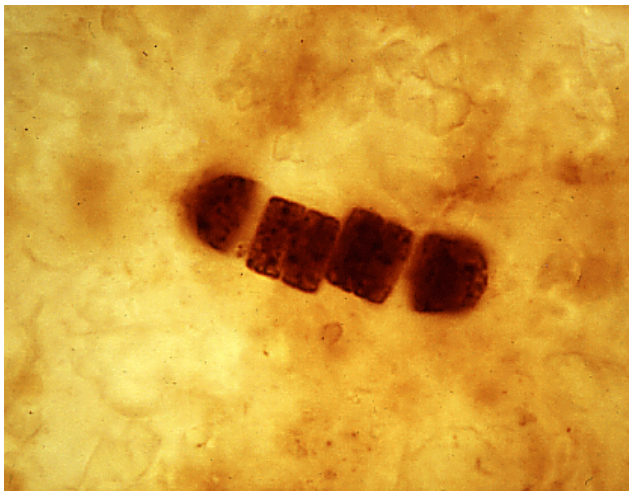
Big Idea II: Evolution



1 billion year old microbe fossil

2.7 billion year old colonial microbe fossil

For Big Idea II, Evolution, we will consider evolution at five biological levels from molecules through ecological systems. You will examine data that has shaped our current understanding so you can construct your own understanding of the origins of life and how it continues to change over time. As you read Chapters 6 – 10, keep in mind these four recurring themes:

- The origin of living systems occurred by natural processes, and life continues to evolve within a changing environment.
- Organisms can be linked by lines of descent from common ancestry.
- Natural selection is a mechanism of evolution that accounts for adaptation.
- Human activity can alter the course of evolution.

Chapter 6 Evolution at the Molecular Scale
Learning Objectives

Learning Objectives

1. Explain how the 5 tenets of natural selection influenced origin of life
2. Define the three fundamental properties of living systems.
3. Describe how RNA molecules can function as enzymes.
4. Discuss how vesicles can grow, compete and store energy.
5. Illustrate how abiotic structures exhibit dynamic and competitive behaviors.
6. Determine how eukaryotes inherited genes from bacteria and archaea.
7. Review the evolutionary origin of the nucleus.

Bio-Math Exploration Learning Objectives

1. Calculate the expected number of events given certain parameters.
2. Determine the probability of two things happening by chance.
3. Measure the rate of change of a biological event.
4. Convert change in pH to linear scale.

Ethical, Legal and Social Implications Learning Objectives (ELSI-LO)

1. Distinguish religion and science as two different ways of understanding the world.
2. Evaluate the difference between "belief" and "acceptance" of evolution.
3. Define the scientific term theory.

Chapter 6 Outline

Introduction
6.1 What is evolution?
Ethical, Legal, and Social Implications: Are Evolution and Religion Compatible?
6.2 Could abiotic molecules produce the first cells?
Bio-Math Exploration 6.1 How many mutations do you expect?
Bio-Math Exploration 6.2 Are you sure this is the best possible sequence?
Bio-Math Exploration 6.3 What is the probability of a highly conserved basepairing?
6.3 Can non-living objects compete and grow?
Bio-Math Exploration 6.4 How fast is the vesicle size changing?
6.4 Can non-living objects harvest and store energy?
Bio-Math Exploration 6.5 Logarithms: The power of pH
6.5 How did the first nucleus come into being?
Conclusions

You Are Here		Organizing, Big Ideas of Biology				
		Information	Evolution	Cells	Emergent Properties	Homeostasis
Levels of the Biological Hierarchy	Molecules	1	Chapter 6	11	16	21
	Cells	2	7	12	17	22
	Organisms	3	8	13	18	23
	Populations	4	9	14	19	24
	Ecological Systems	5	10	15	20	25

A vexing problem in biology is how life first began. Today, all organisms come from preexisting organisms so it is difficult to imagine how the first living cells came into existence. How could **abiotic** (non-living) molecules coalesce to form a living (**biotic**) cell? Just because it is hard to imagine, however, does not mean this problem is beyond scientific investigation. A growing number of scientists (biologists, chemists, biochemists, and biophysicists) have designed very clever experiments to improve our understanding about the origin of life. Chapter 6 focuses on the molecular aspects of evolution with special attention to the formation of complex living cells from simpler abiotic components. After clearly defining evolution, you will examine data that reveals how non-living chemicals can exhibit traits resembling simple cells. Later, you will analyze data that illustrates the origin of eukaryotic cells from prokaryotic ancestors. {Definitions: **Abiotic** refers to chemicals and material that are not living. **Biotic** refers to living cells and higher order structures. }

6.1 What is Evolution?

- Context: Evolution is often misunderstood and discounted based on faith rather than facts.
- Major Themes: Natural selection is a mechanism of evolution that accounts for adaptation.
- Bottom Line: Evolution explains the biological world using scientific principles.

Evolution describes how populations of organisms adapt to a changing world. You will often hear evolution referred to as a **theory**. The term "theory" has two different meanings. The more casual use of theory is defined as a guess, similar to a hypothesis; "I have a theory to explain why we have not heard from Martians yet." The scientific use of theory is defined as a widely accepted concept that has been demonstrated many, many times. For example, the theory of evolution has been supported in at least 100,000 scholarly papers. Because of the dual use of the word theory, more scientists now refer to the *principle* of evolution to help non-scientists understand the degree of confidence in evolution as the only natural explanation for the diversity of life. {Definition: **Evolution** is the scientific explanation for the origin of life and its continual change over time. **Theory** in a science context means a widely accepted concept that has been demonstrated many, many times.}

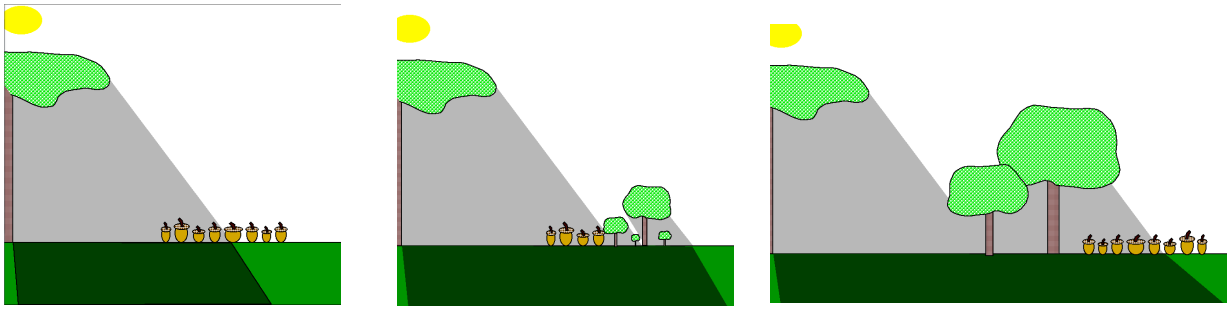


Figure 6.1 Example of natural selection. a) Over production of acorns with variation leads to a competition for the limited resource of light. b) F_1 Acorns that are faster to germinate grow first and shade out their competition. c) Those with the selective advantage survive, reproduce and propagate the trait of fast germination in F_3 progeny

Charles Darwin is usually credited with the discovery of the primary mechanism of evolution, **natural selection**, but the concept of biological change over time had been around for a while. Darwin's famous 1859 publication *On the Origin of Species* spelled out the essential elements of evolution by natural selection as we understand it today. {Definition: **Natural selection** is one mechanism by which evolution takes place and is often summarized by survival of the fittest. } To understand natural selection, you need to appreciate five basic tenets that you can observe around you (Figure 6.1):

1) **Overproduction**. Each generation of an organism produces more offspring than nature can support.
2) **Variation**. With each new generation, individuals have slight differences in characteristics, which means they have slightly different abilities.
3) **Competition**. Overproduction of offspring results in competition for a limited number of resources such as water, food, and shelter.
4) **Selective advantage**. Variation results in some individuals who have an advantage over others, depending on the circumstances. Variations in strength, acquiring energy, stress resistance, or in some other way allows these individuals to out compete others for the limited resources which means they continue living.
5) **Reproduction.** Those who survive the competition are able to reproduce and pass on to the next generation the genetic characteristics that enabled them to out-compete others, though the next generation will exhibit variations of the successful trait.

The five basic tenets of natural selection are all around you – literally. Let's use grade point average (GPA) as a metaphor to represent success. There are many people in your class – way more than your teacher expects will earn a letter grade of A (*overproduction*). Your class is filled with students of varying degrees of intelligence and work ethic (*variation*). Students who exhibit the most of these two qualities earn the top GPA (*competition*). Some students are super smart but uninterested, others work hard but are not as smart. The best grades are earned by students who have the right combination of work ethic and intelligence (*selective advantage*). At this point, the analogy of grades and natural selection begins to break down. Students who earn the top grades can expect to be promoted to the next level and maybe even earn scholarships so they can become teachers. Becoming a teacher is not exactly analogous to *reproduction*, but transmitting information and study habits to the next generation is sufficient for our purposes.

Evolution helps us understand life as it exists today and continues to change over time. We can look back into history to see earlier steps of evolution, but don't be fooled into thinking that evolution has an end goal to accomplish. Evolution is shaped by random changes, some of which are beneficial. It is a common misconception that humans are the pinnacle of evolution, but let's consider a different world. Imagine that 200 years from now, Earth has no oxygen and only CO_2 and nitrogen remain in our atmosphere. Under these oxygen-deficient conditions, humans and all other animals could go extinct. Microbes that can live in the absence of oxygen

would be rulers of the planet. Because environments change all the time, a selective advantage today may not be beneficial tomorrow and so you can never consider adaptive evolution as a series of steps towards an endpoint or goal. In fact, you could think of natural selection as ongoing optimization for changing conditions. Adaptive evolution proceeds because of the process outlined in the five basic tenets above.

In 1973, geneticist Theodosius Dobzhansky famously stated, "Nothing in biology makes sense except in the light of evolution." Dobzhansky meant that the existence of viruses, the size of redwood trees, and the smell of a skunk seem illogical until you see them through the adaptive lens of natural selection. Though what we see in the world today does make sense with respect to natural selection, it is hard to imagine what Earth looked like before any life had evolved. What experimental evidence is available to help us comprehend the origin of life that took place approximately 3.5 billion years ago? Of course, we cannot conduct experiments in the traditional sense; it is impossible to find a second planet Earth where we can observe evolution's beginnings, or manipulate this second Earth to produce new life forms. However, the lack of a spare Earth as an experimental model has not prevented some very clever investigators from conducting research on how life may have evolved.

Since biology is the study of life, it seems logical that biologists should begin by defining the word "life." Though life is a difficult term to define, it satisfies three fundamental properties: life replicates itself; life is contained in a three-dimensional space; and life undergoes changes. However, life is an **emergent property** that occurs when certain molecules come together in very specific ways. {Connection: Emergent Properties are the fourth Big Idea in this book.} Emergent properties occur only when constituent parts are assembled appropriately. You probably use a cell phone with many amazing properties, but its isolated parts of individual electrons, plastic and silicon wafers are not functional and lack valuable emergent properties. Like a cell phone, life is more than the sum of its individual parts. {Definition: **Emergent property** is a characteristic that becomes apparent at one level of biological complexity due to interactions among lower level components.}

Many science fiction movies have used the three fundamental properties of life to explore what might happen if we built robots that could produce more robots. Science fiction stories can lead to interesting philosophical questions, but for now, focus your attention on a process that sounds equally improbable but more amazing than any movie – the origin of life on Earth.

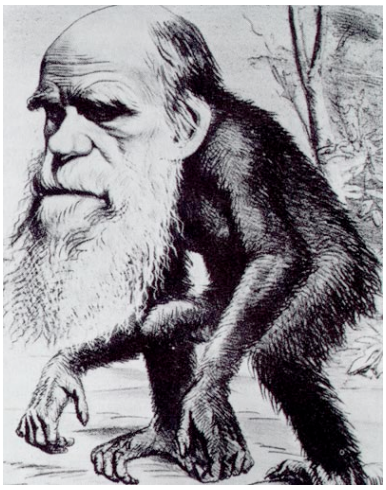
Ethical, Legal, Social Implications: Box 6.1

Believing *v.* Accepting

On October 18, 2004, the school board in Dover, Pennsylvania, voted 6 to 3 in favor of teaching intelligent design (ID) as an alternative to evolution. By December 14 of the same year, eleven parents whose children were affected by the new education policy filed a lawsuit (Kitzmiller *v.* Dover Area School District) to prevent the teaching of religion as science. Five days before Christmas in 2005, Judge John E. Jones III ruled in favor of the parents and against

the Dover School Board. In his ruling, Jones wrote a 139 page opinion explaining why ID cannot be used for science education. He wrote, "The overwhelming evidence at trial established that ID is a religious view, a mere re-labeling of creationism, and not a scientific theory." Jones continued by stating, "ID violates the centuries-old ground rules of science by invoking and permitting supernatural causation…"

When Darwin published his explanation of evolution, he experienced a very negative response from society because he challenged a widely held religious belief of the late 1800's – life on Earth had not changed since divine creation. Darwin's science-based proposal of evolution explained the natural world based on data. A popular political cartoon showed Darwin's head on a monkey because people misinterpreted "descent with modification" to mean we were direct descendants of monkeys (ELSI Figure 6.1). The essential problem was people felt they were being forced to choose either religion *OR* science. However, this choice was a false dichotomy. People can choose to believe in the God of their religion *AND* accept the validity of evolution. Religion and science are two different ways of trying to understand the world but they cannot explain the same concepts.



ELSI Figure 6.1 Nineteenth century political cartoon of Charles Darwin. The concept of descent from a common ancestor was not fully understood by Victorian England so Darwin was ridiculed as being a buffoon.

Any time there is a misunderstanding, it usually begins with a breakdown in communication. To help clarify, let's define some key terms. According to Webster's online dictionary, *religion* is "1) the service and worship of God or the supernatural; 2) a cause, principle, or system of beliefs held to with ardor and faith." *Faith* is "complete trust" or a "firm belief in something for which there is no proof." Notice that religion and faith do not require evidence. Instead, they are based on explanations outside of nature for ideas that cannot be experimentally demonstrated. We cannot conduct experiments to demonstrate the existence of God, or lack of God. People of faith have complete trust in something that is beyond earthly matters.

Faith and religion stand in stark contrast with *science* which is, "knowledge or a system of knowledge… concerned with the physical world and its phenomena… covering general truths or the operation of general laws especially as obtained and tested through scientific method." The scientific method is based on personal observation, experimentation, hypothesis testing, collecting and analyzing data, and reproducibility of observations. Notice that terms such as "belief" or "faith" are not used to describe science. That's why it is inappropriate to ask someone, "Do you *believe* in evolution?" Evolution is the outcome of the scientific method and therefore is a matter of science, not a matter of faith or part of religion. Belief is the wrong term to describe evolution. Instead, you can ask someone if they accept evolution based on the available data. Scientists can reject scientific conclusions and offer alternative explanations, based on different interpretations of the same data. Competing scientific ideas must be founded

on data, not faith, and they must be testable without having to invoke the supernatural. It is unscientific to say I do not *believe* in your scientific conclusion. To say you do not *believe* in evolution is to say you refuse to consider the natural world in a scientific manner. However, if you do not *agree* with the principle of evolution, then we can discuss the evidence based on facts and not faith.

Many famous scientists are deeply religious. They have no problem accepting evolution for the earthly explanations and believing in God to explain their spiritual world and moral values. Dr. Francis Collins (director of the National Institutes of Health in Washington, DC) wrote a book to explain how religion and science are not mutually exclusive for him and other scientists, though the two fields are mutually exclusive in how they understand the world. He wrote, "Science is not threatened by God; it is enhanced," and "God is most certainly not threatened by science." In an interview, Collins elaborated by saying, "Don't misunderstand me, it is clear that the process of evolution by natural selection over hundreds of millions of years is the 'how' that explains the marvelous diversity of life. But that doesn't provide the answer to 'why.' I think God provides that answer."

Parents in the Dover Area School District asked the U.S. government in the form of the U.S. District Court for the Middle District of Pennsylvania, to determine if ID was religion or science. Judge Jones wrote, "We conclude that the religious nature of ID would be readily apparent to an objective observer, adult or child." Jones understood the significance of his opinion and continued, "ID's backers have sought to avoid the scientific scrutiny which we have now determined that it cannot withstand by advocating that the *controversy*, but not ID itself, should be taught in science class.…The goal of ID is not to encourage critical thought, but to foment a revolution which would supplant evolutionary theory with ID." It is worth noting that Judge Jones is conservative, Christian, and a George W. Bush appointee who was not trying to promote a liberal agenda – he was basing his ruling on fact and a clear definition of science. The Judge did not rule on whether ID was correct or not since its validity was not on trial. What was being judged was whether the religious belief in ID could be taught in a science class at an equivalent level with the scientific principle of evolution. Jones emphatically ruled that science and religion are distinct, and faith cannot be employed to formulate scientific conclusions. Therefore, as you read Chapter 6, which addresses the natural world and how life may have evolved, do not consider the data to be a challenge to your personal beliefs or religion because science, by definition, cannot disprove your faith. Religion will always be distinct from science, and *vice versa*. The U.S. government has determined science and religion are separate and should not be confused or substituted for each other.

ELSI Integrating Question

1. Some people *believe* in fortune tellers and others do not. Rephrase this statement using scientific terms. Devise science-based experiments to determine whether fortune tellers are actually seeing the future or not.

6.2 Could abiotic molecules produce the first cells?

- Context: It would be helpful to know whether biological molecules could have been formed prior to the evolution of cells and whether they have any enzymatic activity.
- Major Themes: The origin of living systems occurred by natural processes, and life continues to evolve within a changing environment.
- Bottom Line: Amino acids can form abiotically and RNA molecules can function as enzymes to transmit genetic information.

When you examine life closely, you will discover that it is made from four essential raw materials (Figure 6.2). **Proteins** provide shape to cells and perform functions. **Nucleic acids** such as DNA and RNA carry information from one place to the next and across generations. **Carbohydrates**, in the form of sugars, provide support and relay energy from one place to another. And **lipids** assemble into double layered membranes around cells. From these four basic classes of biological molecules, under the right circumstances and timing, the emergent property of life sprang from our abiotic planet. {Definitions: **Proteins** have shapes that determine their function and are formed by the assembly amino acids. **Nucleic acids** are the building blocks of RNA and DNA and are formed by the assembly of sugars, bases, and phosphate. **Carbohydrates** are energy-rich molecules typically composed of sugars which include carbon, hydrogen and oxygen. **Lipids** are fats and oils that do not dissolve in water and are a critical component of cell membranes.}

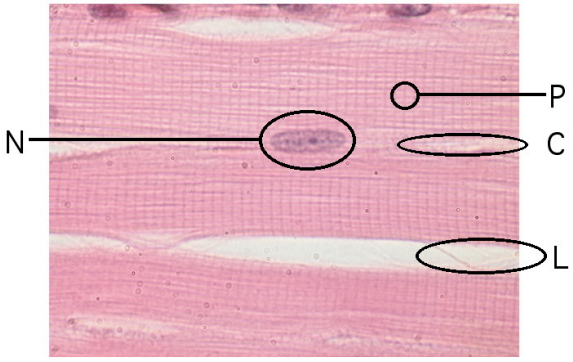
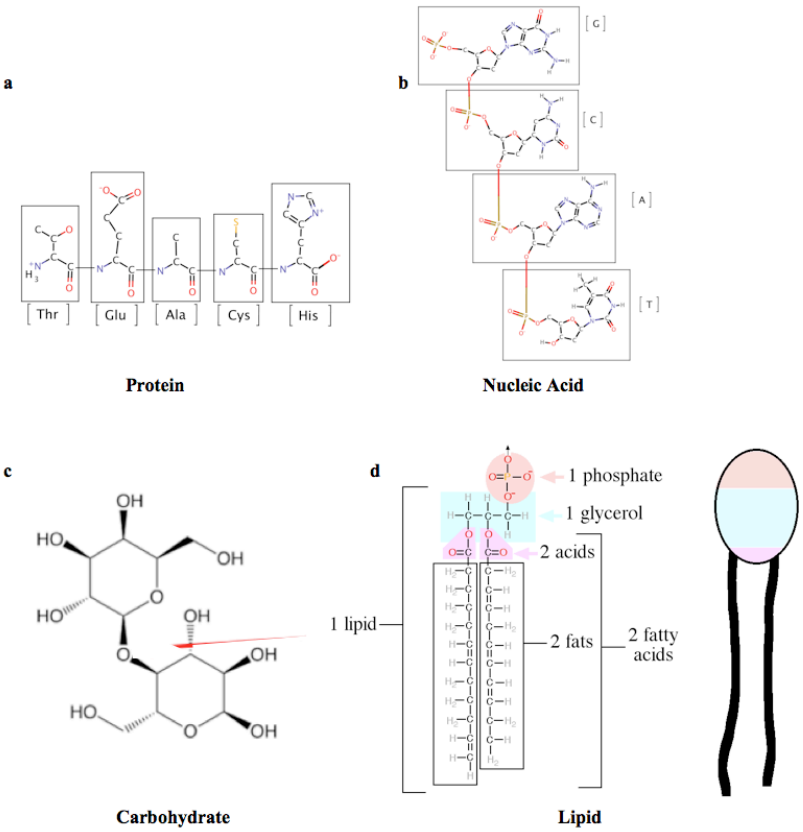


Figure 6.2 Building blocks of a muscle cell. Portions of several skeletal muscle cells are shown with proteins (P) forming the actin and myosin fibers, nucleic acids (N) inside the nucleus, carbohydrates (C) in the cytoplasm, and lipids (L) forming the plasma membranes.

You have learned about nucleic acids and proteins in Chapter 1 (Figure 6.3). In Chapter 2, you learned about carbohydrates, such as the sugar lactose, as sources of energy. The only new category of molecule is lipids. Lipids are often drawn as icons of a circle and two long tails. The circle represents the hydrophilic "head group" that contains the **acid** and the 3-carbon glycerol. Phospholipids also contain a phosphate that may or may not have another molecule attached where the small arrow is drawn in Figure 6.4d. The hydrophilic portion of a lipid interacts with the cytoplasm inside cells and the watery outside world of a cell. You can tell the head group is hydrophilic because the phosphate has negative charges and the acids contain oxygen which is highly **electronegative**, or acts like an electron hog. The elements `N O P S` tend to carry partial negative charges because when they covalently bind to other atoms, `N O P` and `S` tend to hold the electrons a little closer to themselves. When you see one or more of these elements, you can

guess that portion of the molecule will be hydrophilic. The two **hydrocarbon chains** are composed completely of C and H neither of which hog their covalently shared electrons and thus they are **non-polar**, meaning neither element has a partial charge. Water is **polar** so it does not interact well with non-polar molecules such as fats. The fats are attached to the acids to make two fatty acids which are hydrophobic. When lipids are mixed together in the presence of water, the fatty acids stick together in the middle with the hydrophobic head groups facing outwards to the water. Lipids form what looks like an Oreo cookie with hydrophilic cookie outside and hydrophobic fat inside. These lipid sheets can be planar, or more often, they circularize into spheres with an aqueous lumen inside. In addition to being a main components of membranes, fats contain a lot of energy tied up in the covalent bonds in the long hydrocarbon molecules.



{Definitions: **Acids** often contain a characteristic $O^--\overset{O}{\overset{\|}{C}}\rightarrow$ group that donates a H^+ to the solvent and carry a negative charge. **Electronegative** elements `N O P S` pull electrons closer to themselves in covalent bonds. **Hydrocarbon chains** are composed of covalently linked hydrogens and carbons. **Polar** molecules have polar covalent bonds that include N O P or S. **Non-polar** molecules typically have C-H bonds and exclude N O P and S.}

Figure 6.3 Complex biological molecules that form the building blocks of life. a) Five amino acids connected by peptide bonds. b) Four nucleotides connected by phosphodiester bonds. c) The disaccharide lactose. d) A generic phospholipid drawn in chemical detail and as simplified icon.

It is difficult to picture the universe as an expanding collection of elements that coalesced into balls of fire and clay to form suns and planets. Imagine primitive Earth spinning on its axis and orbiting our sun, then contemplate how proteins, nucleic acids, carbohydrates and lipids came into being before life evolved. These four building blocks are made by cells every second of every day now, but before life began, how did these complex biological molecules appear on Earth and coalesce into a living cell?

The universe is composed of the elements you see on a chemistry periodic table. These atoms do not self-assemble into complex biological molecules when you mix them in a test tube containing carbon, nitrogen, oxygen, and hydrogen. A fundamental principle of biology, often

referred to as the **cell theory**, states that all cells come from pre-existing cells. {Connection: Cells are the third Big Idea of Biology covered in this book.} If the cell theory is true, then what produced the first cell? This troubling question has been pondered by philosophers and scientists for hundreds of years. In this Section, you will examine some data that will help you understand how life may have assembled itself from simple elements, to complex molecules, and eventually the four building blocks of life. {Definition: **Cell theory** is universally accepted and states that all cells come from pre-existing cells.}

Integrating Questions

1. Use Figure 6.3 to determine which molecules are hydrophilic or hydrophobic. Use the `NOPS` elements to help you determine the answers.
2. Search the PubChem database to find images of these molecules: the fatty acid called myristoleate; the two amino acids glycine and aspartic acid. Knowing these structures will help you understand the critical experiments in Chapter 6.

Proteins, carbohydrates, nucleic acids and lipids are **organic molecules**, meaning they contain carbon. Scientists looked at these biological molecules and tried to determine how such complex organic molecules could have been produced abiotically. Thinking about the origin of organic molecules is similar to asking which came first, the chicken or the egg. It seems impossible to have life without these complex molecules and yet it seems impossible for these complex molecules to be formed abiotically. But some very clever and persistent investigators have performed experiments to determine if complex organic molecules can be formed in an abiotic world. {Definition: **Organic molecules** contain carbon atoms.}

In 1953, a graduate student at the University of Chicago named Stanley Miller tried to replicate primitive Earth to facilitate abiotic production of amino acids, the basic subunits of proteins. Because no scientist can wait a million years to let things happen naturally, Miller wanted to speed up the process but retain a constructed abiotic world (Figure 6.4a). To conduct his research, he needed a small, self-contained primitive Earth which he built out of Pyrex glass. Geologists had already described which simple, non-organic molecules were present 3 billion years ago before life evolved. Water and all ingredients were poured into Miller's side-arm access point which he later sealed shut. The "modern air" was removed by connecting the apparatus to a vacuum pump and then he filled the entire device with 1 part hydrogen gas H_2, 2 parts methane gas CH_4, and 2 parts ammonia gas NH_3. Heat was applied to the boiler flask; steam rose and moved to a reaction chamber where electrical sparks were applied. Steam condensed in a water-cooled chamber to form liquid which returned to the boiling flask to repeat the cycle. The entire mixture moved in a clockwise direction, and sparks of electricity flashed like prehistoric lightning storms in the reaction chamber.

The electrical stimulation continued for an entire week, nonstop. During the first day, the water became pink and by the end of the week, the water was deep red and bits of glass that wore off during the week to make the solution cloudy. The access point side-arm was cut one week

later to remove the liquid for analysis. The red material was dried and analyzed by two-dimensional thin layer chromatography (see Figure 1.20), a process of chemically separating molecules based on their chemical properties. Miller used a mixture of butyl alcohol and acetic acid for his first solvent and molecules migrated from top to bottom in Figure 6.4b. After allowing the week old reaction products to separate for a while, he dried the paper, rotated it 90 degrees, and then dipped the left edge in the organic solvent phenol to separate the molecules left to right in Figure 6.4b. Once this two-step chromatography was completed, he sprayed the entire paper with a stain to visualize what had been separated from the red solution. As you can see in his hand-written annotation, he discovered that at least 3 different amino acids were formed abiotically under conditions similar to ancient Earth. The identities of the colored spots were later verified by Miller and others using a variety of methods. Essentially, Miller had determined that amino acids, the subunits of proteins, could be synthesized in the absence of life.

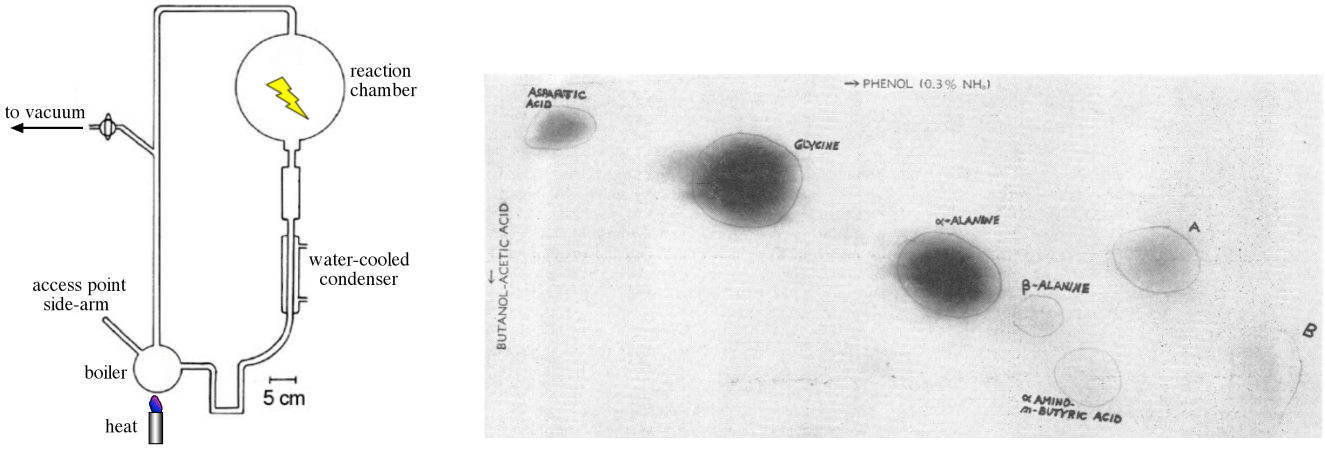


Figure 6.4 Miller's experiment to simulate primitive Earth. Modified drawing from Miller showing the parts of his ancient world device. b) Thin layer chromatography results after running the primitive earth experiment for 1 week. Hand-written labels and circles by Miller.

Miller stimulated new ways of thinking about the origin of life in an abiotic world. Since the 1950s, people have found abiotic sources of lipids and the subunits of nucleic acids. Surprisingly, scientists discovered lipids produced abiotically in space when they examined the Murchison meteorite that landed in Australia in 1969. In 2001, NASA chemists replicated the outer space synthesis of lipids by shining ultraviolet, UV, light on a mixture of simple gases - H_2O:CH_3OH:NH_3:CO in the ratio of 100:50:1:1. Interstellar ice is composed of these four gases which are abundant in space where it is very cold and UV light is abundant. NASA's experiments synthesized **amphiphilic** molecules that could assemble into membrane pieces and even bi-layered, 3-dimensional spheres called **vesicles** (Figure 6.5). {Definitions: **Amphiphilic** molecules are part hydrophobic and part hydrophilic. **Vesicles** are small spheres composed of membranes and engulfing a small space inside.}

a) b)

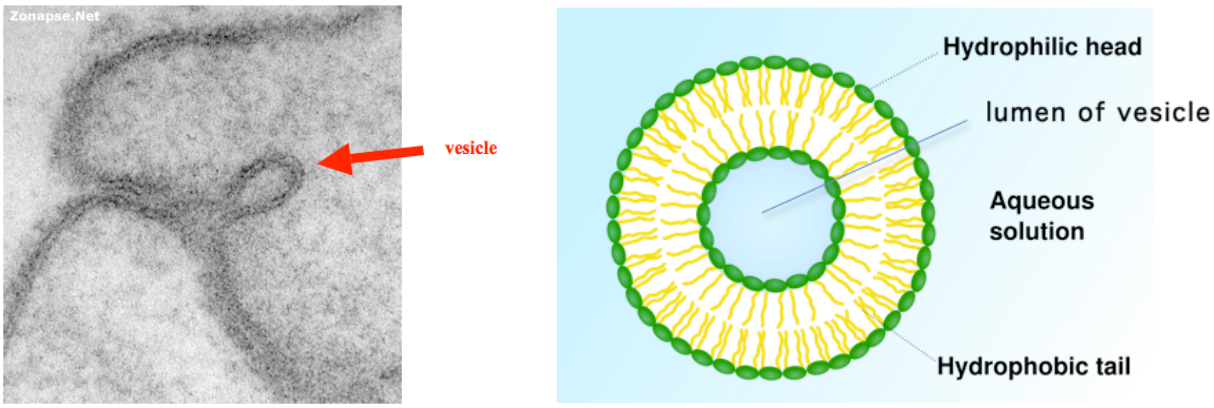


Figure 6.5 Vesicles made of lipids. Electron micrograph (a) and line drawing (b) of vesicles. Vesicles are spheres made of two lipid layers or bilayers; The center cavity of a vesicle is called its lumen.

Synthesizing complex biological molecules using electricity or outer space UV light sounds amazing, but life itself is amazing. We know that entire ecological systems thrive at the bottom of the oceans near thermal vents where microbes use sulfur-containing compounds as their primary source of energy. Perhaps even more amazing was the discovery of bacteria that lived inside 145 million year old rocks, in the absence of oxygen, taken from one mile below the ocean floor (Figure 6.6). These microbes ate tunnels through the rock, leaving clues about their crunchy diet. They never saw any sunlight or what we consider to be normal air. Clearly, life is full of surprises, but Miller and his colleagues raised a major new question: can non-living organic molecules display characteristics of life, and perhaps lead to the formation of the very first cells?

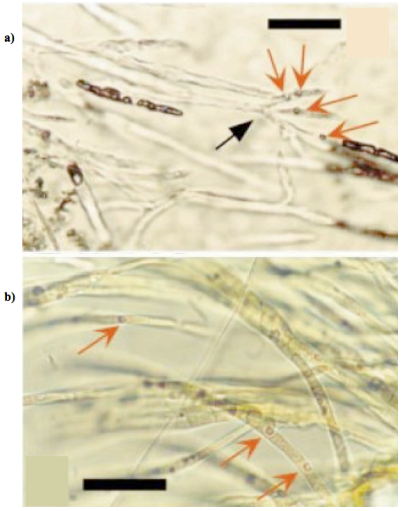


Figure 6.6 Rock-eating bacteria. (a) Tunnels eaten through rocks collected from one mile under the ocean floor. (b) Inside the tunnels are cell-like structures that may be fossil remains of microbes (red arrows). Scale bars are 10 µm.

Integrating Questions

3. What was the major outcome from Miller's primitive Earth experiment? What new question did his research findings stimulate? Search the Internet to uncover the diversity of organic molecules that have been found on the Murchison meteorite.
4. Perform a PubMed search using the phrase "The Miller volcanic spark discharge experiment" and read the 2008 *Science* abstract. Do these more recent results affect your interpretation of Miller's research?
5. It is amazing that scientists found microscopic fossilized bacteria. What is the significance of the ancient bacteria shown in Figure 6.6?

Based on the 2008 publication, we know Miller's primitive earth experiment had produced more than the three amino acids shown in Figure 6.4. Discovering that amino acids and lipid could be produced abioitically means that the first cells did not have to be composed of simple inorganic molecules. Complex, biologically important molecules may have been prevalent prior to the formation of the first cell. The 145 million year old microbial fossils and their rock tunnels demonstrate that the earliest life forms may have exhibited a lifestyle that seems odd to you now. Therefore, you need to keep an open mind when considering how the first cells could have survived without any other cells around.

You have learned that life exhibits at least three properties – change, replication, and occupying three-dimensional space. Did all three properties emerge simultaneously, or did one arise before the other? No one has direct data to answer this question, but perhaps the lack of ancient data is not an insurmountable barrier if modern data can simulate ancient conditions as Miller did. Do abiotic, self-replicating molecules exist today that can copy themselves, transferring information from one "generation" to another? {Connection: see Big Idea I is biological information.} Can abiotic, membrane-bound vesicles be produced that grow and divide? Could biochemists produce abiotic complex molecular structures with the life-like properties of change, replication, and occupying three-dimensional space? You will be able to analyze their data and determine for yourself if primitive life could have evolved from abiotically produced complex molecules.

Membranes and RNA molecules are not living objects and yet they are vital to the existence of life. In fact, if you dissected a living cell and put all its parts in a pile, none of them would be living. Life is an emergent property derived from a particular mixture of inanimate objects assembled within cells. Ribosomes are **self-organizing** molecules in that they are not assembled by a larger device. Self-organizing ribosomes assemble themselves inside cells similar to the way a life raft self inflates when it touches water. Large, multi-subunit molecules are critical to life but are constructed from non-living parts. Is it possible to devise a set of experiments that could demonstrate life-like properties in non-living objects? {Definition: **Self-organization** indicates the subunits of a larger structure assemble themselves without assistance. }

In 1985, biologists proposed a radical hypothesis for the first enzyme – it was composed of RNA and not protein. This hypothesis, called the **RNA-world hypothesis**, proposed that the earliest genomes and enzymatic molecules were all made of RNA. In the 1980s, investigators discovered a new class of RNA molecules called **ribozymes**; ribo- because they were made of ribonucleic acids RNA and -zymes because they functioned like protein-based enzymes (Figure 6.7a). The discovery of ribozymes eventually was awarded a Nobel Prize. The RNA-world hypothesis is simple, but designing experiments to test it has been challenging. {Definition: **RNA-world hypothesis** proposes the first life forms on Earth used RNA as genetic material and enzymes. **Ribozymes** function like protein enzymes but are composed completely of RNA.}

Sometimes politicians and other non-scientists consider researchers to be out-of-touch when they study bizarre organisms that have no obvious relevance to human existence. Such was the

case with biologists who studied the unicellular organism called *Tetrahymena* (Figure 6.7b). This unicellular, fuzzy creature was the origin for much of our understanding of ribozymes because one non-coding intron in a particular RNA molecule has the capacity to excise itself from the immature mRNA. This self-splicing RNA fragment was only 186 nucleotides long and it splices itself out of a larger RNA molecule and **ligates** the two other RNA pieces back together. This ribozyme intron molecule forms new covalent bonds between the two loose ends of mRNA after excising itself. In the cells of your body, you have mixed protein/RNA complexes that perform RNA splicing and ligating, but *Tetrahymena* can catalyze the reaction using only RNA – the first ribozyme discovered. From this odd and ancient organism came the beginnings of the RNA-world hypothesis and much of the experimental evidence supporting RNA as a key player in the origin of life. {Definitions; **RNA-world hypothesis** proposes that the earliest life forms stored genetic information and used enzymatic molecules made of RNA. **Ribozymes** are enzymes that are made of RNA and not amino acids. **Ligation** is the formation of covalent bonds joining two pieces of DNA or RNA.}

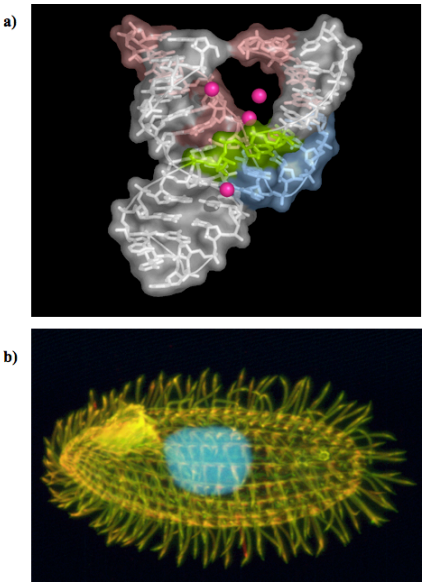


Figure 6.7 Ribozyme and Tetrahymena. a) Ribozymes fold into 3D shapes that facilitate their enzymatic functions. b) Tetrahymena is a ciliated eukaryote, meaning it has many tiny hair-like cilia (yellow) covering its single-celled body and its genetic material is contained in its nucleus (light blue).

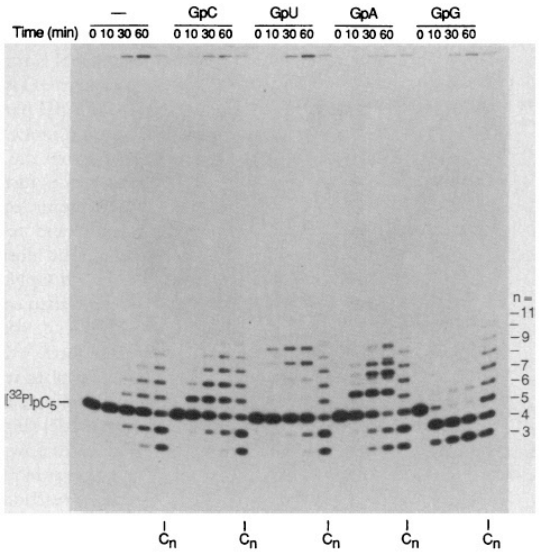
The self-splicing RNA intron does its job in two steps; first it cuts the RNA and the second step is chemically similar to the task performed by RNA polymerase. When given substrates of two different RNA dinucleotides CpU and GpN, the ribozymes produced one trinucleotide and a single base without any phosphate:

CpU + GpN + ribozyme → CpUpN + G + ribozyme.

The p in the dinucleotides represents the phosphodiester linkage between two individual nucleotides while N indicates any of four RNA bases work equally well.

Ligating one base to another is the first step in RNA polymerization and two biochemists wanted to see how many bases could be polymerized by this self-splicing ribozyme (Figure 6.8). Before conducting an experiment, the investigators had to design appropriate controls to perform in conjunction with the experimental conditions. In this case, the investigators chose to include a negative control. The negative control tested whether the self-splicing intron could add bases onto a short piece of RNA, called pC_5 which was 5 cytosine nucleotides in a row, over the course of an hour when no additional nucleotides were added to the mixture. They fully expected the negative control not to extend the length of the pC_5 RNA molecule since the control lacked any RNA dinucleotides. In the experimental conditions, one of four different dinucleotides, GpC, GpU, GpA, or GpG, was added to four separate reactions. These five different polymerization

experiments consisting of one negative control and four experimental conditions were analyzed by gel electrophoresis to separate each polymer based on its size. {Connection with Chapter 2 explaining gel electrophoresis.} Each condition was tested for 0 – 60 minutes of incubation.



Dark places on the gel represent many copies of individual molecules of identical size; the darker the spot, the more molecules of that size. The size of pC_5 is indicated, and sizes of other molecules are denoted by number of bases. Note that the pC_5 starting material was labeled with radioactive ^{32}P so that it and any of its modified products could be detected.

Figure 6.8 X-ray film from gel electrophoresis of radioactive RNA molecules separated by size. RNA molecular size is indicated on the right side by number of bases and the original pC_5 molecule marked on the left. Negative control and 4 different dinucleotides are grouped by the incubation time indicated at the top. Columns labeled C_n are molecular weight markers 3 to 11 bases.

Every time you analyze experimental data, begin with the controls first. For the negative control, you can see that pC_5 is modified over time to produce shorter and longer fragments, at roughly the same rate. Also note the new bands first appear after 30 minutes and are more prominent after an hour. Compare the results for GpC and you can see from the intensity of each band that the RNA polymers primarily get bigger and not smaller. Furthermore, the bands of bigger RNA polymers appear after 10 minutes and can become as long as 9 or 10 bases, doubling pC_5's original length.

Integrating Questions

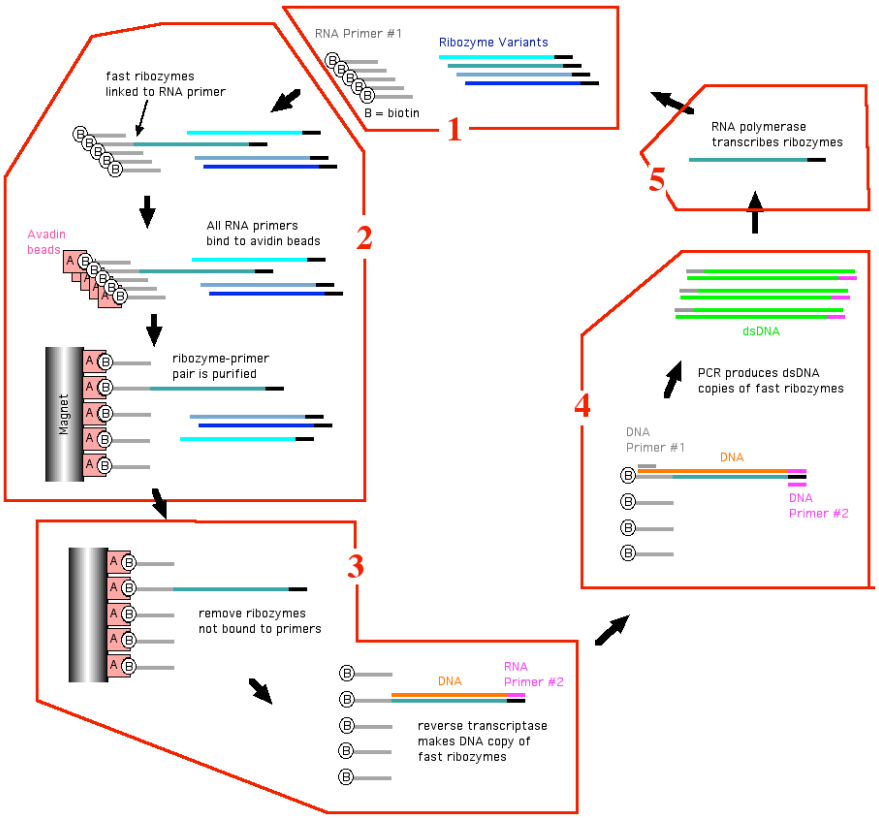
6. Are all four of the dinucleotides used equally well for polymerization by the self-splicing intron? Which base appears to be incorporated best into the growing RNA strand? Do any of the dinucleotides appear to block RNA polymerization, or shorten the pC_5 primer? Support your answers by citing data from Figure 6.8.
7. What is the significance of the experiment in Figure 6.8? How does this help build a case for abiotic molecules producing the first living cell?

From the data above, the investigators had discovered that ribozymes can add bases onto short pieces of RNA and make pC_5 longer. They also realized that this ribozyme cannot add a G to pC_5 as demonstrated by the preponderance of very short molecules in the GpG lanes of the gel. This ribozyme can only polymerize 3 of the 4 bases onto the 3' end of pC_5. Elongating a polymer of nucleic acid is the essence of replicating biological information, as you learned in Chapters 1 and 2. However, adding 4 or 5 bases is not sufficient to be considered a modern polymerase. To be more compelling, you would like to see better ribozymes that can polymerize more bases at a

faster rate.

Some naturally occurring ribozymes have been discovered, but investigators wanted to accelerate the discovery process. You can use the tenets of natural selection in an experimental process called **directed evolution** to derive new molecules that have stronger activity. In directed evolution, investigators generate many mutants**,** sometimes millions of mutants, and then humans choose the best mutants based on the desired function (Figure 6.9). If they started with trillions of ribozymes with slightly different capacities, they had to devise a clever method for isolating only those variants with improved function. They had to conduct several rounds of directed evolution, and here are the five steps that constitute a single round of directed evolution:

1) Start with the population of varied ribozymes and many copies of an RNA "prime #1" similar to pC_5 that has been covalently tagged with a small molecule called **biotin**. All ribozymes variants have a short segment of identical sequence for use in step 3.



2) Briefly incubate the ribozymes and RNA biotin-primers so the fastest ribozymes can generate new covalent bonds with the primers. Later, add to the RNA mixture millions of small magnetic beads that are coated with a protein called **avidin** which binds biotin. Pull all the avidin-beads with biotin-primers to one side of the tube by applying a magnet to the test tube.

Figure. 6.9 One 5-step cycle of directed evolution starting at the top with ribosome variants. In reality, each molecule would be present many more times than shown here. The diagram emphasizes critical steps and removes non-essential components for clarity.

3) Remove old solution and any ribozymes not attached to RNA primers. Add buffer and RNA "primer #2" to the purified primer-ribozyme pairs. The new primer base pairs to the segment of identical sequence shared by all the ribozyme variants. RNA primer #2 permits a newly added enzyme called **reverse transcriptase** to polymerize a new strand of DNA that is complementary to each ribozyme variant.

4) Add DNA primers #1 and #2 that base pair with RNA primers #1 and #2. Include in the mixture a **PCR**-ready DNA polymerase, all the DNA bases (dGTP, dCTP, dATP, and dTTP), and the appropriate buffer. PCR synthesizes billions of copies of double-stranded DNA for each of the isolated ribozyme variants that were able to couple themselves to the original biotin-modified primer.

5) Add RNA polymerase to transcribe a new single strand of RNA from each PCR product. Repeat the process as many times as you want to isolate the optimum ribozyme.

{Definition: **Directed evolution** begins with a known function that the investigators want to improve. **Biotin**, also called vitamin B7, is used to tag proteins for binding to avidin. **Avidin** is a protein that binds biotin with high affinity and is used to purify biotin-tagged molecules. **Reverse transcriptase** is a viral protein that produces a DNA strand using an RNA strand as template. **PCR**, polymerase chain reaction, produces billions of copies of dsDNA from a template DNA and a pair of DNA primers.}

The scientists constructed 10^{14} ribozyme variants derived from the original ribozyme by randomly mutating a subset of the 186 nucleotides (Figure 6.10). Producing 10,000,000,000,000 variants seems excessive until you compare this number to how many variants of length 186 were possible (4^{186}, or approximately 10^{112}; see BME 1.1). In their experiment, the number of possible variants was greatly reduced by the fact that only targeted 20% of the bases for mutation per variant, and fourteen of the 186 bases were never altered. With these limitations, the scientists expected an average of approximately 34 bases to be different in each variant (see BME 6.1). The number of different sequences they could have produced with 34 out of 172 bases changed from the original is approximately 1.7×10^{52} so you can see 10^{14} variants was not excessive (see BME 6.2). Although the scientists constructed an enormous number of variants, it was only a tiny fraction of the total number of possible variants with 20% of the bases changed from the original sequence. In Figure 6.10, green bases were not mutated; pink bases were conserved in all 25 selected sequences; blue bases were either unchanged or replaced by only one of the other three possibilities.

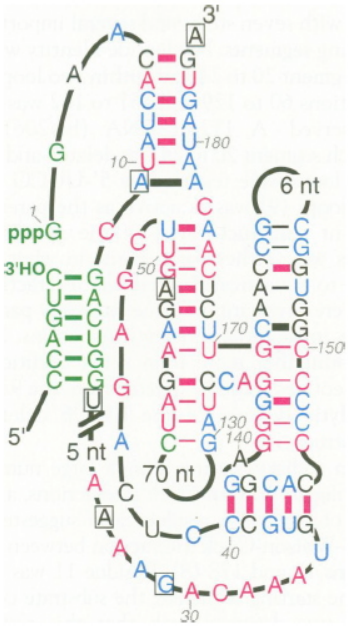


Figure 6.10 Two dimensional diagram of self-splicing ribozymes. Base pairs are denoted by dashes between complementary bases. Thick dashes indicate covariation evidence for pairing. Diagonal lines between base 19 (U) and base 25 (A) indicate where bases were removed to shorten the molecule.

The scientists conducted four cycles of directed evolution on these 10^{14} ribozymes, and at each generation they selected those ribozymes that worked fastest. They were able to isolate 25 ribozyme variants that catalyzed polymerization at least as fast as the original ribozymes from *Tetrahymena*. Pink bases indicate paired nucleotides that were conserved in at least 24 of the 25 sequences; black pairings were conserved in at least 22 of 25 sequences. After further

modification of the 25 best evolved ribozyme sequences, the investigators eventually found one human-designed ribozyme that could catalyze polymerization 700 times faster than the biological one isolated from *Tetrahymena*.

Bio-Math Exploration 6.1 How many mutations do you expect?

- Concept: Expected values predict the outcome of random events.
- Objective: Calculate the expected number of mutations in ribozyme variants.
- Required Skills: Multiplication, reading a histogram.

{*Editorial Note: This BME has been rated "green," meaning that most students should be able to master the techniques.*}

Prediction is an important part of designing and interpreting the outcomes of experiments. In this experiment, scientists needed to predict how many bases would vary from the original sequence if 14 of the 186 bases were never mutated, and the mutation rate was 20 percent for the remaining 172 bases. This is like predicting how many flips of a coin would come up heads in 172 flips of an unbalanced coin that only shows heads 20% of the time. In both experiments, we are interested in the number of "successes," and the probability of a success is 20%, or 0.2. Let's explore the concept of expected value in the more familiar context of flipping coins.

To find the expected number of successes in repeated independent trials of an experiment, multiply the probability of a success on a single trial by the number of trials. In the coin flipping example, flipping heads is a success, so you multiply the probability of getting heads on a single flip (0.2) by the number of flips (172) for an expected value of 0.2 x 172 = 34.4 heads. Of course, you would never get exactly 34.4 heads in 172 flips. But don't make the mistake of rounding off this number and claiming that the expected number of heads is 34. If it seems strange that the "expected" number is a number of heads that you can't possibly get, it might help to think of the expected value as the long run average number of heads in 172 flips if you repeat this coin-flipping experiment thousands of times. For example, if you repeat the experiment 1000 times, flipping the coin 172 times each, you will have flipped the coin a total of 172,000 times. If you average the numbers of heads you get in each of the 1000 repetitions, you should get approximately 34.4. The more times you repeat the experiment, the closer the average will be to 34.4. Computing the expected number of heads by multiplying 0.2 x 172 is a shortcut way to get this same result.

Bio-Math Integrating Questions

BME IQ 6.1 Go to the coin simulator at http://www.math.uah.edu/stat/applets/BinomialCoinExperiment.xhtml Move the first slider to n=43, which will give you 43 coins to flip. Move the second slider to p=0.2, which will make each coin come up heads 20% of the time. Calculate the expected number of heads and continue to the next question.

BME IQ 6.2 In the top right panel is the theoretical distribution, illustrating what proportion of the time you should expect to get each possible number of heads, from 0 to 43. These proportions are also given in table form in the bottom right panel. The expected value is marked with a line under the theoretical distribution, and given at the bottom of the table in the line labeled "Mean," another word meaning "expected value." To flip all 43 coins once, set the update frequency to "Update 1" and press the "single step" button that looks like a play button on a media player. Each time you hit the single step button, the coins are flipped again, and the number of heads are recorded in the table on the bottom left, graphed in a histogram superimposed on the theoretical distribution, and averaged into the "Data" column of the table on the bottom right. Continue to step through the flips, 43 at a time, until you feel comfortable with how the simulation is working. Now change the stop frequency to "Stop 1000" and hit the "Play" button (which looks like a fast forward button) to flip the 43 coins an additional 1000 times. Describe what happens to the histogram of outcomes and the average number of heads as you repeat the experiment more and more times. {*Editorial Note: These detailed instructions could be moved to WileyPLUS, and the question shortened to just the last sentence.*}

Let's apply what you learned about outcomes of random events to mutations of a DNA sequence, in which a "success" is a base being changed. Each base is a "trial" of the simple experiment to change a base, and we assume that each base has the same probability of being changed, independent of all the other bases. The expected number of changes is therefore the probability of a change at a single base (0.2) multiplied by the number of bases mutated (172). Therefore, you expect 0.2 x 172 = 34.4 base changes. As in the coin-flipping example, you would never get exactly 34.4 bases changed. But you don't always get 34 or 35 changes, either. Although it is extremely unlikely, it is even possible that zero bases would be changed. The only certainty is that the number of changes will be between 0 and 172.

Now you know why the scientists expected about 34 bases to be changed in each ribosome variant. BME 6.2 shows how to determine the total number of sequences with exactly 34 bases changed from the original.

---------- End of BME 6.1 ----------

Bio-Math Exploration 6.2 Are you sure this is the best possible sequence?

- Concept: Counting possible outcomes using a two-step process.
- Objective: Determine the number of possible sequence variants when 34 out of 172 bases are mutated
- Required Skills: Multiplication principle (BME 1.1).

{*Editorial Note: This BME is rated "yellow," because it combines concepts from two previous BME's.*}

The multiplication principle and binomial coefficients can be used to count the number of different sequences with 34 out of 172 nucleotides changed from the original. To count the number of such sequences, we break down the process of mutating sequences into the following two steps, and count the number of possibilities at each step.

Step 1: Choose which 34 out of 172 nucleotides are changed. The number of ways to make this choice is given by the binomial coefficient $\binom{172}{34} = \frac{172!}{34!(172-34)!}$, approximately 1 x 10^{36}

Step 2: Count the number of ways to change the selected 34 nucleotides. Because there are 3 ways to change each nucleotide, we can use the multiplication principle (see BME 1.1) and determine that there are 3^{34} possible ways to change the selected 34 nucleotides.

Because this is a two-step process, the multiplication principle can be applied to count the number of choices in the entire process. We multiply the number of ways to do each step in the sequence mutation process, and determine that there are $\binom{172}{34} 3^{34}$, or approximately 1.7 x 10^{52}, possible sequence variants with 34 out of 172 nucleotides changed from the original.

---------- End of BME 6.2 ----------

One of the recurring rules of biology is that "form meets function" which means that the shape of an object determines what processes an object can perform. {Connections to Chapters 11 and 12 which describe several critical structure/function relationships.} By comparing the RNA sequences of the 25 most efficient ribozyme variants, the investigators predicted the 2D structure of the ribozyme (see Figure 6.10). One piece of experimental evidence that helped predict the 2D structure was **covariation**, in which some pairs of bases either both stayed the same or both changed to maintain complementarity meaning bases A and U formed pairs and bases G and C formed pairs. The probability of a particular base pairing being conserved by chance in at least 24 out of 25 sequence variants is approximately 0.0021 (see BME 6.3). Therefore, when several consecutive base pairings are conserved in at least 24 of the faster ribozymes, it was logical to deduce that the bases form pairs in the folded RNA molecule. You can see conserved base pairs in Figure 6.10 such as: 36 – 79; 37 – 78; 38 – 77; 39 – 76; and 40 – 75. {Definition: **Covariation** describes the degree to which two events are synchronized.}

Bio-Math Exploration 6.3 What is the probability of a highly conserved base pairing?

- Concepts: Multiplication and addition rules for computing probabilities.
- Objective: Deduce likely RNA structural elements by determining probability of chance covariation.
- Required Skills: Multiplication, addition and complement rules for computing probabilities (see BME 3.3).

{Editorial Note: This BME is rated "yellow," because it applies probability concepts from another BME. This BME reinforces the fundamental concepts of multiplication and addition rules, first introduced in the context of Mendelian genetics.}

The probabilities of base conservation and base pairing conservation helped scientists predict the critical nucleotides in the ribozyme, as well as its 2D structure. For example, the rules for computing probabilities, introducing the addition rule, multiplication rule, and complement rule in BME 3.3, can be used to compute the probability that a particular base pairing would be conserved by chance alone in at least 24 out of 25 randomly mutated sequences. We begin by breaking the problem into two parts, because the probability of a base pairing being conserved in at least 24 of the sequences is the probability it is conserved in exactly 24 sequences *or* it is conserved in all 25 sequences. Because the key word here is "or," we need to add these two probabilities.

The probability of a particular base pairing being conserved after random mutations in a single sequence is 0.71 (see *WileyPLUS* for a derivation of this probability), and the probability that a particular base pairing is *not* conserved (the only other choice) is (1 – 0.71) = 0.29. The probability that a particular base pairing is conserved in all 25 sequences (*i.e.*, conserved in the first one *and* the second one *and* the third one, and so on) is 0.71^{25} by the multiplication rule. Similarly, the probability that a particular base pairing is conserved in a particular set of 24 sequences, and not the 25th sequence, is 0.71^{24} x 0.29. There is a hidden "or" in this part of the problem, because there are 25 different ways to choose which one of the 25 sequences will not have the base pairing conserved (the first one *or* the second one *or* the third one, etc.). Therefore, we must add all 25 of these probabilities. Because all 25 probabilities have the same value of 0.71^{24} x 0.29, the probability that a particular base pairing is conserved in exactly 24 out of 25 sequences is 25 x 0.71^{24} x 0.29. The final answer for the probability of a particular base pairing being conserved in at least 24 out of 25 variants by chance is 0.71^{25} + 25 x 0.71^{24} x 0.29 ≈ 0.0021.

---------- End of BME 6.3 ----------

Integrating Questions

8. Go to NCBI and retrieve the evolved ribozyme version 2.0 using the accession number U26413. How many bases are in this ribozyme? Now use the sequence comparison program nucleotide BLAST to determine if any organisms have a similar sequence to the one produced through directed evolution. Paste U26413 into the blank field and choose "nucleotide collection" for the database, then hit the "BLAST" button. How many good matches did you find? Are any of them natural products, or only engineered ribozymes?
9. Do you think it is possible that ribozymes with faster activity than the one from Tetrahymena could evolve in nature given enough time? Do you think your BLAST database results indicate the complete range of naturally occurring ribozymes? Explain your answer.

10. Scientists often try to predict structures based on base pairing rules and nucleotide sequences so they can understand how ribozymes work. View the structure of a different ribozyme found in nature to determine if you think we have a good chance of predicting structure of ribozymes. Use the Jmol tutorial to view the 3D structure of a different ribozyme numbered 2oeu. Does the RNA always fold into a double helix? Does RNA follow the same rules as DNA when it comes to base pairing? What effect does this structure have on your confidence that anyone can predict the folding of RNA? {*Editorial Note: these directions will be incorporated into an online tutorial so students will not have to type these commands. What Jmol will do for us: Right click on the molecule and open the console window. Copy and paste these commands:*

    select :a
    cartoon off
    wireframe 100
    spacefill 100
    color brown
    select :b
    cartoon off
    wireframe 100
    spacefill 100
    color gray
    select A21
    color CPK
    select G36
    color CPK

The new, human-evolved ribozyme in Figure 6.10 polymerizes RNA at rates comparable to protein-based enzymes (on the order of 100 bases per minute). Some bases in the ribozyme were always conserved after the directed evolution. Using the mutation rate of 20% and the multiplication rule for probability calculations (BME 6.3), the probability that a base was conserved in all 25 variants by chance alone is $0.71^{25} \approx 0.0002$. Therefore, it seems likely that these conserved bases were critical for the ribozyme's function. The evolved bases in the best ribozymes probably contribute to the enhanced capacity as would be predicted by the form-meets-function rule. Comparing the old and new ribozyme sequences helped us understand which bases were critical for the enhanced function. However, we know so little about ribozymes structure, it is impossible to predict which bases actually catalyze the ligation reaction. Since non-standard base pairing can happen, it would be impossible to BLAST a sequence to find a functional structure since RNAs can fold in unpredictable ways. Furthermore, the BLAST database does not contain every sequence on Earth, so any searches you perform will not survey the true diversity that exists in nature. Therefore, it is possible that faster ribozymes are still functional in a species we have not sequenced, or uses different nucleotides to perform its

function. Either way, all we can do now is continue to search through odd creatures who may hold ancient clues about early ribozymes that could represent the earliest forms of transmitting genetic information from one generation to the next.

Earlier, you learned that the existence of a RNA-based polymerase was first hypothesized and later found in nature. Through directed evolution, biologists were able to produce a ribozyme with improved capacity to form RNA polymers. In short, the hypothesized molecule does exist and a naturally produced counterpart may have been the first biological information encoder. Ribozyme RNA polymerases satisfy two of our criteria for the bare minimal requirements for the origin of life - replication and change over time. Now you need to discover whether nature could have produced lipid membranes capable of growing in the absence of enzymes or living cells.

6.3 Can non-living objects compete and grow?

- Context: Abiotic processes seem devoid of qualities associated with life such as growth and competition for limited resources.
- Major Themes: The origin of living systems occurred by natural processes, and life continues to evolve within a changing environment; organisms can be linked by lines of descent from common ancestry; natural selection is a mechanism of evolution that accounts for adaptation.
- Bottom Line: Abiotic vesicles can grow, reproduce and compete for limited resources.

Perhaps the most challenging aspect for us to understand about the origin of life is the formation of a cell membrane before cells existed. It seems nearly impossible for abiotic forces to organize lipids into a 3D sphere, and allow these spheres to grow and produce more spheres. Could one or more lipids self-organize and replicate solely based on the biophysical properties determined by the shape of the molecules? If such lipids exist, are they capable of encapsulating any cargo such as self-replicating ribozymes? Perhaps the greatest challenge will determining whether abiotic lipid vesicles could harvest and store energy. If a vesicle possessed all of these properties, you could imagine how life might have evolved over a long period of time through abiotic actions.

As described earlier, amphiphilic lipids and fatty acids have been synthesized on meteors, on ice exposed to UV light, and at deep sea thermal vents. But with Earth being so large, it is hard to imagine how enough of these fatty acids could congregate in one place to coalesce into membranes. One particular fatty acid called **myristoleate** (Figure 6.11a) does congregate on a type of natural clay called montmorillonite. To get a sense how chemicals can cluster on particular surfaces, drop a small piece of food such as a pretzel or cracker crumb into a soda and watch the bubbles form on the surface of the food particle, causing it to rise and fall in your drink. Similarly, myristoleate and other fatty acids can cluster onto the surface of clay when both are immersed in a watery environment. Lipids and fatty acids can form solid balls of lipids called **micelles** when mixed with water (Figure 6.11b). Clay serves as a catalyst for micelles because it enhances their formation, but the clay is not consumed in the process. The formation of vesicles

is detected by measuring the amount of light absorbed by a solution of clay and lipids. Just as light cannot shine through a glass of muddy water, vesicles prevent light from passing through the solution. The amount of vesicles present is measured by the amount of light absorbed by the sample. Vesicle formation from 10 mM myristoleate miscelles was tested with different solid surfaces such as tiny plastic beads called microspheres and different concentrations of clay (Figure 6.11c). They tested montmorillonite clay, ceramic microspheres with a high density of negative charges, ceramic microspheres with low density of negative charges, and buffer alone as a negative control. The team measured vesicle formation using 10 mM myristoleate miscelles and different concentrations of montmorillonite clay; 0.02 mg/mL, 0.1 mg/mL, 0.5 mg/mL, and a negative control of buffer (Figure 6.11d). In Figure 6.11e, they plotted the initial rates of vesicle formation using different amounts of clay using the data from Figure 6.11d. {Definition: **Myristoleate** is a fatty acid that self-organizes on clay surfaces and can form larger lipid molecules. **Micelles** are non-hollow spheres of lipids where the fatty tails fill the interior of the ball while the hydrophilic portions form the outer layer that interacts with the water. }

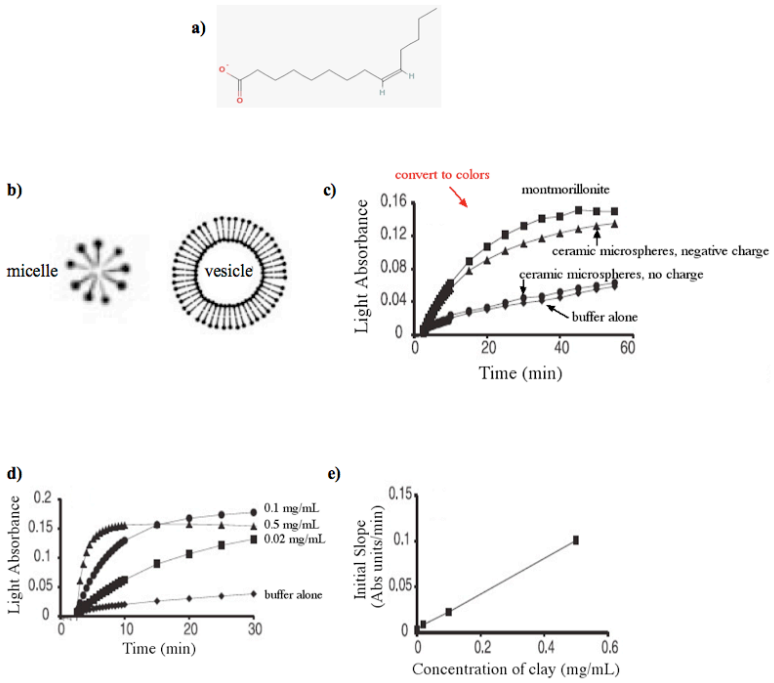


Figure 6.12 Fatty acid structures and parameters that affect vesicle formation rates. a) Myristoleate chemical structure. b) Fatty acid monolayer forming a micelle and a bilayer forming a membrane vesicle. c) Effects of different solid surfaces for vesicle formation. d) Effects of different concentrations of clay on vesicle formation. e) Initial rates of vesicle formation based on clay concentrations.

Integrating Questions

11. Lipids and fatty acids can self-organize into membranes because of their chemical and physical properties. To understand membranes better, go to http://biomodel.uah.es/en/model3/index.htm, and click on "Lipid bilayer" to see a small piece of a membrane. Rotate the block of membrane. Describe the amount of space between each molecule. Use the check box to display the water. Which atoms interact with the water and which atoms are hydrophobic? How deep does the water penetrate into the membrane? Note: palmitic is a particular type of fatty acid in this lipid.
12. Which type of solid surface used in Figure 6.11c is the best at forming vesicles? What effect does charge have on vesicle formation? Which concentration of clay leads to the fastest vesicle formation? To determine the best concentration for stimulating vesicle

formation, should you look at the last time point, or the first few time points? Explain your answer using Figures 6.11d and 6.11e.

13. Based on the relationship between the concentration of clay and the rate of vesicle formation, predict what would happen if you added a little more clay? Predict what would happen in Figure 6.11d if you added more lipid to the experiment containing 0.5 mg/mL of clay (triangles) at time 20 minutes?

From these data in Figure 6.11, you saw that many solids are capable of catalyzing vesicle formation. If you performed the cracker crumb and soda experiment, the rate of bubble formation increased as you added more surface area of thin crackers flakes, not fat chunks of cracker. In other words, two small flakes are more effective than one larger crumb because of the increased surface area for bubble formation. {Connection with Chapter 11 on surface area and volume.} You can also see that the higher two concentrations of clay consumed all the micelles into vesicles and the amount of light absorbed by the solution stopped increasing after 10 minutes. Chemical reactions often stop after a while, so you should always use the earliest time points to examine rates of reactions. Furthermore, it is the slope of these initial time points that tells you how fast a reaction is proceeding, not the value at the final time point.

What is not evident from Figure 6.11 is that bits of clay and microspheres can become entrapped inside the growing vesicles (Figure 6.12). Green-stained fatty acids were mixed with negatively charged ceramic microspheres and viewed in different ways by microscopy to see the different components. This was the first time anyone had shown that self-organizing vesicles could spontaneously trap objects inside the lumen of the vesicles, including smaller vesicles (Figure 6.12d). These results indicated abiotic vesicles could contain organic molecules that perhaps had enzymatic activity. The investigators wanted to determine whether the abiotic vesicles could entrap RNA molecules as well, such as ribozymes which you have already learned can perform simple genetic replication. When mixed with RNA stained red and attached to clay (Figure 6.12e) or free floating in solution (Figure 6.12f), the green-stained fatty acid vesicles were able to provide a 3D protective layer around the red nucleic acids. Figures 6.12e and 6.12f look a lot like membrane-bound genomes as exhibited by prokaryotes today.

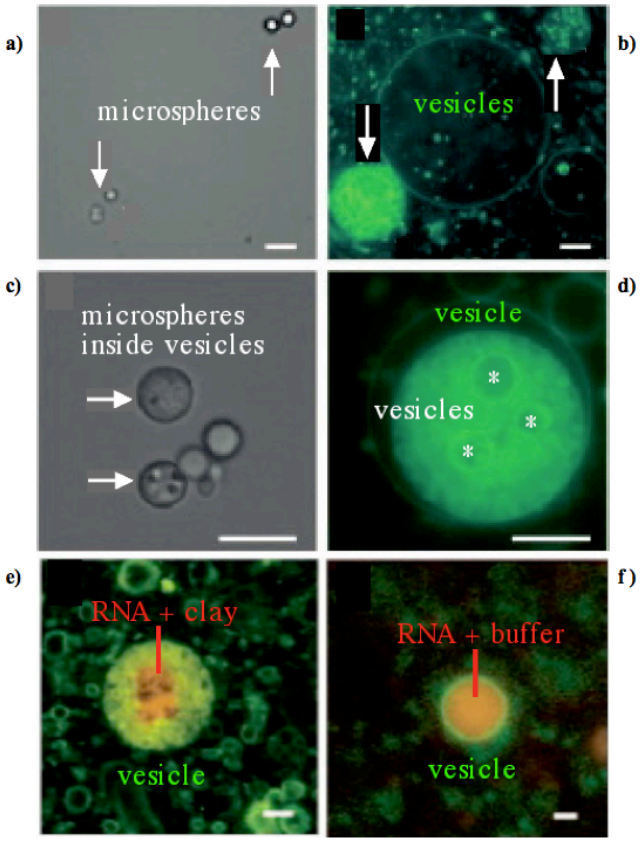


Figure 6.12 Microscopy of solid particles and vesicles. Green-stained fatty acids mixed with negatively charged

ceramic microspheres. White light in panels (a) and (c) to emphasize the microspheres. Green fluorescence in panels (b) and (d) to highlight the vesicles. Red-stained RNA attached to clay (e) or in solution (f) trapped inside vesicles. Size bars are 5 µm in a-d, and 1 µm in e and f.

Living organisms grow, and the first cell would need to be able to grow and eventually divide to produce two cells. To measure whether the experimentally produced abiotic vesicles could grow spontaneously, the investigators measured the size of the vesicles at time zero as well as different times after the addition of more fatty acid micelles (Figure 6.13). They used light scattering to measure the size of vesicles and count the number of vesicles before and after adding the micelles. As you can see in Figure 6.13a, the entire population of vesicles got bigger. The investigators produced a narrow distribution of starting size vesicles by pushing all the membranes through a filter similar to a coffee filter. After adding micelles and waiting for 4 hours, the vesicles were measured again and found to be larger. Growth of myristoleate vesicles was linear over four hours in response to the gradual addition of myristoleate micelles. Therefore, once a vesicle is formed, it can become "greedy" and incorporate more and more fatty acid to increase its size.

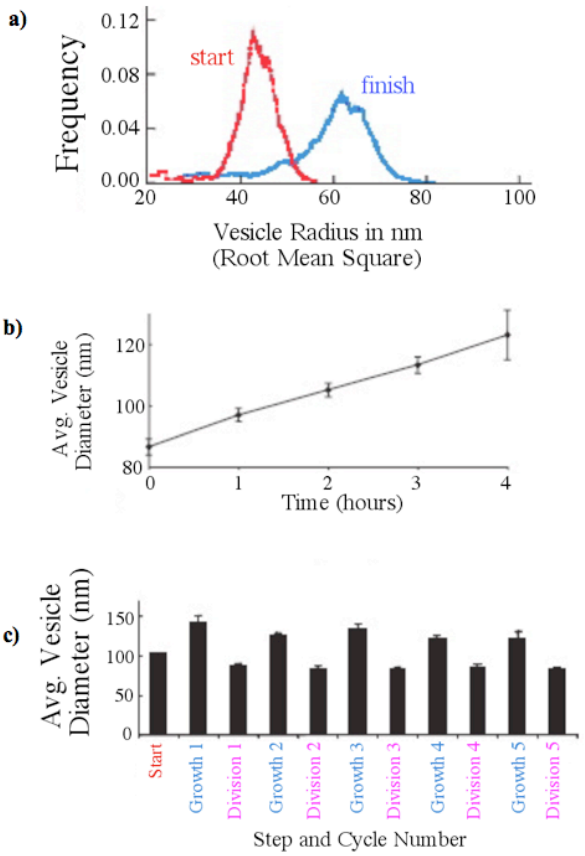


Figure 6.13 Changes in vesicle sizes. a) Small starting vesicles grew larger after the addition of micelles. b) A similar experiment with periodic sampling. c) Mean diameter + standard deviation of four measurements demonstrates reproducibility of vesicle growth.

Another characteristic of life is the ability to reproduce. Reproduction for abiotic vesicles would mean producing more vesicles using abiotic processes. Since these myristoleate vesicles lacked any proteins or enzymes for cell division, you would need to determine whether physical and chemical forces alone are sufficient to generate new, and thus smaller, vesicles. The investigators pushed the larger vesicles through the filter again, and you can see what happened to the average size of a vesicle (Figure 6.13c). Pushing large vesicles through smaller holes might be similar to the physical pressure exerted on vesicles in shallow pools that evaporated over time. The membrane vesicles were essentially squeezed until they budded off smaller daughter vesicles from the maternal vesicles. The investigators repeated the cycle of growth through accumulation of more fatty acid and division by extrusion through filter paper. You can recognize the 5-cycle alternating size pattern similar to the life cycle of organisms that reproduce asexually. {Connection with Chapter 3 on cell division.} The following Integrating Questions will ask you to apply logic and mathematics to understand the consequences of abiotic growth of

vesicles as demonstrated in Figure 6.13. Your answers will have an impact on the experiments that you will analyze immediately after the Integrating Questions.

Integrating Questions

14. As a vesicle grows in size, its volume and surface area change at different rates. What was the percent change in vesicle radius over four hours in Figure 6.13a? Calculate the volumes and surface areas for the two major size classes of vesicles. The formula for the surface area of a sphere is $4\pi r^2$ and the formula for the volume of a sphere is $4/3\ \pi r^3$. Use the highest point in the two graphs to estimate the radii for the two populations of vesicles. If you round to the nearest hundred, what was the percent change in vesicle surface area and volume?

Vesicle Type	Radius (nm)	Surface Area (nm^2)	Volume (nm^3)
Small vesicle			
Large vesicle			
Percent Change			

15. In Figure 6.13c, the large maternal vesicles and the newly formed daughter vesicles had average radii similar to those in Figure 6.13a. When a vesicle divides by extrusion through a filter, does its membrane surface area or its total enclosed volume remain constant? If an RNA genome were replicating inside one of the maternal abiotic vesicles, what could happen to the genome with each round of membrane division? Revisit Figure 6.12f and what you just learned about volume and surface areas. Speculate about the consequences of daughter vesicle formation if the maternal vesicles contained self-replicating RNA genomes?
16. Propose a mechanism for primitive cells composed only of greedy lipid vesicles and self-replicating RNA genomes could become a population of cells with variation. Utilize your answers to the previous two Integrating Questions and propose how a particular genome could be encapsulated by membranes composed of different fatty acids and lipids.

At this point, you have learned that lipids and fatty acids can assemble into membranes and form vesicles. These vesicles are catalyzed on the surfaces of clays found in nature today. You also found that abiotic vesicles can grow in size through the accumulation of more fatty acids. These vesicles can be broken into smaller vesicles as a consequence of physical forces such as being pushed through small holes. Furthermore, you saw that RNA can be trapped inside the vesicles. In the previous Section, you learned that some RNA molecules are enzymes and these ribozymes can polymerize RNA. Perhaps there are ribozymes that can completely replicate themselves, though no one has found self-replicating ribozymes, yet. Many of the abiotic vesicle characteristics sound very similar to living cells, though the experiments have utilized only physical and chemical processes driven by the shape and charge of non-living systems.

When you answered the Integrating Questions above, your mathematics proved that the vesicles maintain a surface area that approximately doubles before dividing in half. At the same time, you mathematically proved the volume increases about 300% and is reduced to its original 100% with each round of vesicle division. The consequence of each abiotic vesicle formation is a loss of about one third of a vesicle's lumenal contents. If the vesicle contained a self-replicating ribozyme, some of these molecules would be spilled into the surrounding media. As you saw in Figure 6.12f, RNA can be encapsulated in newly formed vesicles which means different copies of an RNA genome could end up inside vesicles composed of different fatty acids or lipids. It is striking that with only two types of abiotic molecules, you deduced a mechanism to produce diversity within a population of growing and dividing abiotic vesicles that contain RNA genomes. In short, you have discovered the first tenets of natural selection using fatty acids, nucleic acids and mathematics.

Three billion years ago when life was first evolving, the building blocks of cells were present but in limited concentrations which means resources were limited. Limited resources, variation in the population and over production are the first three tenets to natural selection. The only missing piece is competition. Could these RNA/fatty acid primitive "cells" compete for limited resources needed to make more copies of themselves? The number of lipid molecules formed abiotically on Earth was probably limited and concentrated on clay surfaces. In order for life to evolve, you would predict the earliest cells would have to compete with one another for the few available lipids. It would be very informative if abiotic vesicle competition could be experimentally replicated. If vesicles can compete, then there would be a selective advantage for a few vesicles over the many other smaller ones and all the components of natural selection could have been in place 3 billion years ago when we know from fossil evidence life first evolved.

To investigate the possibility of vesicle competition, the scientists produced two populations of vesicles made of an amphiphilic fatty acid called **oleate**. One set of vesicles were filled with one molar, 1M, sucrose sugar made by dissolving 342g of sucrose into 1 L of water. As you may remember from chemistry classes, when a substance diffuses, it moves down its concentration gradient to an area of lower concentration. In the case of 1M sucrose, water tries to passively diffuse from outside the vesicle into the lumen where there is less water due to the high concentration of sugar. In other words, the water will dilute the 1M sucrose inside the vesicle causing it to swell as water traveled across the membrane. When water is drawn across a membrane into areas of concentrated dissolved particles, we use the term **osmosis** to describe water's passive movement. The second set of "relaxed vesicles" was prepared with just buffer inside them so there was no osmotic pressure for water to move in either direction. The investigators wanted to know if the sucrose-filled, "osmotically stressed" vesicle would be able to steal lipids from the relaxed, buffer-filled vesicle. {Definition: **Oleate** is a fatty acid that is prevalent in olive oil. The diffusional force exerted by the water is called **osmotic pressure** or stress. }

The two populations of vesicles were labeled different colors so the investigators could distinguish which vesicles were stressed and which ones were relaxed. After mixing the two types of vesicles in a 1:1 ratio, they measured the surface area of both types of vesicles (Figure 6.14). Each dot represents a different measurement of one vesicle at a single time point. Relaxed vesicles were measured after adding additional relaxed or osmotically stressed oleate vesicles (Figure 6.14a-b). Osmotically stressed vesicles were measured after adding equally stressed or relaxed vesicles (Figure 6.14c-d). When vesicles with the same osmotic pressure were mixed together, they did not grow any bigger (Figure 6.14a and c) but mixing stressed and relaxed vesicles produced rapid changes in vesicle sizes (Figure 6.14b and d). By finding exponential curves that best fit the observed data (solid lines in Figure 6.14b and d), the investigators determined that relaxed vesicles shrank at about the same rate that the stressed vesicles grew, supporting their conclusion that fatty acids moved from one type of vesicle to the other.

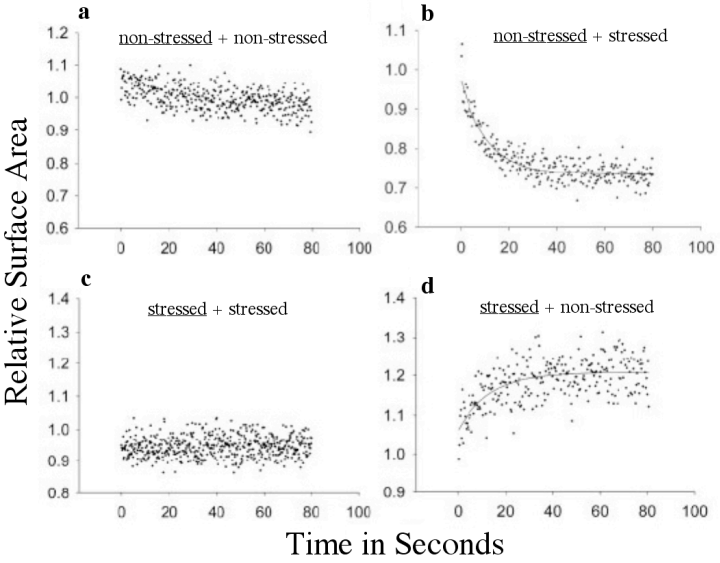


Figure 6.14 Measuring surface areas of vesicles relaxed vesicles were measured after adding relaxed (**a**) or osmotically stressed (**b**) vesicles. Osmotically stressed vesicles were measured after adding equally stressed (**c**) or relaxed (**d**) vesicles. Solid lines indicate exponential curves with best fit to the data.

This elegant experiment was the first one to show that the physical and chemical properties of lipids could lead to a competition driven by the osmotic pressure contained within a vesicle. Therefore, if any vesicle increased its osmotic pressure, it would steal away lipids from vesicles with lower osmotic pressure. All of this competition happened in the absence of life or any enzymes. These investigators demonstrated the existence of abiotic competition for limited resources, the remaining tenet of natural selection.

Bio-Math Exploration 6.4
How fast is the vesicle size changing?

- Concept: Fitting an exponential function to data.
- Objective: Quantify the rate of change in vesicle size.
- Required Skills: change values in a spreadsheet (alternatively, use the exponential and graphing functions on a graphing calculator)

{Editorial Note: This BME is rated "amber" because it uses concepts from pre-calculus and requires independent exploration.}

The relaxed vesicles in Figure 6.15b are getting smaller, and the stressed vesicles in Figure 6.15d are getting larger. The investigators quantified the rates of change in their sizes by fitting

curves to the data. The shape of the curve in Figure 6.15b is called **exponential decay**. An exponential decay curve drops off quickly at first, and gradually levels off. {Definition: **Exponential decay** is the loss of a substance at a rate proportional to the amount of substance currently present. }

The equation for an exponential decay curve is $y = ae^{-kt} + b$, where y is the dependent variable, t is the independent variable, and a, b, and k are constants. In this example, y is the relative surface area, and t is time. This equation is called exponential because the independent variable (t) is in the exponent. The number e is approximately 2.1718, and coded into scientific calculators and spreadsheet software. It is traditional to use e as the base (the number raised to the power $-kt$) because it has some useful mathematical properties (beyond the scope of this BME).

Bio-Math Integrating Questions

BME IQ 6.4: Use a graphing calculator or the Excel file exponential_graphs.xls to sketch the graph of $y = ae^{-kt} + b$ for the values of a between 0 and 1, values of b between 0 and 1, and values of k between 0 and 5. Describe how the value of each constant affects the shape of the exponential decay curve.

The values of the constants a, b, and k can be determined so that the curve "fits" the data as closely as possible. In Figure 6.15, the value of k represents the rate of change in vesicle size. By finding the rates at which the sizes are changing, the investigators confirmed their hypothesis that the two populations of vesicles are exchanging lipids rather than changing sizes by two independent mechanisms.

---------- End of BME 6.4 ----------

The biologists realized that 1M sucrose was not a biologically relevant solution to generate osmotic pressure inside abiotic vesicles because 1M sucrose does not exist inside cells. They decided to use tRNA as a more biologically relevant source of a large molecule that could generate osmotic pressure. tRNA is easy to purify and about the same size at the ribozymes you studied in Section 6.2. Once again, they measured the change in surface area of relaxed vesicles (Figure 6.15a) and stressed vesicles (Figure 6.15b) when mixed with vesicles of similar osmotic stress or the opposite osmotic stresses. The biological relevance of this final competitive experiment should be clear because you have already seen that RNA molecules can become trapped inside abiotically formed vesicles. This team of investigators has demonstrated:

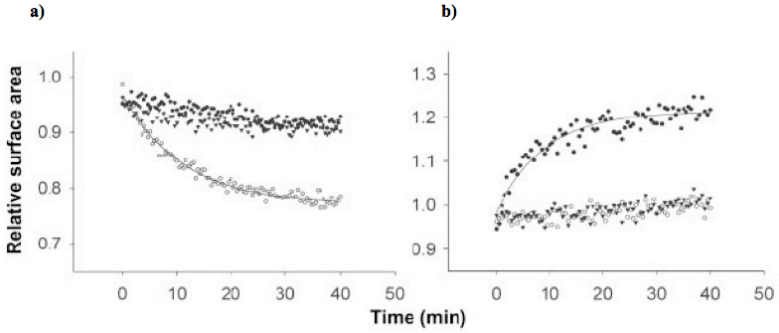


Figure 6.15 Lipid competition using tRNA to swell vesicles. Non-stressed (a) or stressed (b) vesicles were measured after mixing with stressed vesicles pressurized by tRNA (open circles), non-stressed vesicles (solid circles), or buffer only (triangles).

- abiotic vesicles can form spontaneously
- abiotic vesicles can trap RNA inside them
- abiotic vesicles can compete for fatty acids
- abiotic vesicles can grow through successful competition
- abiotic vesicles can divide and regrow in size.

These biologists claim that their data support their original hypothesis that abiotic factors could have led to the origin of life. Earlier in this Chapter, you read that 3 billion year old data of the first cells is impossible to obtain, but if primitive cells could be produced on a simulated ancient Earth, then abiotic origins of life would be a realistic possibility. You have analyzed data that demonstrate the tenets of natural selections are possible in the absence of life which begs the question – how do you define a living cell?

Integrating Questions

17. Interpret the results from Figure 6.15a and Figure 6.15b. Explain why some vesicles grew and others shrank. What is the biologically significant difference between Figure 6.15 and Figure 6.15?
18. BME 6.4 can help you describe the shapes of the curves shown in Figures 6.15b and Figure 6.15d. Hypothesize why stressed vesicles stopped growing once they had increase by ~ 30%. What physical forces would have set this limit? Try to design an experiment to test whether 30% growth can be exceeded or not.

All of these experiments use the chemical and physical properties of osmosis to direct the movement of fatty acids from one vesicle to another. As a vesicle shrinks when outcompeted for fatty acids, the osmotic pressure should increase inside the shrinking vesicles because its volume is reduced but the chemicals inside the lumen are trapped, either sucrose or tRNA. Conversely, growing vesicles should decrease their osmotic pressure at the same rate as shrinking vesicles increase their osmotic pressure. These predictions could be tested experimentally by adding more micelles to those vesicles already changed in size to determine if they continue to change or they maintain their size. Furthermore, you could alter the osmotic balance of the solution they were floating in to see if this could affect vesicle growth or not.

Science is a discipline of asking and answering questions. You have analyzed the best data in the world trying to determine how life could have evolved on Earth for the first time. Abiotic factors of chemistry and physics are sufficient to explain many of the traits you associate with living cells. Once the simplest cells were present, natural selection would have led to an ever increasing diversity of life forms. Keep in mind, evolution is an explanation of the natural world using data while faith and religion explain spiritual and personal beliefs and independent of data. Religion and science are not mutually exclusive and many people of faith believe the chemical

and physical properties were God's creation and life simply evolved from the big bang initiated by God.

6.4 Can non-living objects harvest and store energy?

- Context: Life consumes energy and the first cells needed an abiotic way to harvest and store energy.
- Major Theme: Organisms can be linked by lines of descent from common ancestry.
- Bottom Line: Non-living vesicles can accumulate energy in the form of a pH gradient.

At this point, you have seen RNA molecules perform enzymatic reactions that led to growing strands of nucleic acid and these RNA molecules could become entrapped inside abiotic vesicles. RNA trapped inside a vesicle produces sufficient osmotic force to cause the stressed vesicle to out-compete relaxed vesicles for the limited resource of fatty acids. Life also requires energy and so far you have not seen any data indicating these primitive cells could harvest or store energy. Is it possible to store energy in an abiotic world? Can growing vesicles sequester energy that could be used to do work at a later time? The final set of experiments in this section examines this highly improbable question.

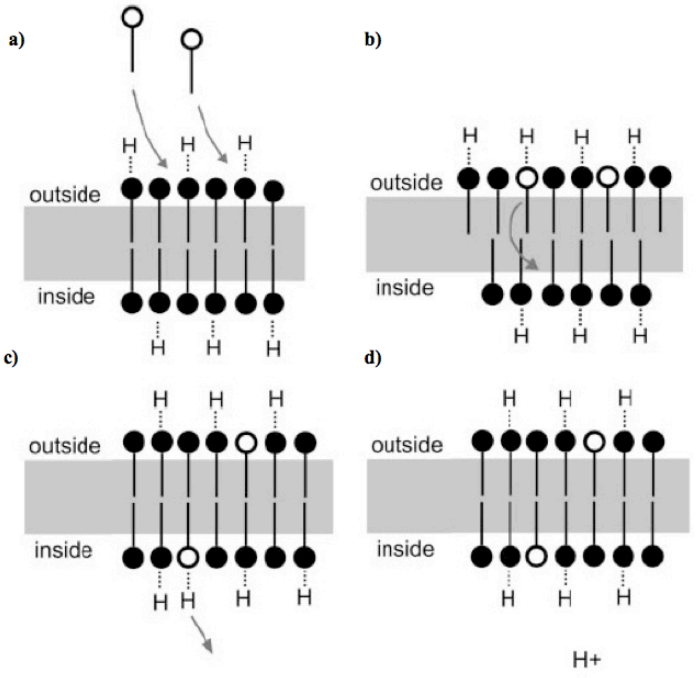


Figure 6.17 Model to explain the accumulation of pH gradient inside the vesicle lumen. The process proceeds from (a) through (d) over time to gradually accumulate H^+ ions inside the vesicle which lowers the internal pH.

The idea behind these experiments was pretty simple (Figure 6.16). If you add more fatty acids to vesicles capable of growing in size, and the added lipids have a slight positive charge on them due to the low external **pH** of the solution, will the liquid inside the abiotic vesicle develop a lower pH? The investigators hypothesized that their vesicles could abiotically accumulate a H^+ gradient inside the vesicles. They predicted that if they added new fatty acids to a solution containing many vesicles, the new hydrophobic molecules would quickly join the exterior of the vesicle to initiate growth. The new fatty acids would have a negatively charged head group but roughly half of the lipid head groups would pick up a H^+ ion from the buffered solution to maintain a neutral overall charge for the outer layer of the bilayer vesicle. As the outer layer of the vesicle accumulated the new fatty acids, half of them would flip into the inner layer to maintain mass balance in the bilayered membrane. The outer layer's overall neutrality could act

as a reservoir of positive charges that accumulate on the inside of the vesicle. Inside the vesicle, the new lipid molecules would equilibrate to the luminal pH which would result in approximately half of the new fatty acids releasing a proton into vesicle. An accumulation of H^+ ions would lower the pH and produce a proton gradient stored inside these primitive cells made of RNA and fatty acids. {Connection with Chapters 21 and 22 for proton gradient.} {Definition: **pH** is a negative log scale from 0 to 14 that describes how many hydrogen ions are in a solution.}

You may recall from previous chemistry courses that a H^+ ion is the same thing as a proton because H atoms are composed of a proton and an electron and H^+ has lost its only electron. pH is a negative log scale from 0 to 14 that describes how many hydrogen ions, H^+, are in a solution. A pH of 0 is a strong acid with many H^+ ions while a pH of 14 is very basic with no H^+ ions. Neutral pH is 7 with an intermediate H^+ ion concentration. Because the volume of the vesicle is tiny compared to exterior buffer, the pH is lowered inside the vesicle quickly while the external pH is unchanged in the much larger volume of extracellular liquid. This is analogous to filling a boat with lake water – the water level in the lake is essentially unaffected while the boat rapidly fills to the top.

Figure 6.16 depicted their hypothesis and this visualization led investigators to a second hypothesis. As the pH changes, the vesicle would exhibit an unexpected, or emergent, property. The lower internal pH would limit the maximum growth of the vesicle because the lower pH would generate new chemical and electrical gradients. Since abiotic, passive diffusion allows components to move down their concentration gradients, a second gradient might stop the movement. Subsequent fatty acids trying to join the vesicle would have to combat the newly accumulating positive charge and passive diffusion may not offer enough energy to overcome the new pH gradient. In other words, their non-living vesicles could accumulate a separation of charges and grow until the H^+ ion gradient is of sufficient strength to limit further growth. Regulated growth is sophisticated behavior for an abiotic sphere of lipids. As you know from your own life, each species has a characteristic size range and individuals cannot exceed these genetically determined sizes.

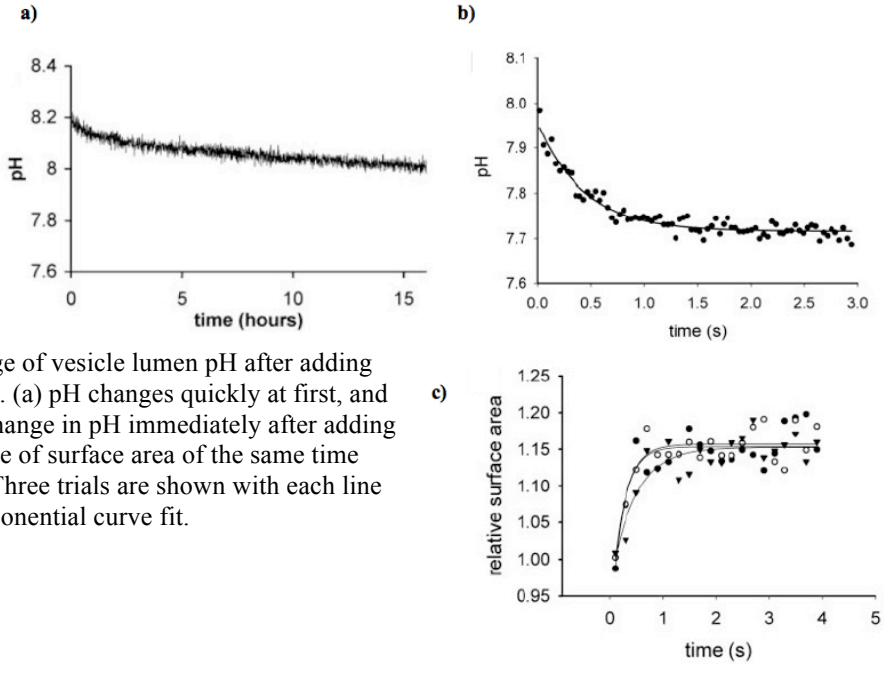


Figure 6.17 Change of vesicle lumen pH after adding micelles to vesicles. (a) pH changes quickly at first, and then slowly. (b) Change in pH immediately after adding micelles. (c) Change of surface area of the same time course as panel b. Three trials are shown with each line representing an exponential curve fit.

Now that they had formulated testable hypotheses, the investigators designed and performed their experiments (Figure 6.17). They added an amount of micelles equivalent to the amount of fatty acids already in the vesicles with interior and exterior initially at pH 8.2. They measured the internal pH after adding the micelles and saw an initial rapid drop in pH over the first 10 minutes, probably to be due to diffusible charged impurities in the lipids, but the pH drop continued slowly over several hours (Figure 6.17a). They measured the internal pH over the first few seconds after adding more micelles to some preexisting vesicles (Figure 6.17b). At the same time, they measured the change in surface area of the vesicles as the lipids incorporated into the membranes. The size of the vesicles changed at exactly the same rate as the pH dropped inside the vesicles (Figure 6.17c). Because pH is a negative $\log_{10}$ scale of H^+ ions (see BME 6.5), one unit of change in pH represents a 10^1, or ten-fold, change in H^+ ions, and a decrease in pH represents an increase in H^+ ions. Therefore, the decrease of 0.3 pH units in Figure 6.13 represents a $10^{0.3} \approx 2$, or two-fold, increase in H^+ ions, which corresponds to storage of 2.2×10^{-17} joules (unit of energy) for vesicles with an initial diameter of 100 nm. The energy storage in the form of a pH gradient represents about 12% efficiency, which compares favorably to the most famous energy pathway of photosynthesis (about 34% efficient). {Connection: Photosynthesis is described under the Big Idea of Energy, in Chapter 22.}

Bio-Math Exploration 6.5
Logarithms: The power of pH

- Concepts: Logarithms are exponents, or powers, of a selected base.
- Objective: Understand the pH scale.
- Required skills: Compute the value of a whole number raised to a negative or decimal power.

{Editorial Note: This BME is rated "green," meaning that most students should be able to master the techniques.}

Logarithms are a useful tool for presenting and interpreting quantities that vary over many orders of magnitude, or multiples of ten, such as 10^3, 10^7 and 10^{15}. The pH scale is a logarithmic scale, one of several examples of quantities measured on a log scale. Others include the Richter scale for earthquakes and decibels for sound.

The logarithm of a number is the power to which a base is raised to get the number. Many logarithms can be found without a calculator by using the power and base rules. For example:

$$\log_{10} 0.001 = \log_{10} 10^{-3} = -3$$
$$\log_{10} 100{,}000 = \log_{10} 10^{5} = 5$$

The first line is read "the log base 10 of 0.001 equals log base 10 of ten to the negative 3 power equals negative 3." Note that the base of a logarithm is written as a subscript. If the subscript is omitted, the base is assumed to be 10. We will use base 10 throughout this BME.

If the log of a unknown number is known, you can solve for the number by reversing the above process. For example, in the first equation above, knowing that the power is -3 and the base is 10 tells us that the original value was $10^{-3} = 0.001$. In symbols, if $\log_{10} x = y$, then $x = 10^y$.

Bio-Math Integrating Questions

BME IQ 6.5a: Find $\log_{10}$ 0.00001

BME IQ 6.5b: Find x if $\log_{10} x = -7$

A useful property of logarithms is that the log of the ratio of two values is the difference between the two logs. For example, the ratio between the two values above (100,000 / 0.001) is 10^8, and the log of the ratio is 8. This value is the same as the difference between the logs: 5 – (-3) = 8. In symbols, $\log_{10}$ 100,000/0.001 = $\log_{10}$ 100,000 - $\log_{10}$ 0.001 = 5 – (-3) = 8. In general, this "difference rule" is written as $\log_{10} (a/b) = \log_{10} a - \log_{10} b$.

Bio-Math Integrating Question

BME IQ 6.5c: If $\log_{10} a = -0.5$ and $\log_{10} b = -0.8$, find (a/b) in two different ways:

(1) by finding a and b and dividing the two numbers

(2) by using the difference rule to find $\log_{10} (a/b)$, then solving for (a/b).

The definition of pH is the negative of the power to which 10 must be raised to produce the hydrogen ion concentration (moles per liter). In symbols, you'd write $pH = -\log_{10}[H^+]$. The reason for the negative sign is just so that typical pH values end up being positive. For example, the hydrogen ion concentration in pure water is, on average, 10^{-7}M, corresponding to a pH of $-\log_{10}(10^{-7}) = 7$. The difference rule helps us see that the difference of 0.3 in the pH of vesicles (Figure 6.17b) corresponds to a $10^{0.3}$, or two-fold change in H^+ concentration.

---------- End of BME 6.5 ----------

Perhaps the most disappointing aspect of data in Figure 6.17b and c is the incredibly short time scale. If the goal is to demonstrate events that could lead to the formation of life, then these phenomena would need to persist longer than just a few seconds. Having the horizontal axis in seconds helps you discern the rate of membrane growth and pH gradient production, but it seems unlikely that a gradient that survives only a few seconds could evolve into a biological power source. However, the investigators did conduct a series of experiments to determine the maximum pH gradient they could produce and how long it would be sustained (Figure 6.17a). In experiments not show here, the investigators were able to measure a pH change that lasted at least 16 hours if they loaded the vesicles with an amino acid to help counterbalance the protons. They continued this research to fine tuned the internal and external pH differences and could store even more energy within vesicles. You should remember that in the 1950s Miller had shown amino acids could be produced abiotically so counterbalancing the luminal pH of a vesicle would have been feasible in ancient Earth.

Integrating Questions

19. Summarize how vesicles can harvest and store energy while they grow. Use Figures 6.13 – 6.18 to support your explanation. What do you think would happen to the stored energy when vesicles divided as you read in Section 6.3? What change in vesicle reproduction would need to happen before this stored energy could be used to support life?
20. Can you think of any ways a pH gradient across membranes can be used to perform any work? Is there a way to convert the potential energy of a pH gradient into chemical bonds? If you do not know these answers, you will learn how pH is a powerful potential energy source upon which all of life is dependent.

At this point, you have seen a lot of data demonstrating that abiotic factors can lead to the formation of a vesicle composed of a lipid bilayer capable of growing, competing with other vesicles, and generating potential energy in the form of a pH gradient. These same vesicles can enclose nucleic acids that accelerate growth of vesicles, and some RNA molecules can perform simple replication functions of polymerization of ribonucleic acids. Proton gradients in chloroplasts are responsible for photosynthesis at the base of the food chain. {Connections: Photosynthesis is covered in Chapter 22.} Mitochondria and bacteria use proton gradients for extracting energy from food sources. {Connections: Cellular respiration is covered in Chapter 21.} Abiotic vesicles that can generate proton gradients are one step closer to resembling living cells. As you have seen in previous chapters, science makes progress by offering rational explanations of natural phenomena that can be tested experimentally. The more times a hypothesis is supported by additional experimentation, the more we accept the hypothesis as the causal explanation. Eventually, a hypothesis can be described as a theory if it has been supported so many times that its status is considered secure, but never proven.

For skeptics who demand that a human be present to observe the first cell's formation or else they refuse to accept an abiotic origin of life, they will never be satisfied by scientific evidence. But if you consider the data from this chapter, what initially sounded outlandish can seem much more feasible – the formation of a living cell from physical and chemical processes. With the competing vesicle experiments, you have seen abiotic competition that resembles natural selection. Interestingly, those who reject evolutionary explanations for the origin of life on earth cannot produce a witness for alternative origins of life either – it is a logical impossibility. How could a multi-cellular human record the creation of the first cells when humans did not exist, much less know how to write? As a scientist, you must rely upon observable facts today and extrapolate the simplest natural way to explain the origin of life on earth. Once the first cell exists, all subsequent cells are easier to explain. However, you now have a new challenge to address. If the first cell was a prokaryote, how did the nucleus evolve? In the next section, you will consider genome sequence data to help you understand how eukaryotes came into being about 1.5 billion years ago.

6.5 How did the first nucleus come into being?

- Context: Once prokaryotes were abundant, the evolution of the first eukaryotes was sudden and not gradual.
- Major Themes: Organisms can be linked by lines of descent from common ancestry; and natural selection is a mechanism of evolution that accounts for adaptation.
- Bottom Line: Eukaryotes are the product of fused archaeal and bacterial cells.

When Charles Darwin was formulating his ideas about evolution, he drew a branching tree to indicate how one species can change and give rise to several related species, just as we use tree-like pedigrees to show family relationships among humans (Figure 6.18). In Darwin's tree, the letters A through M represent different species in a large genus. The dotted branched lines represent the offspring of A, each with some variation. Species B – D did not diverge and went extinct, as noted by their lines terminating before reaching the top of the page. Though his drawing was simple, Darwin's diagram had a profound effect on the way biologists think about evolution. Only now, about 150 years later, is this way of thinking slowly giving way to a new image. Figure 6.18 was Darwin's only figure in the entire 556 page edition of *On the Origin of Species* published in 1859, and it appeared as a fold-out between pages 116 and 117. As everyone knows, trees sprout from seeds and form a trunk with branches diverging in many directions. If life on Earth is drawn as a tree, then what was the seed and where is the trunk? The tree of life predisposed biologists to think of evolution as a series of events with gradual, linear progression towards the existing species. Whenever you go looking for evidence to support a particular idea, you can bias yourself to find what you were seeking even if alternatives exist.

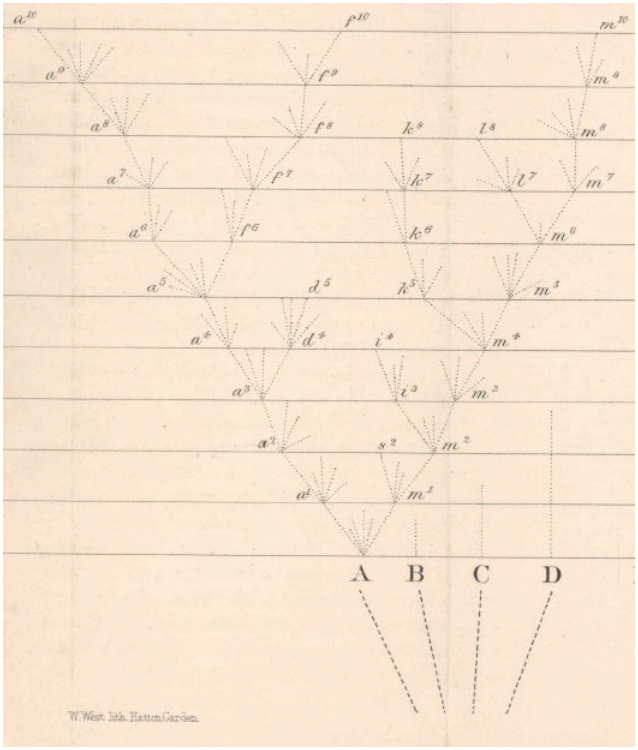


Figure 6.18 Tree of life from Darwin's *On the Origin of Species.* Species B, C and D went extinct, while species A continued to evolve over time (vertical axis) to give rise to three species listed at the top.

In Darwin's explanation of evolution, he stated, "Only those variations which are in some way profitable will be preserved or naturally selected." If a dotted line reaches a horizontal line representing 1000 generations, then enough variation would have accumulated that the offspring of A would look different from the original parent as denoted by a^1 and m^1. If a branch reaches the top most line where you can see a^{10}, f^{10}, and m^{10}, then line of evolving species avoided

extinction. Darwin recognized that his diagram looked too regular. He explained, "I do not suppose that the process ever goes on so regularly as is represented in the diagram."

Darwin knew it would be difficult to formulate rules in a field where the object of study changes over time. Biology is an unusual science in that it is very difficult to make broad statements without citing exceptions. For example, plants do not move, except algae that can swim; birds fly, except kiwi, ostrich, and penguins; lizards have four legs, except for the legless ones, *etc*. At the beginning of this chapter, you tried to define life and found it pretty difficult. Similarly, it is difficult to construct a set of rules that can distinguish all eukaryotes with nuclei from the other two **domains** of life - **bacteria** and **archaea**. The most common misconception cited as unique to eukaryotes is internal membrane-bound organelles while prokaryotes lack internal membranes (Figure 6.19). As you can see for yourself, the misconception is incorrect and some prokaryotes do have internal membranes. Figure 6.19a shows a very common prokaryote called cyanobacteria that have stacks of photosynthetic membranes similar to the structures found inside chloroplasts. Figure 6.19b is an electron micrograph of a large eubacterium called *Gemmata obscuriglobus* with its internal double-membrane (M) that is attached to the inner cell membrane and surrounds a fibrous genome (F) and a granular matrix (G). These prokaryotes with membrane-bound organelle reinforce the truism in biology – rules without exceptions are rare. {Definitions: All of life is divided into three **domains** or classifications: **eukaryotes** are cells with nuclei; **bacteria** are unicellular cells without nuclei that live in common places; **archaea** are similar to bacteria except they live in very harsh conditions such as heat and high salt.}

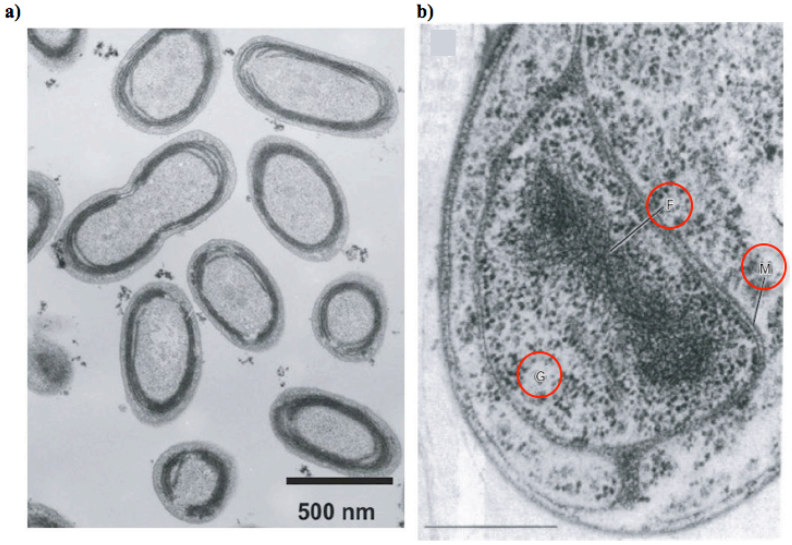


Figure 6.19 Electron micrographs of two prokaryotes that contain internal membranes. a) Photosynthetic bacteria contain internal membranes in think peripheral bands. b) The larger bacterium *Gemmata obscuriglobus* contains internal membranes similar to a nuclear envelop. Bar = 0.5 µm.

In *G. obscuriglobus* (Figure 6.19b), you observed a structure that looks transitional, as if the bacterium is trapped between its evolutionary past and a future species yet to come. Transitional species are sometimes referred to as "missing links" as if all of evolution were a chain and each species a link. If you were holding two halves of a chain, you would suspect a link is missing that used to exist. Similarly, lungfish have lungs and gills and are perceived to be links in the evolution of fish from water to land dwellers and the beginning of amphibians. However, you cannot realistically expect every transitional species to survive unchanged for millions of years

while other species continue to change and compete for limited resources. In the absence of living "missing links," you need alternative data to search for clues about the first eukaryotes and how they came into being from prokaryotes. In the absence of missing link of cellular fossils, you will analyze DNA evidence because genes contain clues about how our eukaryotic relatives evolved the first nucleus.

When DNA sequencing became readily available, biologists searched for the best gene to sequence in order to uncover the evolutionary relationships of all species to complete Darwin's tree of life. They reasoned that if each species evolved from a previous species, then it should be possible to discern the relationship between ancestral and derived species by comparing their DNA. Let's look at a simplistic example of five hypothetical species and their DNA data.

Species Names	DNA Sequences
S. singularis	GCAT**GCAT**GCAT
S. doubletus	GCAT**GCTT**GCAT
S. tripletus	GCAT**GTTT**GCAT
S. reversus	GCAT**TACG**GCAT
S. reversusdoublecus	GCAT**TACC**GCAT

You can see how the DNAs are similar to each other, indicating they are closely related, but the middle four bases have changed over time (Figure 6.20). It should be possible to use these sequences to determine which species evolved from which ancestor. If we use Darwin's tree imagery and focus on the central four bases, it is easier to understand.

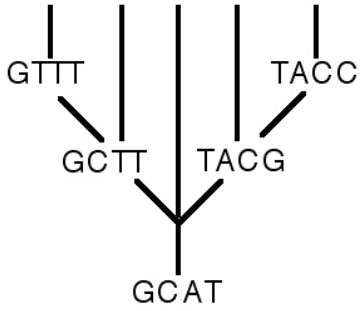


Figure 6.20 Evolutionary tree showing the changes in DNA over time, vertical dimension. The original species at the base is represented by the sequence GCAT. Only the four central bases are shown for clarity.

To read the 5 species tree in Figure 6.20, look at its base to find the oldest species/sequence GCAT. From this ancestral sequence, three events happened. On the middle branch, the original species remained unchanged, GCAT. On the right branch, the sequence was first inverted to TACG and later mutated at a single base to TACC. The left branch mutated A to T and later mutated C to T. Each branch in this overly simplified diagram indicates a new species evolving due to a change in the DNA. All 5 species evolved at different times in the past as indicated by the different lengths of the lines and the location of the branch points. You can see which species are more closely related because they share a common ancestor, and conversely, which species are more distantly related. You could imagine that producing an accurate tree would be much more difficult if you did not have data for every related species (Figure 6.21). When DNA analysis was conducted on highly conserved ribosomal genes from many more species, a more

complex tree emerged. The rDNA-based tree showed eukaryotes evolved from archaea with eukaryotes and archaea evolving from a common ancestor that also gave rise to eubacteria.

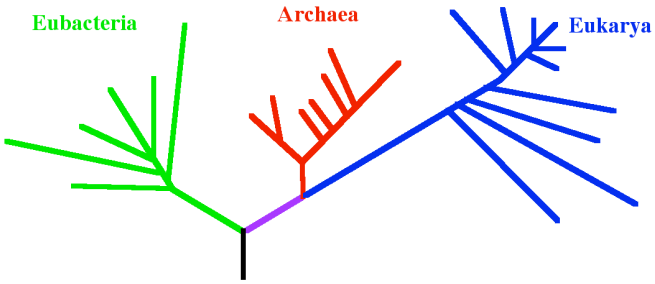


Figure 6.21 Evolutionary tree based only on one ribosomal gene from many species. Each domain is indicated, with Archaea and Eukarya sharing a common ancestor.

Integrating Questions

21. In an electron micrograph like those in Figure 6.19, each dark line is a phospholipid bilayer but you cannot see the two layers at these magnifications. How many membranes can you count on the outside of the cell in Figure 6.19b? Now search the internet for the term "nuclear envelop" and determine how many membranes surround the nucleus. What do prokaryotes and eukaryotic nuclei have in common?
22. Go to this ClustlW analysis web site (www.bio.davidson.edu/people/macampbell/111/ClustlW_search.html) and submit the rDNA gene sequences below to draw your own evolutionary tree. Can you deduce the evolutionary history of these organisms? Any surprises? Can you deduce the relationship among the mammals? What does your tree tell you about the traditional way of using a single gene to deduce evolutionary relationships of species?
23. Read these BLAST results (see UN 6.2) when some human protein sequences were used to find the best microbial matches. Can you deduce any patterns in this list, even though it is a short list?

Human Protein #	Protein Function	Protein Location	Best Match Domain
NP_001009	Translation	Cytoplasm/ rER	Archaea
NP_003185.1	Transcription Factor	Nucleus	Archaea
NP_001001937	ATP synthase	Mitochondria	Bacteria
NP_005521	Energy Harvesting	Mitochondria	Bacteria
NP_000393	Energy Harvesting	Cytoplasm	Bacteria
NP_004138	Cell Signaling	Cytoplasm	Archaea
NP_061816	Cytoskeleton	Cytoplasm	Bacteria

UN 6.2 Summary results from BLASTing human protein sequences against microbial sequences.

The nuclei of eukaryotic cells are surrounded by a pair of phospholipid bilayers, similar to the pair of membranes surrounding prokaryotes. Paired membranes also surround mitochondria and chloroplast, each of which also contain DNA. For now, you should just note the similarities as you continue learning about the origins of nuclei. When you performed the CLUSTLW evolutionary tree, you may have noticed some odd things. Most obviously, all the mammals were indistinguishable because their sequences did not vary for this highly conserved gene. The first lesson is that some genes are better at revealing distant relationships while other genes would be better at distinguishing closely related species. Therefore, it is nearly impossible to find one gene that could be used to determine the relationships of all species in the world. For more distantly

related species, the CLUSTLW tree was inconsistent with what we know about evolution. For example, the tree put birds closer to amphibians than reptiles even though we are certain that birds evolved from reptiles. It is good to remember that sequence-based evolutionary trees should not be over interpreted to indicate the relationships of species but limited to the conservation of sequences for particular genes.

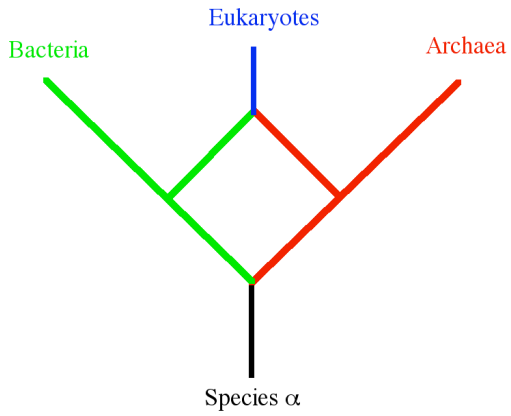


Figure 6.22 Diagram illustrating how a primitive species α could diverge into two branches which merged genomes later to produce a third branch Note the atypical tree structure with a loop in the middle.

From the summarized BLAST results, you can see that some human proteins are more closely related to Eubacteria while others are more closely related to Archaea. Energy production was closer to bacteria while transcription and translation were closer to archaea. Signaling and cytoskeleton are both cytoplasmic communication systems and they were split between the two microbial domains. With conflicting information similar to what you analyzed in the BLAST summary, scientists found it difficult to draw a single tree that accurately represented the evolutionary origins of all human proteins. Which gene would be the best gene to choose to help understand the evolutionary origin of eukaryotes? When investigators finally let go of their bias towards a Darwin-esque tree, they realized what had happened (Figure 6.22). This "tree" looks very different from a typical Darwinian tree. Humans and all eukaryotes contain genes from both Eubacteria and Archaea which diverged from a common ancestor long ago, species α. As you saw in the electron micrographs, nuclei and bacterial membranes are composed of two layers, each a phospholipid bilayer. The simplest solution to make sense of all these data was to propose that an archaeal cell and a eubacterial cell combined genomes with one cell engulfing the other. {Connection: The consequences of genome duplication is addressed in Chapter 23.} The original, abiotically produced cell was species alpha, which evolved and became a true prokaryote. Later, a subset of prokaryotes diverged and evolved into Archaea and Eubacteria. Millions of years later, a pair of Archaea and Eubacteria cells fused their genomes and gave rise to Eukarya (Figure 6.23).

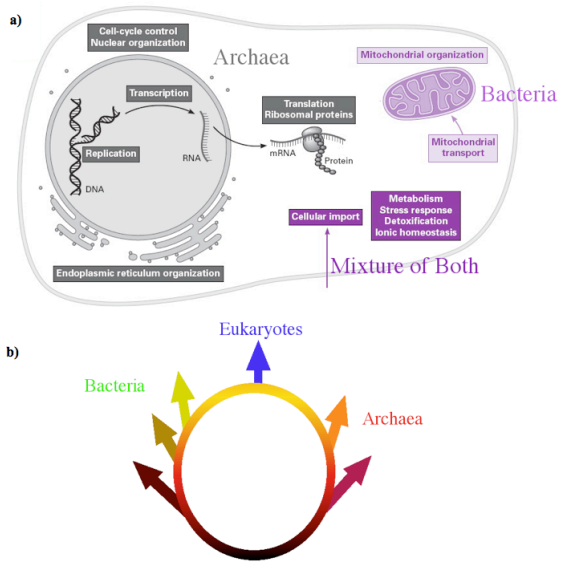


Figure 6.24 Origins of eukaryotes. a) Diagram showing the evolutionary origins of genes that originated in Archaea , Eubacteria , or a mixture of both. b) Ring of life diagram showing the intimate relationships between Archaea and Bacteria that fused genomes to produce Eukaryotes.

Over time, with duplicate genes in this new domain of life, sections of DNA were ejected to minimized the redundancy and be more efficient. It appears that whole pathways were retained rather than individual genes. For example, pathways related to DNA, transcription and translation all appear to be Archaea in origin. Energy harvesting genes were retained predominantly from the ancestral Eubacterial genome that contributed to the first eukaryote. {Connection with Chapter 12 for details of DNA ejection.} Some processes were more difficult to categorize as either Archaea or Eubacteria and are probably a mixture of both. It is time for us to stop using the metaphor "tree of life" and time to consider the "ring of life" which represents more accurately how the three domains evolved (Figure 6.23b). Humans tend to prefer tidy rules, black and white concepts and other clear distinctions. However, evolution does not follow a set plan nor is it linear in the adaptations that are selected over time. The "Ring Of Life" is probably more accurate that a tree of life, and so Darwin was off a bit on his imagery. However, a scientific model does not have to be correct in order to be useful. Darwin helped society see how species were related to each other through heredity and the species alive today provide a partial clue to the evolution of life. The genome sequences available today are similar to a puzzle with several pieces missing. It may be impossible to determine all the details, but with each experiment it gets easier understand the origin of life and general evolutionary relationships between species.

Conclusions

The definition of life at the beginning of this Chapter focused on three properties: transference of information from one generation to the next; containment in a 3D space; and change over time. By definition, humans did not record the origin of life on Earth, so we have to deduce what physical and chemical properties contributed to that amazing first cell. RNA is the most widely accepted original molecule that provided both the genetic information and the enzymatic capacity to replicate. Lipids can form membranes that can grow, divide, and store energy. Finally, genomes can merge and eject excess genetic baggage in order to adapt to changing environments through natural selection. Science is driven by data and scientists interpret nature after observation and experimentation (see ELSI 6.1). Chapter 6 provided you with the data necessary to understand how cells could have evolved and produced the three domains of life. Simply noting that some data are missing is insufficient to refute our understanding of evolution today. All data were missing until a scientist discovered them. To overturn a hypothesis, you need to analyze data that directly contradict your current interpretations. The data need to be reproducible and founded on scientific standards of credibility. Hearsay and secondhand accounts are unacceptable to be considered scientific data. However, if the origin of life hypotheses you discovered in this Chapter are repeatedly supported for many years, then they may be considered scientific theories similar to gravity, relativity, and evolution. It is OK to be wrong in science as long as you submit your ideas to peer review so others can test your claims. Even Nobel laureates have been wrong more times than right. We will never prove with mathematical certainty how life evolved on Earth the first time, but we can offer plausible explanations that are supported by data.

End Of Chapter Review Material

Review Questions

1. What are the five tenets of evolution?
2. Compare the use of the word "theory" when used in science vs every day usage.
3. Describe the experimental apparatus Miller used to synthesize amino acids in an abiotic environment.
4. What prediction did theoretical scientists make that was subsequently observed in nature?
5. How did investigators produce ribozymes that were more effective than ones found in nature?
6. Explain how lipids and fatty acids can be greedy.
7. List as many examples as you can of internal membranes NOT found in eukaryotic cells.
8. How do abiotic vesicles grow in size?
9. Describe how positive charges can accumulate inside a vesicle.
10. Give one or more examples of evolutionary missing links.
11. Compare the expressions "tree of life" and "ring of life".
12. What is the significance of amino acids being produced abiotically?
13. What is the significance of vesicles being able to grow and divide abiotically?
14. What are the implications for RNA molecules being trapped inside a vesicle?
15. Why is it important that RNA molecules can polymerize RNA as well as proteins can?
16. Why is capturing a pH gradient inside lipid vesicles a significant step forward in the origin of life?
17. Explain the relationship between structure and function in biological molecules.
18. How are surface area and volume related and why is this an important feature of living systems?
19. Explain why the concept of a missing link is both futile and informative when considering evolution.

Apply What you Know

1. Do you think biological molecules can change their shape and function over long periods of times and through many generations? Hypothesize how this can happen.
2. If a trait is adaptive today, will it always be adaptive? Speculate how a single trait can differ in its ability to confer selective advantage.
3. Where and when did humans evolve? What color was the skin on the earliest humans? Search for the human gene SLC24A5 to find out specific details.
4. Can the evolution of one species affect the evolution of another species?
5. Does one species remain constant over time, or does it change? If it changes, is it still the same species?

References

Evolution Overview

Charles Darwin. 1859. On the Origin of Species. Free online version at Darwin Online <http://darwin-online.org.uk/pdf/1859_Origin_F373.pdf>. Accessed 20 November, 2007.

Martin R. Fisk, Stephen J. Giovannoni, Ingunn H. Thorseth. 1998. Alteration of Oceanic Volcanic Glass: Textural Evidence of Microbial Activity. *Science*. 281: 978 – 980.

Giovanni Murtas. 2007. Question 7: Construction of a Semi-Synthetic Minimal Cell: A Model for Early Living Cells. *Orig Life Evol Biosph*. 37: 419 – 422.

Leslie E. Orgel. 2008. The Implausibility of Metabolic Cycles on the Prebiotic Earth. *PLoS Biology*. 6(1): e18. doi:10.1371/journal.pbio.0060018

Robert Shapiro. 2006. Small molecule interactions were central to the origin of life. *The Quarterly Review of Biology*. 81(2): 105 – 125.

Jack W Szostak, David P Bartel, P. Luigi Luisi. 2001. Synthesizing life. *Nature*. 409: 387 – 390.

Woese, Carl R., Otto Kandler, and Mark Wheelis. 1990. Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya. PNAS. Vol. 87:4576-4579.

First Organic Molecules

Jason P. Dworkin, David W. Deamer, Scott A. Sandford, and Louis J. Allamandola. 2001. Self-assembling amphiphilic molecules: Synthesis in simulated interstellar/precometary ices. *PNAS*. 98 (3): 815 – 819.

Adam P. Johnson, H James Cleaves, Jason P. Dworkin, Daniel P. Glavin, Antonio Lazcano, Jeffrey L. Bada. 2008. The Miller volcanic spark discharge experiment. Science. Vol. 322 pp. 404.

Stanley L. Miller. 1953. A Production of amino acids under possible primitive earth conditions. *Science*. 117: 528 – 529.

Stanley L. Miller. 1955. Production of some organic compounds under possible primitive earth conditions. *Journal of the American Chemical Society*. 77 (9): 2351 – 2361.

Stanely L. Miller. 1957. The mechanism of synthesis of amino acids by electric discharges. *Biochemica et Bophysica ACTA*. 23: 480 – 489.

Cyril Ponnamperuma and F. Woeller. 1967. a-Aminonitriles formed by an electric discharge through a mixture of anhydrous methane and ammonia. *Currents in Modern Biology*. 1: 156 – 158.

Rushdi AI, Simoneit BR. 2001. Lipid formation by aqueous Fischer-Tropsch-type synthesis over a temperature range of 100 to 400 degrees C. *Orig Life Evol Biosph*. Feb-Apr;31(1-2):103-18.

George U. Yuen and Keith A. Kvenvolden. 1973. Monocarboxylic Acids in Murray and Murchison Carbonaceous Meteorites. *Nature* 246, 301 – 303.

RNA Hypothesis

Michael D. Been, and Thomas R. Cech. 1998. RNA as an RNA Polymerase: Net Elongation of an RNA Primer Catalyzed by the *Tetrahymena* Ribozyme. *Science*. 239: 1412 – 1416.

Nicholas H. Bergman, Wendy K. Johnston, and David P. Bartel. 2000. Kinetic Framework for Ligation by an Efficient RNA Ligase Ribozyme. *Biochemistry*. 39: 3115 – 3123.

Eric H. Ekland and David P. Bartel. 1995. The secondary structure and sequence optimization of an RNA ligase ribozyme. *Nucleic Acids Research*. 23 (16): 3231 – 3238.

Eric H. Ekland, Jack W. Szostak, David P. Bartel. 1995. Structurally Complex and Highly Active RNA Ligases Derived from Random RNA Sequences. *Science*. 269: 364 – 370.

Peter Walde, Ayako Goto, Pierre-Alain Monnard, Michaela Wessicken, and Pier Luigi Luisist. 1994. Oparin's Reactions Revisited: Enzymatic Synthesis of Poly(adenylic acid) in Micelles and Self-Reproducing Vesicles. *Journal of the American Chemical Society*. 116: 7541 - 7547.

Vesicle Growth

Irene A Chen, Richard W Roberts, Jack W Szostak. 2004. The emergence of competition between model protocells. *Science*. 305: 1474-6.

Irene A. Chen and Jack W. Szostak. 2004a. Membrane growth can generate a transmembrane pH gradient in fatty acid vesicles. *PNAS*. 101 (21): 7965 – 7970.

Irene A. Chen and Jack W. Szostak. 2004b. A Kinetic Study of the Growth of Fatty Acid Vesicles. *Biophysical Journal*. 87: 988 – 998.

Irene A Chen, Kourosh Salehi-Ashtiani, Jack W Szostak. 2005. RNA catalysis in model protocell vesicles. *Journal of the American Chemical Society*. 127 (38): 13213-9.

Martin M. Hanczyc, Shelly M. Fujikawa, Jack W. Szostak. 2003. Experimental Models of Primitive Cellular Compartments: Encapsulation, Growth, and Division. *Science*. 302: 618 – 622.

Giovanni Murtas, Yutetsu Kuruma, Paolo Bianchini, Alberto Diaspro, Pier Luigi Luisi. 2007. Protein synthesis in liposomes with a minimal set of enzymes. *Biochemical and Biophysical Research Communications*. 363 (1): 12-17.

Eukaryote Evolution

Fuerst, John A. and Richard I. Webb. 1991. Membrane-bounded nucleoid in the eubacterium *Gemmata obscuriglobus*. *PNAS*. 88: 8184 – 8188.

Horiike, Tokumassa, Kazuo Hamada, et al. 2001. Origin of eukaryotic cell nuclei by symbiosis of Archaea in bacteria is revealed by homology-hit analysis. *Nature Cell Biology*. 3: 210–214.

Martin, William and T. Martin Embley. 2004. Early evolution comes full circle. *Nature*. 431: 134 – 137.

Rivera, Maria C. and James A. Lake. 2004. The ring of life provides evidence for a genome fusion origin of eukaryotes. *Nature*. 431: 152 – 155.

ELSI

Francis Collins. 2007. The Language of God: A Scientist Presents Evidence for Belief. Free Press. 294 pages.

John E. Jones III. 2005. Opinion from TAMMY KITZMILLER, *et al.* (Plaintiffs) v. DOVER AREA SCHOOL DISTRICT, *et al.* (Defendants). Case No. 04cv2688. US District Court for the Middle District of Pennsylvania. December 20, 2005.