Author's note to online readers: The three chapters on life are a bit less in depth than many chapters before and after. Actually, with these chapters I actually achieved about the right amount of detail. I will probably simplify the other chapters. However, this chapter needs some work. It is a bit dry, for one thing, and I still need to write the section on evolution.

UNIVERSAL FEATURES

UNIVERSALS OF LIFE

Around the time of the appearance of the first continents, about 3,700 million years ago, life had emerged in the form of simple single-celled organisms. Today we call such organisms **prokaryotes**. This name refers to the fact that they do not have a nucleus, like more recent cells do. *Karyo* is Greek for kernel or nucleus, so *prokaryote* means *before nuclei*. Prokaryotes are the simplest true life forms, but it would be a mistake to think of this beginning as inauspicious. Even today, prokaryotes are probably the most important class of living things on earth in terms of the role they play in the biosphere. We would not and could not exist without them. They are the ancestors of every living thing on earth today, and they helped shape the face of the Earth and atmosphere, as we will see. Because they have been around for so long, prokaryotes have diversified more than any other form of life. They have filled an unbelievable number of niches--they exist deep in the earth's crust, high in the atmosphere, at the bottoms of glaciers, and inside other organisms. They invented every kind of metabolism used by more recent organisms, and then some.

Because they were the first life forms, and the simplest, prokaryotes can tell us about life's essential features. They have all of the basic features of life, without many embellishments. Here are some of the major features they display; features that are common to every living thing on earth: 1. They are made mostly of six basic elements: Carbon, Hydrogen, Oxygen, Nitrogen, Sulfur, and Phosphorous (which can be abbreviated as CHONSP) 2. These elements combine into the four basic types of organic molecule common to all life, all of which based on carbon. These are: **carbohydrates**, **lipids**, **nucleic acids** (like DNA and RNA), and **amino acids**, which combine into **proteins**. 3. Prokaryotes are **cells**, which close themselves off from their environment, maintaining an internal organization different from their surroundings. 4. They maintain their organization by harnessing matter and energy from their environment, and exporting entropy and waste. In other words, they are open systems. 5. Prokaryotes also maintain their organization by retaining information, in the form of **genes**, which provide instructions for building and maintaining their bodies. Genes are encoded in DNA. 6. Living cells are able to reproduce; producing new individuals like themselves. They are able to do this because DNA, in addition to storing information, can replicate itself. Simple cells reproduce by copying their DNA and dividing in two, leaving a copy of each new cell. 7. Prokaryotes adapt over time. Those with more useful genes produce more copies of themselves than those with less useful genes. So, over time, successive generations accumulate useful genetic structures, which code for useful characteristics.

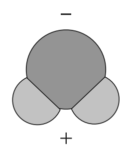
ELEMENTS

First of all, let's look at the elements that are common to all living things. Of the 92 naturally occurring elements, six happen to have properties that made them useful as building blocks for living things. These six are carbon, hydrogen, oxygen, nitrogen, phosphorus, and sulfur. These elements were, in a sense, selected from the rest of the elements--not selected in the sense that the rest were eliminated or cut back, but selected in the sense that they were the ones which would be incorporated into the processes and structures of life.

Now, of these six elements, the first three are the most common and the most important, so they are worth a closer look. Let's start with hydrogen, which is surely becoming familiar, since it has been with us from the beginning of our story. Hydrogen has one electron surrounding its nucleus, but it tends to gain another, to give it the less reactive noble gas shell of helium. This means that hydrogen often forms a single bond with another atom. Since many other atoms form bonds with several atoms at once, such atoms, like carbon and oxygen, may surround themselves with hydrogen atoms. Both carbon and oxygen are most stable when they have eight electrons in their outer shell, so that they resemble neon. Carbon has four electrons in its outer shell, so it tends to form four bonds. Oxygen has six electrons in its outer shell, so it tends to form two bonds.

Oxygen often fills its need for electrons by bonding with two hydrogen atoms to form water molecules. Each hydrogen shares two electrons with oxygen--one of its own and one from oxygen. This gives oxygen the two it needs, and each hydrogen the one they need. But the term "share" is not quite accurate, because a water molecule is not an equitable arrangement. Oxygen's eight protons cause it to greedily attract nearby electrons in order to fill its shell. Hydrogen, on the other hand, has just a lone proton, so it does not attract other electrons strongly. In fact, it tends to lose its grip on its own. So, in the water molecule, the electrons in each hydrogen atom are partly pulled away, toward the oxygen atom. This makes the oxygen side of the molecule negatively charged, and the hydrogen side positively charged (Figure .1). Water molecules, in other words, are **polar**; with opposite charges at each end. This makes groups of water molecules stick together; with the negatively charged end of one attracting the positively charged end of another. Many substances with charged atoms or molecules dissolve in water, because the water molecules pull the charged particles apart. In other words, water is an excellent solvent. This is one of the properties that makes water the most important substance for life on Earth. Life probably emerged in a watery environment full of dissolved materials. These materials, and water itself, were life's original matrix.

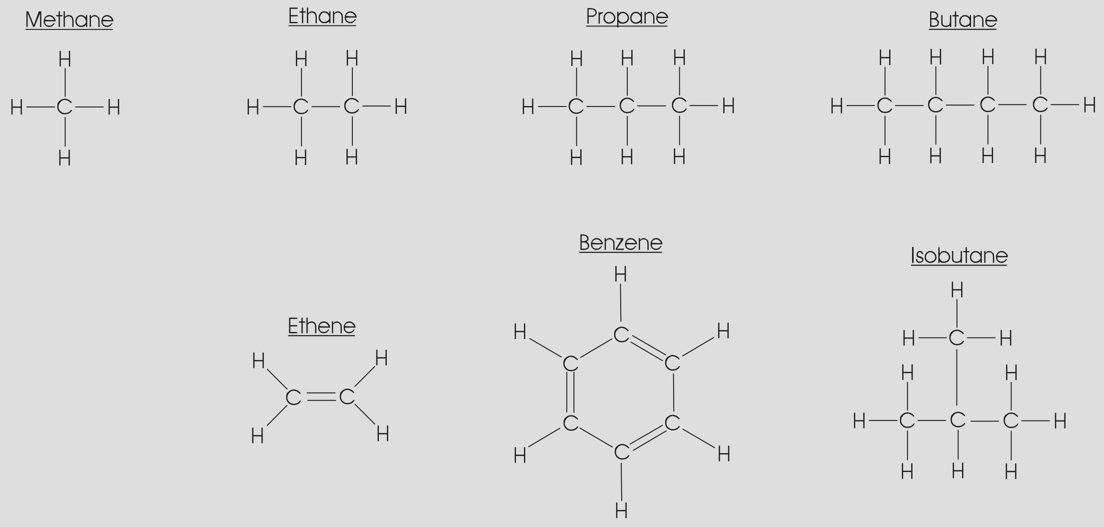
Figure .1 A Water Molecule



Hydrogen also shares its electrons with other atoms. Since its hold on electrons is much weaker than many other atoms, hydrogen tends to become polarized. Hydrogen atoms in molecules are often positively charged, which means they attract negative ends of other molecules; forming weak bonds called **hydrogen bonds**. Hydrogen bonds are extremely important for life, because they form bonds that can be easily broken if need be. The double helices of DNA, for example, are held together by hydrogen bonds.

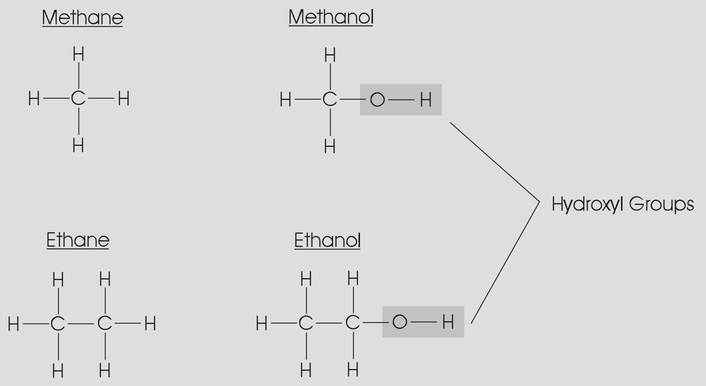
Carbon is a bit less greedy than oxygen. It needs four electrons to fill its outer shell, so it

Figure .2 Some Simple Hydrocarbons


Methane
Ethane
Propane
Butane
Ethene
Benzene
Isobutane

tends to form four bonds. But since it has fewer protons than oxygen, it does not pull as hard, and it is not as likely to form polar molecules. Carbon's mild, sharing nature makes it the most versatile of elements, serving as the backbone for a practically endless variety of molecules. Large molecules based on carbon are called **organic molecules**, and the study of organic molecules constitutes an entire field called *organic chemistry*. The diversity of carbon-based molecules is based on single family of molecules, called **hydrocarbons.** As the name suggests, hydrocarbons consist entirely of carbon and hydrogen. The carbon atoms link together to form a backbone, with hydrogen atoms bonded around the outside. As Figure .2 shows, some hydrocarbons, such as butane, are long strings. Some have branching carbon backbones, such as isobutane. Hydrocarbons like benzene loop back on themselves to form rings. Whatever the basic structure of hydrocarbons, their most important feature is that they are modular- they can be fitted with various attachments that change their characteristics. These attachments are various molecules called **functional groups**. A *hydroxyl group*, for example, contains an oxygen and a hydrogen. It can bond to a hydrocarbon backbone, forming an alcohol, such as methanol (wood alcohol, which is toxic), or ethanol (grain alcohol, which is intoxicating). Life is based on large, complex molecules which are formed from hydrocarbons modified by functional groups. Now let's look at these molecules.

Figure .3 Functional Groups Change the Properties of Hydrocarbons



MOLECULES

Life is based on the fact that the diverse elements, especially carbon, can combine into even more diverse molecules, some of which have properties that are extremely useful. Of these diverse molecules, there are four main types on which life absolutely depends: *carbohydrates*, *lipids*, *proteins*, and *nucleic acids*. The first three make up the bulk of every living body, which is why they are so prominent in the nutrition information boxes on food packages. The fourth category, nucleic acids, are less bulky, but just as important. They include DNA and RNA, the molecules that carry the genetic code.

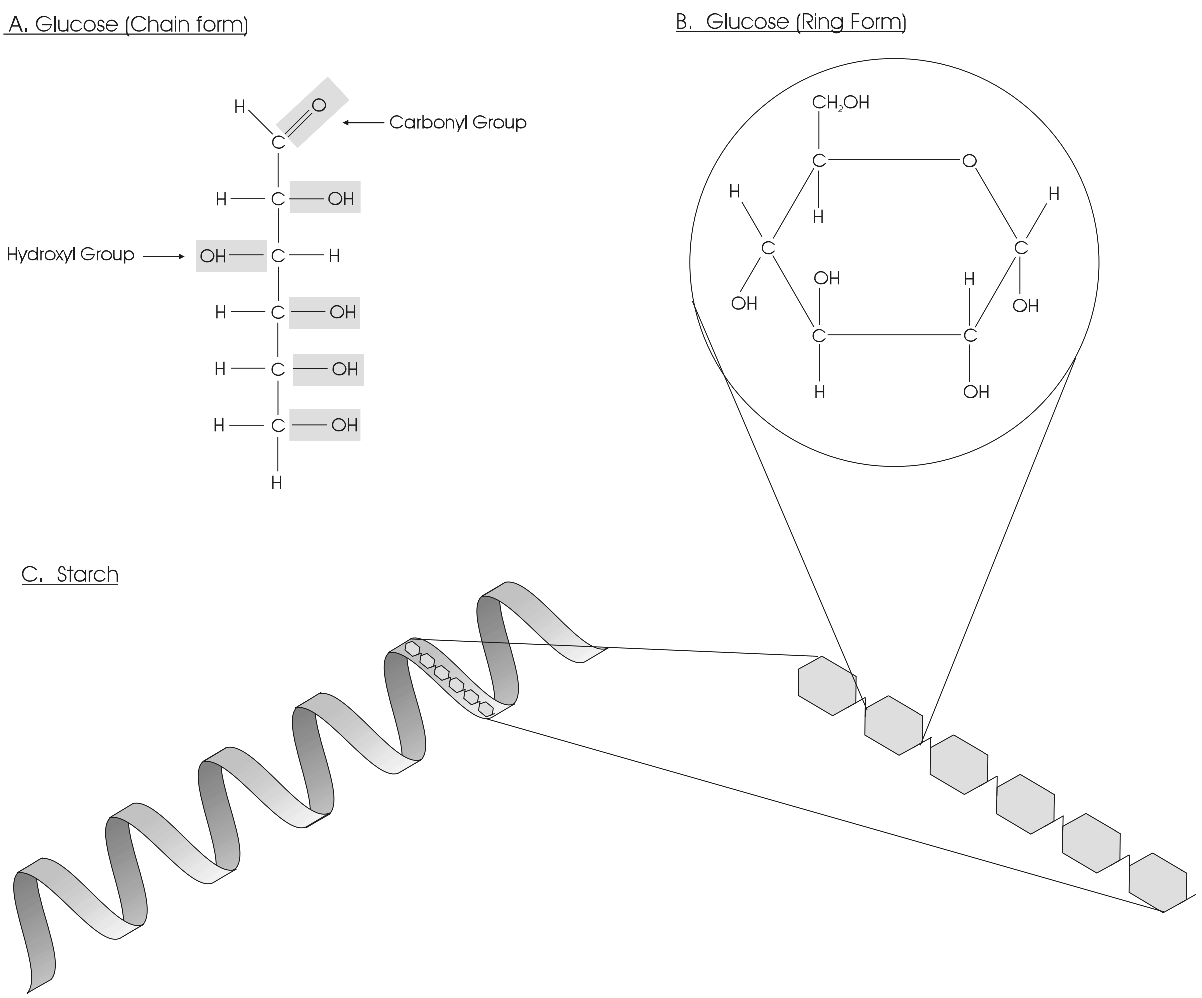
CARBOHYDRATES

The word *carbohydrate* is a familiar one; bringing to mind images of bread, candy bars, potatoes, or beer. In fact, the word is so familiar that people don't stop to take a look at it. *Carbo-hydrate* is a word laden with meaning. *Carbo* indicates that carbohydrates contain carbon. The meaning of the second part is suggested by words like *hydraulic* and *hydration--hydro* means water. And generally speaking, carbohydrates are a combination of carbon and water. Actually, the water is often broken into its constituents--two hydrogen atoms for every oxygen atom. But it is useful to think of carbohydrates as combining water and carbon, because they tend to have one carbon atom for every water molecule. For example, a molecule of **glucose**, a simple carbohydrate (Figure .4 A), can be disassembled to make 6 carbon atoms and 6 water molecules. Since 6 water molecules contain 12 hydrogen atoms and 6 oxygen atoms, the chemical formula for glucose is $C_6H_{12}O_6$. This is an important fact to remember, because glucose is central to the nutrition of most living things, as we will see.

Now, Figure .4 A shows glucose as a linear chain. Most of the time, however, one end of the chain links to another to form a ring, as in Figure .4 B. Now, a single glucose molecule is a kind of carbohydrate called a **monosaccharide**, or a simple sugar. There are also many other simple sugars, including *fructose*, a sugar in fruit, and *galactose*, a sugar in milk. Simple sugars can themselves be parts of larger carbohydrate molecules. For example, a glucose and a fructose can combine to form a molecule of *sucrose*, or table sugar. Such double sugars are called **disaccharides**. Simple sugars can also form the links of larger chains, called **polysaccharides**. For example, many glucoses can link into a long chain to form starch (Figure .4 C).

Carbohydrates have two basic functions in living things--fuel storage and structural reinforcement. Plants store glucose in the form of starch, which they stock away in pockets within their cells. Some plants even have modified organs for storing starch. A potato, for example, is a root that specializes in starch storage. This starch is broken down when glucose is needed for energy. Animals store glucose in another polysaccharide, an extensively branched

Figure .4 Carbohydrate Structure. Glucose combining into Starch



chain called *glycogen*. *Cellulose* is a structural carbohydrate. It consists of long chains of glucose banded together to form fibers, which weave together to form sturdy walls around the cells of plants, helping plants stay rigid even though they have no bones. Many boneless animals, such as insects, have external skeletons made with another complex carbohydrate chain called *chitin*.

PROTEINS: CELLULAR HARDWARE

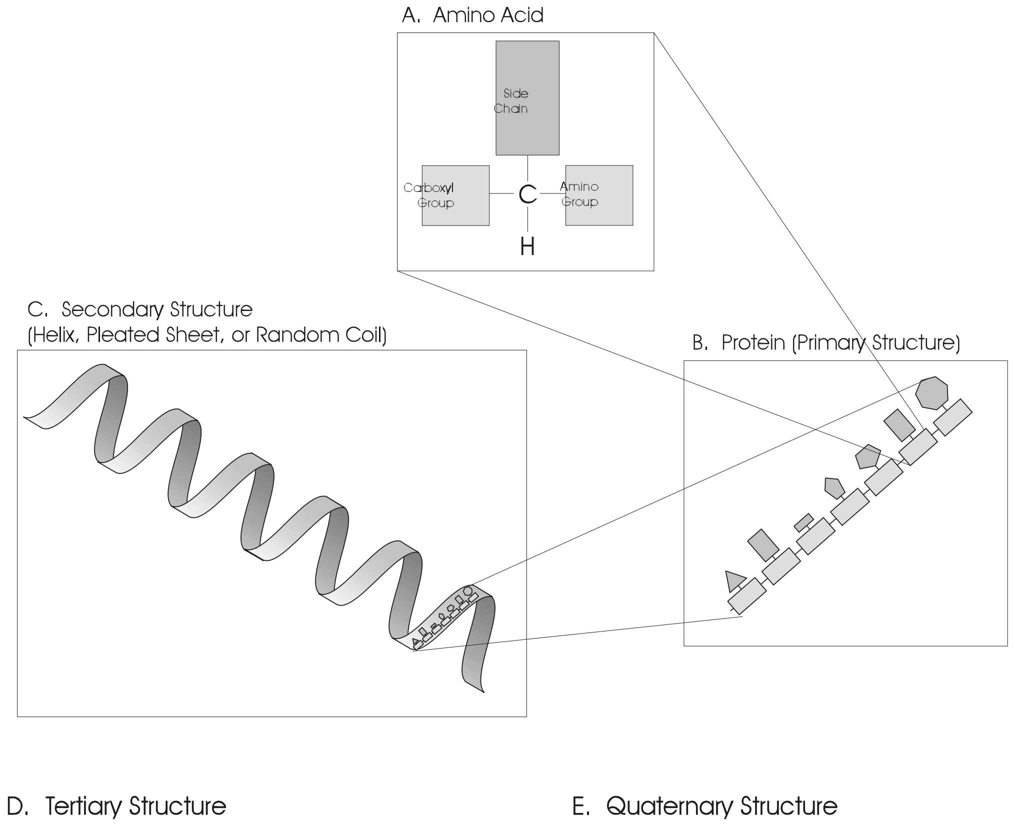
Proteins take the modular, hierarchical form a step farther than carbohydrates. Like polysaccharides, proteins are long chains whose links are smaller molecules. The links in protein chains are called **amino acids**. Whereas most carbohydrates in living things are composed of just a handful of simple sugars, proteins may be composed of 20 different amino acids. Amino acids are a larger alphabet than simple sugars, and the sentences they form--proteins--are consequently quite diverse.

Like the different letters in an alphabet, the different amino acids have different properties. Each amino acid is built around a central carbon atom, as in Figure .5 A. At one of the four bonding sites is a hydrogen atom. Two more sites are occupied by functional groups--an *amino group* and a *carboxyl group*. This structure is the same in every amino acid. What makes amino acids different is the functional group across from the hydrogen atom, called a **side chain**. Some side chains are small and simple, while others are complex, with functional groups of their own. Depending on their functional groups, some amino acids are attracted to water, while some are repelled. Some are acidic, and some more basic.

Amino acids are diverse, but the real diversity appears when they combine into larger wholes. Like many whole systems, proteins are fewer in number, but far greater in kind than their parts. Amino acids combine into proteins by linking the carboxyl groups on one side to the amino groups on the other (Figure .5 B). This leaves the side chains sticking out on one side. Each protein is composed of a different sequence of amino acids, whose properties combine in different ways to help determine more complex, emergent properties. Since 20 amino acids can be combined in countless different sequences, the number of possible proteins is enormous, especially since many of them are composed of thousands of amino acids.

But the emergent properties of proteins don't end with the specific sequence of amino acids. The particular sequence of amino acids is a protein's **primary structure**. After a protein is formed, bonds between hydrogen atoms cause the protein chain to fold. Parts of the chain may spiral around to form a helix (Figure .5 C), while other parts may bunch up into a shape called a *pleated sheet*. These emergent shapes are the protein's **secondary structure**. And it doesn't stop there. The entire chain of amino acids, with its helical and pleated sections, folds in on itself based on bonds between different side chains. The particular shape that results is called the **tertiary structure** (Figure .5 D) **(Awaiting permission for image)**. Finally, two or more separate chains may come together to form an even higher-order structure, called the **quaternary structure** (Figure .5 E)**(Awaiting permission for image)**.

Figure .5 Levels of Structure in Proteins



Proteins, then, are complex wholes whose diverse properties depend on many layers of structural organization. Individual atoms, such as hydrogen and carbon, contribute their characteristic features. At a higher level, different amino acids contribute their individual properties, which depend on their side chains. These combine into long chains whose emergent properties depend on their particular sequence as well as the properties of individual amino acids. One emergent property of these sequences is the way the chains fold up; into primary, secondary and tertiary structures. The particular shapes that arise from this folding provide a further level of emergent function.

Most proteins can do a particular job because they have a particular shape, which fits in with other molecules in a particular way. For example, many proteins are **enzymes**, which means that they speed up reactions between other substances. Figure .5 **(awaiting permission for image)** shows how this works with the protein sucrase, which catalyzes the breakdown of sucrose (a disaccharide) into its components, glucose and fructose. Sucrose is attracted by electric charges in the sucrase molecule, causing it to lodge in a perfectly-shaped site in the larger molecule. In fact, the sucrase actually changes shape to hold the sucrose tightly. A water molecule is also attracted to a nearby site, where it breaks the bonds between glucose and fructose in a process called *hydrolysis*. The fructose and glucose are then released, ready for further processing, and another sucrose enters the site on the sucrase molecule.

Most enzymes work in similar ways, by bringing molecules close enough together to facilitate the rearrangement of their atoms. Enzymes display a close coupling of shape and function. But not all proteins are enzymes; some take different shapes, and serve different functions. Collagen, for example, is a *structural protein*. It consists of three helices wrapped around each other like strands in a rope. This gives it a fibrous structure that is ideal as a component of tendons, ligaments, and other connective tissue. Silk is also a structural protein, as is *keratin*, the main component of fingernails and hair. Some proteins, such as insulin, serve as hormones--chemical messengers within the body. Actin and myosin are *contractile proteins*, whose changes in shape make our muscles move. Proteins like hemoglobin carry substances from place to place in the body. Some proteins are lodged in cell membranes, to help move molecules into or out of the cell. Some of these are simply tubes, through which molecules move from low to high concentrations, while others actively change shape, spitting molecules into or out of the cell. In short, proteins are a cell's cogs, gears, motors, templates, funnels, and more, each performing the function for which it is shaped.

NUCLEIC ACIDS

DNA AND RNA

If proteins are the hardware of the cell, then nucleic acids store the software. DNA, or *deoxyribonucleic acid*, is the medium that carries genetic information. This information is used for two main purposes. First, it orchestrates the building and maintenance of living bodies. All organisms, from single-celled bacteria, to cheetahs, to giant redwoods, are organized and administered by the instructions in each of their cells. Second, DNA is capable of replicating, or making copies of itself, which is what allows all living things to reproduce.

Actually, there are two kinds of nucleic acids in living cells--DNA and **RNA**, or *ribonucleic acid.* To see how these molecules function, let's look at how they are put together. The basic unit of both DNA and RNA are molecules called **nucleotides**. Nucleotides are analogous to amino acids and simple sugars--they are molecular beads which join together to form polymer strings. Like amino acids and simple sugars, nucleotides are themselves wholes composed of simpler parts. A nucleotide has three basic parts (Figure .6 A) **(Still working on image)**. On one end is a functional group called a phosphate group, consisting of a phosphorous atom surrounded by four oxygen atoms. This is attached to a simple sugar in the middle. In RNA, that sugar is ribose, and in DNA that sugar is deoxyribose- which is ribose minus one oxygen atom. The third molecule in a nucleotide is called a **base**. A nucleotide of DNA could have four different bases--**guanine**, **adenine**, **thymine**, or **cytosine**. These are often called by their first letters- G, A, T, and C. (A science fiction movie about discrimination based on genetics was cleverly titled *GATTACA*.) In RNA, thymine is replaced with **uracil**. As Figure .6 B shows, thymine, cytosine and uracil have a central, hexagon-shaped ring of atoms. Adenine and guanine are larger, with a pentagon shaped ring attached to the hexagon.

Now we can begin to see how nucleotides fit together to form nucleic acids. In adjacent nucleotides phosphate groups bind with sugars, as in Figure .6 C. This forms a backbone linking many, many nucleotides (often millions), where the phosphate groups alternate with sugars in a repeating pattern. This pattern of sugars and phosphate groups is the same wherever we look on the nucleic acid. Lots of order, but not much organization. The pattern of bases is far more interesting. Each nucleotide along the backbone might have any one of the four bases characteristic of that nucleic acid. So, different strands of nucleic acids can have very different sequences of bases, even though they have the same backbone. It is this flexible sequence of bases that allows nucleic acids to carry information.

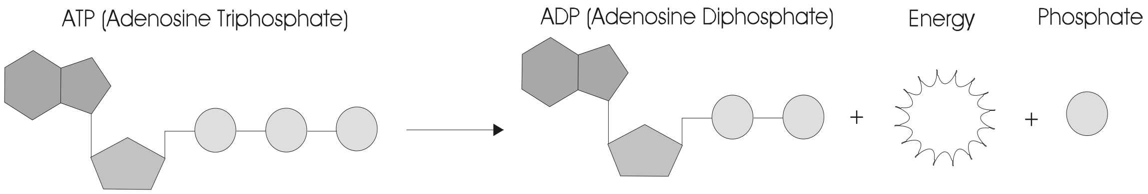
Now, each of the bases can form weak bonds with the bases of other nucleotides. Guanine bonds with cytosine, and adenine bonds with thymine (or uracil in RNA). So, each base on one side is mirrored by a complementary base on the other side (Figure .6 D). This means that if you know the sequence on one side, you can figure out the sequence on the other. For example, if a sequence on one side is G,A,T,T,A,C,A, the sequence on the other will be

C,T,A,A,T,G,T. In DNA, the nucleotides always occur in pairs, with a backbone on either side. Each backbone winds itself into a helix, giving DNA its famous double helix shape. RNA usually has just a single strand, although it often attracts complementary bases for short times, as we will see.

ATP: A MOLECULAR POWER PACK

One of the nucleotides, adenine, is closely related to another molecule that plays a very different, but equally vital, role. Like all nucleotides, adenine has one phosphate group attached to its central sugar. If two more phosphate groups are added, so that it has a chain of three of them, the molecule is called **adenine triphosphate**, which is usually abbreviated as **ATP**. The phosphate groups, especially the last two, are not very strongly bonded, which makes ATP a rather high-energy system. It isn't very stable, and that is why it is so useful. ATP easily loses its third phosphate group, becoming **adenine diphospate**, or **ADP**. This releases energy, which, to simplify a bit, is used to drive cellular work. Mainly, it provides the energy for proteins to do their thing; changing shape to do some simple task, like squeezing other molecules together or pumping ions out of the cell. Each change in the protein is powered by the energy released from a single ATP molecule. Most proteins perform their task many times each second, propelled by a series of ATP molecules. Its as though a machine on an assembly line were powered by a bin full of firecrackers. These firecrackers are rechargeable, however. In every cell, hordes of ATP molecules are constantly being broken apart, and then put back together again, ready to drive another reaction.

Figure .7 ATP Loses a Phosphate, becoming ADP and releasing Energy



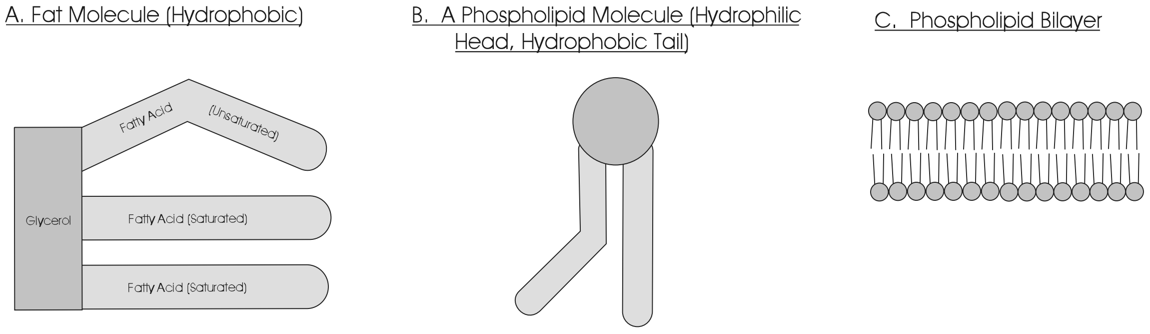
LIPIDS

The third class of molecules essential to all life are lipids. Wax, fat, grease, many oils, margarine, and cholesterol--all these things are made of lipids. Unlike carbohydrates and proteins, lipids do not share a single molecular architecture. They are a motley assortment with many different molecular shapes. But as the list of lipids above suggests, they do share one major feature--they don't dissolve in water. It is often said that oil and water don't mix. A more sweeping statement is that lipids and water don't mix. Water beads up on a newly waxed car, oil beads up or floats on water, and fat separates out of cold soup to form that icky yellow overlay.

FATS

Lipids also differ from other classes of biological molecules in that they do not form long polymers. They are more compact molecules. Let's look at two of these: **fats** and **phospholipids**. Like other complex organic molecules, these have a hierarchical structure--they are composed of smaller molecules, which are in turn composed of carbon backbones and functional groups. As Figure .8 A shows, a fat consists of a molecule called glycerol attached to three molecules called fatty acids. Fatty acids consist mainly of a long string of carbon and hydrogen molecules. Fatty acids are what makes fats *hydrophobic*, or insoluble in water. Actually, it is more accurate to say that water is fat-phobic. Water molecules attract each other because they are polar--they are positively charged on one end and negatively charged on the

Figure .8 Lipids


A. Fat Molecule (Hydrophobic)

B. A Phospholipid Molecule (Hydrophilic Head, Hydrophobic Tail)

C. Phospholipid Bilayer

other end. Fatty acids, on the other hand, are not polar. Water molecules simply cling together so tightly that they squeeze out the fats. The more charged and polarized constituents exclude the more neutral ones. It's a bit like politics.

Besides being hydrophobic, the other chief characteristic of fats is that they can store a great deal of energy--twice as much as carbohydrates. This is why animals, which need to move around, keep long-term energy stores locked away as fat, instead of relying on bulky carbohydrates like starch, as plants do. Of course, plants do have fats, but they tend to be different than animal fats. Plant fats, like olive oil or corn oil, are **unsaturated**, which means that their hydrocarbon chains are not saturated with hydrogen atoms. Not all of the spots for hydrogen atoms are filled. In those spots in the chain, carbon atoms have no hydrogen to bond to, so they bond together in double bonds, which creates a kink in the chain. This means that unsaturated fats have kinks in their tails, and **saturated fats** do not. These kinks keep unsaturated fats from packing together closely enough to form a solid, so they tend be liquids (oils) at room temperature. Saturated fats, on the other hand, can pack tightly together to form solids, like butter or lard. Saturated fats are associated with clogging of the arteries, so unsaturated vegetable fats are generally healthier than animal fats. But beware. Many vegetable fats are "partially hydrogenated' so that they will be solid at room temperature. This is another way of saying they have been "re-saturated", and made less healthy. (Transfats)

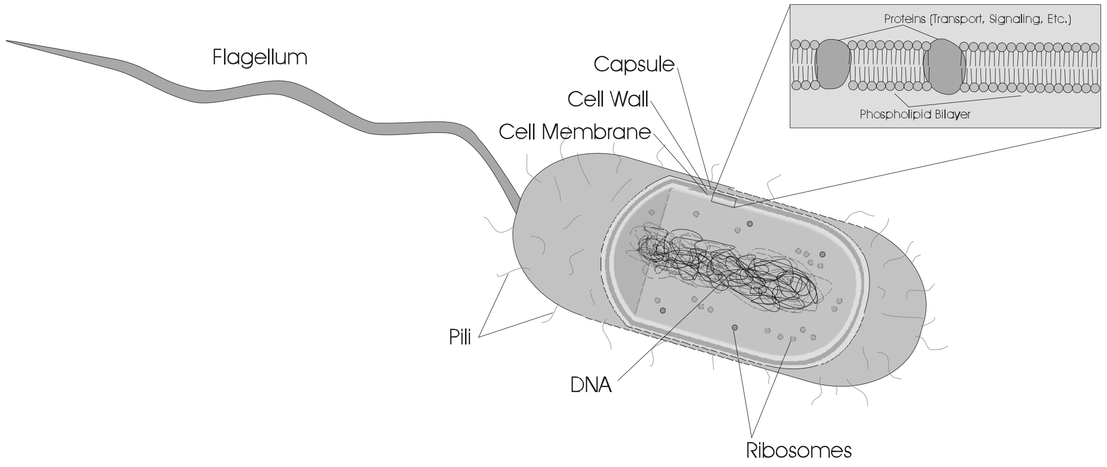
PHOSPHOLIPIDS

Another important class of lipids is the phospholipids (Figure .8 B). Phospholipids consist of a phosphate head and two fatty acid tails. This arrangement is what gives phospholipids their most important property: their tails are hydophobic, moving away from water, while their heads are hydrophilic, moving toward water. Cell membranes are composed mainly of two layers of phospolipids, called a phospholipid bilayer (Figure .8 C). The heads of one layer point toward the watery interior of a cell, while the heads of the other layer point toward the (often watery) exterior. The tails point toward each other in the middle. Since the phospholipids in each layer can flow past one another, the membrane behaves a bit like a liquid and a bit like a solid. It is strong but flexible, and permeable to some things but closed to others. These properties make it an effective boundary between the living cell and the world outside.

THE SIMPLEST CELLS

Now that we have looked at the major building blocks of life, let's begin to put the parts together into the simplest whole that can truly be considered living--a simple cell. Prokaryotes, the single-celled organisms that are still the most numerous living things on Earth, represent the simplest known cells. Figure .9 shows a generic prokaryote. A cell membrane composed of two layers of phospholipids helps separate the inside of the cell from the world outside. Many prokaryotes have a rigid cell wall around their cell membrane, and a gel-like coating around that. Inside the cell is a soup of organic molecules and tiny cellular organs. The organs are called **organelles**, and the soup as a whole is called **cytoplasm**. Prokaryotes have few organelles, however, except for numerous **ribosomes**, which help translate genes into proteins. The DNA is scrunched up into a tangle in the middle of the cell, because if uncoiled it would be far longer than the cell itself.

Figure .9 Cutaway View of a Generalized Prokaryotic Cell



METABOLISM

The simplest cell is a marvel of biochemical organization. Outside the membrane and cell wall is the wide, unpredictable world. Inside is a very well-orchestrated set of chemical reactions and complex molecules, working together in an emergent process called life. Cells create and maintain this high level of organization, beating the flow of entropy by exporting it elsewhere. They are open systems: patterns of organization that persist as matter and energy flow through. All living things take in high-grade energy and expel low-grade energy. This allows them to take in matter and rearrange it into a form that they can use. Some is recycled, and some is expelled as high-entropy waste.

The set of chemical reactions by which an organism constructs and maintains itself in the face of entropy is called **metabolism**. Metabolism has two parts. **Catabolism** is the breaking apart of molecules in order to extract energy or simpler building blocks. **Anabolism** is the building up of complex molecules from simpler ones. In other words, metabolism is the processing of two basic resources: energy and raw materials. How this processing is accomplished is surprisingly variable, especially among prokaryotes. But there are some universal features. For example, every known living thing uses ATP as its energy currency. An essential process in the metabolism of any organism harnessing energy for immediate use by converting ADP into ATP. Some get this energy from light, and some from breaking down chemicals, but they all convert it into the universal currency of ATP. This probably means ATP has been used since the very beginnings of life.

Another feature almost as universal as the use of ATP is that simple sugars, especially glucose, are central to the metabolism of most organisms. Plants, for example, use light to generate ATP, and then use ATP to construct sugar molecules out of carbon dioxide. Plants, and many other organisms, also break down sugars in order to extract energy and raw materials. In both processes, simple sugars are converted, in a series of steps, into other molecules. Many biology classes teach this process by tracing how glucose in transformed through this sequence. Glucose is the simple sugar that most organisms use internally; storing it as complex carbohydrates like starch or glycogen, and then retrieving it when needed. But while sugars like glucose are central to the metabolic process, they are not the entire metabolic process. Far from it.

When we eat, we eat a complex mix of lipids, carbohydrates, proteins, and other molecules, not pure glucose. Most of these things can also be broken down for energy and raw materials. Metabolism is a complex network, where many molecules are converted to many others, often in a series of steps called pathways. Figure .10 (awaiting permission for image) illustrates this, by showing the most central reactions the metabolism of an animal cell. Each dot represents a molecule, and each line represents a transformation from one molecule to another, which is usually catalyzed by an enzyme that exists just for that purpose. The lines and dots in bold show the central sequence of sugar metabolism. Other pathways show how other kinds of molecules are built up or broken down in side processes.

Like many networks, this one is capable of a great deal of self-regulation; seeking a stable, dynamic steady state. Many of the pathways work in both directions, so a balance of chemicals may be maintained through a simple equilibrium. For example, if a complex molecule formed by a catalyst becomes too numerous, and its molecular parts become too scarce, the anabolic (building up) process may go into reverse, becoming a catabolic (breaking down) process. This can happen as a spontaneous restoration of equilibrium, when the concentration of wholes vs. parts settles into a stable ratio. It can also happen as a result of a feedback process. If too much of a certain molecule is produced, it may trigger the formation of another molecule that counteracts it, perhaps by blocking the sites on the enzyme that catalyzes it. In short, the internal business of the cell is a complex network, full of feedback loops and equilibria, relying partly on bottom up, self-organizing processes. But a cell is too complex to rely entirely on such bottom-up processes. It also relies on top down control, in the form of a set of instructions that has built up through adaptation. And that brings us back to DNA and the genetic code.

THE GENETIC PROGRAM

REPLICATION: COPYING THE PROGRAM

In 1953, James Watson and Francis Crick published a 900-word letter in the scientific journal *Nature*, explaining the structure of DNA. DNA had been known for years to be the molecule that carries hereditary information, but before its structure was explained, nobody knew how DNA was able to copy itself. In one of the most brashly understated remarks in history, Watson and Crick wrote that “It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material". Once the structure of DNA was known, the way it copies itself was clear. It simply unzips down the middle, splitting each base pair **(Figure .11 Awaiting permission for image)**. If there are other nucleotides around, the bases in each half of the double helix will then attract their complementary bases. Each separate helical backbone thus serves as a template for constructing the complementary helix, and where there was one double helix, now there are two. The DNA has replicated, and when the cell divides, each new cell keeps a copy.

Of course, it is not nearly this simple. A great many complex processes have to occur to pull the double helix apart in just the right way at just the right time. Still, the basic mechanism of replication is clear. A much harder question, one that still has not been entirely answered, is how the genetic information encoded in DNA is used to build and maintain organisms. It is a fiendishly complex process. The first step in this process, however, is straightforward. DNA and RNA cooperate to make proteins.

BUILDING PROTEINS: RUNNING THE PROGRAM

We have discussed the structure of DNA, but we have not yet defined what a **gene** is. A gene is a sequence of DNA that acts as a code for a sequence of amino acids strung together as a protein. Each of the 20 amino acids used in proteins is specified by a particular string of three bases in the DNA sequence. The sequence GCC, for example, codes for an amino acid called alanine. Each three-letter sequence that codes for an amino acid is called a **codon**. Codons are like words whose meaning is an amino acid[1]. A string of these words forms a sentence whose meaning is a certain protein.

TRANSCRIPTION: FROM DNA TO RNA

Now, how is this code deciphered and put into effect in the creation of real proteins?

---

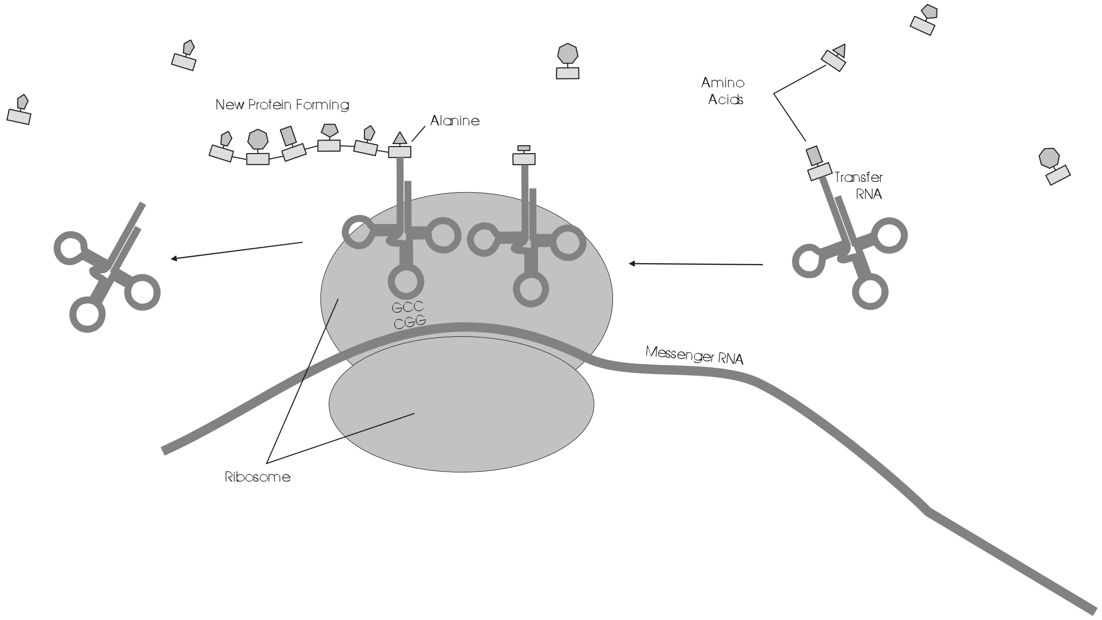
[1] Actually, most amino acids can be specified by more than one codon. Four different three-letter sequences code for alanine, for example. These four codons are like synonyms, all with the same meaning.

Here's where RNA comes in. The first step is a process called **transcription**. With the help of enzymes, a section of DNA in the region of a gene unzips (Figure .12 **Awaiting permission for image**). The side that carries the code attracts nucleotides of RNA. For example, the codon GCC, which codes for alanine, attracts the complementary bases CGG. When the sequence of DNA is transcribed into the complementary sequence of RNA (again with the help of enzymes), the RNA detaches. This strand is called **messenger RNA,** often abreviated *mRNA*. The mRNA carries its message to a **ribosome**, a piece of cellular machinery made out of proteins and RNA, through which the mRNA feeds like tape through a recording head.

TRANSLATION: FROM RNA TO PROTEIN

The ribosome is the site of the next step, translation (Figure .13). Out in the cellular fluid are other kinds of RNA called **transfer RNA**, or tRNA. At one end of each transfer RNA is a sequence of three bases that is complementary to a three-base sequence on the messenger RNA. At the other end is the amino acid coded for by that sequence. So, it is transfer RNA that connects each codon to its amino acid. For example, if the messenger RNA feeding through the

Figure .13 Translation From RNA to Protein (Highly Schematic)



ribosome comes to the sequence for alanine (CGG, the complement of GCC, the codon in the DNA), a transfer RNA with the sequence GCC will be attracted to it, bringing along its alanine molecule. The next three letters of RNA will attract another t-RNA, carrying another amino acid. These amino acids are joined together. As this process is repeated again and again, the DNA sequence transcribed by the messenger RNA is translated into a long string of amino acids--a protein. When the protein is finished, it begins to fold into its secondary, tertiary, and perhaps quaternary structures. A protein has been constructed, based on the instructions in DNA, and it is ready to go off and do the cellular work for which it is shaped.

PROTEINS, THE GENOME, AND THE LIFE OF THE CELL

Of course, this is only a skeletal description of the real process. In real life, each step is a complex orchestration of processes mediated by enzymes and other protein machinery. One set of enzymes helps unzip DNA and transcribe m-RNA. Another edits out meaningless strands of RNA called *introns*. Others help fold the newborn protein into the right shape. Naked DNA and RNA can't do much by themselves--they require an army of molecular helpers. And so far, we have only described the way proteins are built. The genetic code does much more than simply construct proteins and send them on their way. The information in DNA also oversees the working of the cell (or system of cells) through birth, life, reproduction, and death.

The construction of proteins is the beginning of a cascade of processes that combine into the smooth functioning of the cell. The genome responds to conditions within the cell through feedback, turning different genes on and off as different proteins are needed. Example. So, while the genome is the ultimate top-down authority, the cell's functioning is a complex network of feedback loops between genes, proteins, other molecules, and other factors, all influencing each other. Some of these processes run by themselves, based on self-organization or even basic laws of chemistry. Some rely on explicit instructions from the genome. It is a complex system combining many of the patterns we have been discussing in this book: adaptation, networks, feedback, pre-existing regularities and laws of physics, top-down and bottom-up control. This is typical of life--it uses whatever works. And while it may be hard to unravel and understand, most of the time it works quite well.

EVOLUTION