of systems biology data. Because proteins are the direct products of genes and reflect activity at the genomic level, it is natural to use proteomes to compare the protein profiles of different cells to identify proteins and genes involved in disease processes. Most pharmaceutical drug trials target proteins. Information obtained from proteomics is being used to identify novel drugs and understand their mechanisms of action.



If the bait protein interacts with the prey protein, the transcription factor's activator domain binds to the binding domain, and transcription occurs.



If the prey doesn't catch the bait, no transcription occurs.

Two-hybrid screening is used to determine whether two proteins interact. In this method, a transcription factor is split into a DNA-binding domain (BD) and an activator domain (AD). The binding domain is able to bind the promoter in the absence of the activator domain, but it does not turn on transcription. A protein called the bait is attached to the BD, and a protein called the prey is attached to the AD. Transcription occurs only if the prey "catches" the bait.

The challenge of techniques used for proteomic analyses is the difficulty in detecting small quantities of proteins. Although mass spectrometry is good for detecting small amounts of proteins, variations in protein expression in diseased states can be difficult to discern. Proteins are naturally unstable molecules, which makes proteomic analysis much more difficult than genomic analysis.

Cancer Proteomics

Genomes and proteomes of patients suffering from specific diseases are being studied to understand the genetic basis of the disease. The most prominent disease being studied with proteomic approaches is cancer. Proteomic approaches are being used to improve screening and early detection of cancer; this is achieved by identifying proteins whose expression is affected by the disease process. An individual protein is called a biomarker, whereas a set of proteins with altered expression levels is called aprotein signature. For a biomarker or protein signature to be useful as a candidate for early screening and detection of a cancer, it must be secreted in body fluids, such as sweat, blood, or urine, such that large-scale screenings can be performed in a non-invasive fashion. The current problem with using biomarkers for the early detection of cancer is the high rate of false-negative results. A false negative is an incorrect test result that should have been positive. In other words, many cases of cancer go undetected, which makes biomarkers unreliable. Some examples of protein biomarkers used in cancer detection are CA-125 for ovarian cancer and PSA for prostate cancer. Protein signatures may be more reliable than biomarkers to detect cancer cells. Proteomics is also being used to develop individualized treatment plans, which involves the prediction of whether or not an individual will respond to specific drugs and the side effects that the individual may experience. Proteomics is also being used to predict the possibility of disease recurrence.

The National Cancer Institute has developed programs to improve the detection and treatment of cancer. The Clinical Proteomic Technologies for Cancer and the Early Detection Research Network are efforts to identify protein signatures specific to different types of cancers. The Biomedical Proteomics Program is designed to identify protein signatures and design effective therapies for cancer patients.

Section Summary

Proteomics is the study of the entire set of proteins expressed by a given type of cell under certain environmental conditions. In a multicellular organism, different cell types will have different proteomes, and these will vary with changes in the environment. Unlike a genome, a proteome is dynamic and in constant flux, which makes it both more complicated and more useful than the knowledge of genomes alone.

Proteomics approaches rely on protein analysis; these techniques are constantly being upgraded. Proteomics has been used to study different types of cancer. Different biomarkers and protein signatures are being used to analyze each type of cancer. The future goal is to have a personalized treatment plan for each individual.

Review Questions

What is a biomarker?

1. the color coding of different genes
2. a protein that is uniquely produced in a diseased state
3. a molecule in the genome or proteome
4. a marker that is genetically inherited

A protein signature is:

1. the path followed by a protein after it is synthesized in the nucleus
2. the path followed by a protein in the cytoplasm
3. a protein expressed on the cell surface
4. a unique set of proteins present in a diseased state

Free Response

How has proteomics been used in cancer detection and treatment?

What is personalized medicine?

Glossary

biomarker individual protein that is uniquely produced in a diseased state

false negative incorrect test result that should have been positive

metabolome complete set of metabolites which are related to the genetic makeup of an organism

metabolomics study of small molecule metabolites found in an organism

protein signature set of uniquely expressed proteins in the diseased state

proteome entire set of proteins produced by a cell type

proteomics study of the function of proteomes

systems biology study of whole biological systems (genomes and proteomes) based on interactions within the system

Licensing & Attributions

CC licensed content, Shared previously

- Biology. **Authored by**: Open Stax. **Located at**: http://cnx.org/contents/GFy_h8cu@10.59:fVAf83sY@12/Preface-to-Biologytt. **License**: *CC BY: Attribution*