# 2

# Molecules and Origins of Life

Ecology seeks to explain patterns in species abundance and distributions. Evolutionary ecology considers how the ecology of an organism affects its evolution and more specifically the transfer of genes from one generation to the next. The genealogy of an organism is the complete uninterrupted linking of beginnings to the present. Life on this planet had some beginning. All organisms that are alive today can trace their genealogy to some time when there was no life. This is an extremely disconcerting concept. It conjures up a scenario in which, at one moment, the chemicals of life are floating around and, at the next moment, these same chemicals are working together inside some primordial life form that can reproduce and make copies of itself. That concept takes a leap of faith, but life did originate. It did begin somehow and somewhere. Although the pathways which led to life may or may not still be among us, it seems appropriate that we discuss some of the numerous models of the origin of life. These discussions are essential in helping us understand the basis of life and to understand what natural selection can act on that has resulted in the diversity of life we have before us today.

All early life was microbial but not necessarily prokaryotic. By this I mean that the complex prokaryotic organisms that microbiologists study today were probably not the first living things. Prokaryotes are much simpler than many higher organisms but nevertheless the diversity of metabolism and function within the prokaryotes is immense and very complicated and complex. The first organisms had to be much simpler and easier to construct than even the simplest microbes of today.

All models on the origin of life are based on there being a distribution of various molecules that either are required for life or promote the existence of life. It is very important that microbial ecologists understand the chemicals of life as a precursor to understanding life. The absence of some of these molecules would have made life either nonexistent or perhaps very different from the life forms we see today. After our discussion of the molecules of life, we will discuss various models that seek to describe the origin of life.

These models and concepts are at best guesses about how life originated, because no one knows the exact conditions that prevailed at the site where life first began. However, based on our knowledge of the chemicals of life, these models give us insight into primordial conditions and show us some of the interactions that had to occur before a replicating life form could arise.

# Chemistry of Life

## Water

The first molecule that is important in the evolution of life is water. This is truly a remarkable molecule that has properties that allow life to form and survive. All life, regardless of where it is found, requires liquid water. Fortunately, water is the most common liquid found on the earth, and it has several properties that have allowed life to evolve and to perpetuate.

Water is made up of only two elements: one atom of oxygen and two atoms of hydrogen. Each atom of hydrogen is linked to the oxygen atom by a *covalent bond.* Because the molecule has equal numbers of electrons and protons, it is on paper neutral. However, oxygen, because of its mass, has a greater attraction for electrons than the hydrogen that makes the region around the hydrogen atoms slightly positive and the region around the oxygen slightly negative. The water molecule is *polar*. When two water molecules come close to each other, the oppositely charged regions form a weak bond known as a *hydrogen bond.* Each water molecule can form four hydrogen bonds. It is the hydrogen bond (hydrogen bonds are not found exclusively in water) that gives the water molecule properties that are important to life. In the following paragraphs, we discuss some of the properties of water affected by hydrogen bonding.

*Surface tension* is brought about by the *cohesion* of water molecules. Water holds together at its surface because of its attraction to itself. Water striders and other insects can walk across the surface of water. Many people have placed a needle on the surface of water and observed the float. Surface tension or the hydrogen bonds between the water molecules prevents the needle from sinking. Water forms spherical drops because of surface tension. As soon as a drop of water breaks free from a stream of water it immediately forms a spherical shape. Because water is a polar molecule it will be attracted to any charged surface. The ability of water to wet a surface is due to *adhesion* or the attraction of two different substances. In this case it is the attraction of water molecules to either positive or negatively charged surfaces.

Water can move up tiny tubes or spaces because of cohesion and adhesion. This movement is termed capillary action and it is important in the movement of water in plants but also in the movement of water through various inorganic and organic matrices like soils. The water is attracted to the surface of the tube by adhesion and pulls other water molecules up through cohesion.

Hydrogen bonds in water are also responsible for the high *specific heat* of water. The specific heat is the amount of heat required to raise the temperature of a substance a given amount. It requires one calorie of heat to raise the temperature of 1 cubic centimeter (=1 gram) of water by 1°C. There are not many other substances with as high a specific heat as water. This is important biologically because it means it takes a lot of heat to raise the temperature of water. This means that organisms living in water will experience fairly constant temperatures especially in large bodies of water like oceans and lakes. Also large organisms are composed of significant amounts of water, changes in the temperature of those organisms will be a function of the water content and surface to volume ratios of the organism.

Water also has interesting properties as it freezes. As water gets colder the density increases to around 4°C. At that temperature water molecules are moving so slowly

that they can fo
molecules have
making frozen w
property becaus
would sink to th
of water with sc

Water is also a
in living organis
food. The polari
ecules that are in
molecules and d

## Biological El

Six elements ma
nitrogen, phosph
rally occurring e
among the most
advantage of otl
such a few elem
ments requires a
all are able to fo
means that the
ments (except hy
possible number
Nature has prod
has synthesized

Carbon in pa
figurations and s
of organic comp
pounds can also
other elements a
are the products

Every living c
ganic. Many of t
all living tissue i
principally comp
$Ca^{2+}$, $Mg^{2+}$, $K^{+}$,
The positively ch

Microorganisr
ecules. Living th
including lipids,
Each of these m
isms, is the prod
of the simplest l
lism of these m
originating).

that they can form the maximum number of four hydrogen bonds but to do so the molecules have to move apart. This moving apart continues as the water freezes making frozen water less dense than the water at 4°C. This is an incredibly important property because it means that ice floats. If ice were denser than liquid water the ice would sink to the bottom of a body of water, accumulate, and eventually fill the body of water with solid ice. This is a condition that would make life difficult.

Water is also a good solvent because of its polar nature. Many important substances in living organisms are in solutions. These substances include gases, nutrients, and food. The polarity of water facilitates the separation of ionic molecules. Many molecules that are important to life, such as sugars, are also polar, and they attract water molecules and dissolve in it, making their distribution possible.

## Biological Elements

Six elements make up nearly 99% of all living tissue. These six elements are carbon, nitrogen, phosphorous, hydrogen, oxygen, and sulfur. Considering there are 92 naturally occurring elements, this number seems quite small. These six elements are not among the most abundant elements at the earth's surface. Life did not evolve to take advantage of other extremely abundant elements like silica. Why is life made up of such a few elements and why these six? Part of the answer is that each of these elements requires an addition of electrons to complete the outer energy levels, and they all are able to form covalent bonds. These elements are also relatively small, and that means that the bonds they form result in tight stable molecules. Each of these elements (except hydrogen) is also able to form bonds with more than one atom. The possible number of combinations among these six elements is immense and diverse. Nature has produced thousands of compounds based on these six elements and man has synthesized many more.

Carbon in particular is able to bond with other carbon atoms in a variety of configurations and sizes. This diversity in form results in diversity in function. Millions of organic compounds (i.e., contain carbon) have been identified. These organic compounds can also include hydrogen, oxygen, nitrogen, sulfur, phosphorus, and many other elements and salts. All of these complex compounds made by living organisms are the products of specific genes and/or other gene products such as enzymes.

Every living cell contains a variety of molecules including both organic and inorganic. Many of these molecules are present as charged ions. Considering that 99% of all living tissue is made up of six elements the remaining 1% of biological mass is principally composed of inorganic ions. The positively charged ions are mostly $Na^+$, $Ca^{2+}$, $Mg^{2+}$, $K^+$, and $Fe^{2+}$, and the negatively charged ions are $SO_4^{2-}$, $PO_4^{3-}$, and $Cl^-$. The positively charged ions are important in many enzymatic reactions and functions.

Microorganisms have evolved ways to capture and sequester these important molecules. Living things have many other types of molecules that are essential for life including lipids, fatty acids, proteins, numerous enzymes, and vitamins to name a few. Each of these molecules of life, while fundamental to the survival of living organisms, is the product of cell metabolism. Any model that seeks to describe the origin of the simplest life form must be based on chemical interactions before the metabolism of these molecules (i.e., the constituent molecules need to exist prior to life originating).

# Early Atmosphere and the Beginnings of Life

Geological evidence suggests that the atmospheric chemistry of today is very different from that found 4 billion years ago. Most of the building block chemicals of life including oxygen, nitrogen, hydrogen, sulfur, and carbon were present in this earliest of atmospheres but not in the forms that much of life today uses. The actual forms of each of these elements is still being debated but it is generally accepted that free oxygen ($O_2$) was basically unavailable and that the other elements that make up life were present in the atmosphere or in the waters that covered the earth in simple molecules. Much of what follows is modified from Casti (1989).

In 1922, the Russian biochemist A. I. Oparin came up with the first testable hypothesis about the conditions and events that preceded the first living things. Many people had ideas of how life came to be found on this earth, but few of these ideas were testable. Scientific hypotheses must be testable and generate questions that can be answered through experiments. Oparin reasoned that the primordial atmosphere was reducing rather than oxidizing and as such was filled with methane, ammonia, hydrogen and water vapor. If energy in the form of lightening, volcanic heat, ultraviolet light and other sources of radiation were introduced into mixtures of these gases, he hypothesized that organic molecules would form. Not just any organic molecules but amino and nucleic acids, the basic building blocks of living organisms. Given time and the absence of oxygen these organic compounds could accumulate in the oceans until sufficiently concentrated that the first living organisms could form. In England, J. B. S. Haldane a few years later formulated a similar hypothesis and called the resulting mixture a "hot dilute soup" which has been modernized into the Primordial Soup Theory.

The Oparin-Haldane hypothesis was presented in the early to mid 1920s but was not seriously tested until the 1950s. Why? Remember that spontaneous generation had been laid to rest by Pasteur in 1864. Embodied in this hypothesis was the essence of spontaneous generation. Life could arise from inorganic materials that had been converted into organic compounds. Who wanted to go down that path again and face a scientific audience that had finally accepted that life came from life? In the early 1950s, a graduate student at the University of Chicago was willing to test the hypothesis. This young student was Stanley Miller. Miller was a student of Harold Urey, who had argued in a much more convincing and thorough manner than Oparin that the earth's early atmosphere was reducing and a good place to synthesize the molecules of life.

## Miller Flask Experiment

The basic experimental design Miller used to test the hypothesis is illustrated in Figure 2.1. The primordial atmosphere was simulated using ammonia ($NH_4$), methane ($CH_4$), hydrogen ($H_2$), and water vapor ($H_2O$). An electrode attached to a power supply supplied the energy required. The sparks created by the electrode were meant to mimic lightning. The whole mixture was cycled through a cooling tube that condensed the gases and resulted in a simulated rainfall. The water was slightly heated to promote evaporation. After one week, Miller analyzed the water and found significant amounts

---

**Figure 2.1** Diagramn
experiments.

of various amino
blocks of protein
studies have been
water, the starting
ferent organic mol
strate that the bas
The studies remai
molecules came in
of the process and
place in the early
Once the chemi
the sum of the ch
death, the molecu
and yet there is
through these am
into being.

## Which Molec

The origin of life
to catalyze or code

s

day is very differ-
chemicals of life
ent in this earliest
The actual forms
accepted that free
that make up life
th in simple mol-

the first testable
ing things. Many
ew of these ideas
uestions that can
rdial atmosphere
thane, ammonia,
canic heat, ultra-
nixtures of these
any organic mol-
living organisms.
ld accumulate in
isms could form.
r hypothesis and
dernized into the

id 1920s but was
s generation had
as the essence of
at had been con-
again and face a
n the early 1950s,
hypothesis. This
l Urey, who had
n that the earth's
olecules of life.

strated in Figure
, methane ($CH_4$),
ower supply sup-
meant to mimic
t condensed the
ated to promote
nificant amounts

---

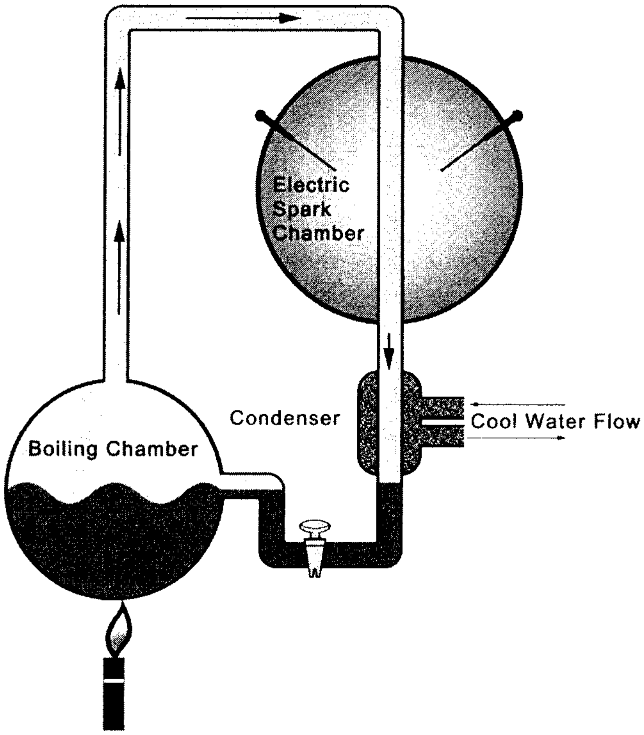


**Figure 2.1** Diagrammatic representation of the experimental apparatus used in the Miller-Urey experiments.

of various amino acids—specifically glycine and alanine, two of the basic building blocks of proteins. Since this experiment was initially conducted, numerous other studies have been performed altering the sources of energy, the temperatures of the water, the starting mixture of gases and each experiment has produced slightly different organic molecules. These studies are extremely important because they demonstrate that the basic molecules of life could originate through totally abiotic means. The studies remain theoretical because we are unable to show that this is the way these molecules came into existence. There is a tremendous amount of evidence in support of the process and most biochemists agree that something akin to these reactions took place in the early earth's history.

Once the chemicals of life were formed we still did not have life. Life is more than the sum of the chemicals that make it up. Every organism that dies still contains, at death, the molecules of life in pretty much the same proportions and concentrations and yet there is not life. Therefore just because certain molecules can be formed through these amazing processes does not mean that we understand how life came into being.

## Which Molecule Came First?

The origin of life is a chicken-or-egg type of problem. Certain molecules are needed to catalyze or code for the formation of other molecules, which are catalyzed or coded for by the other molecules. Let us examine this problem in more detail, because evolutionary ecology is based on these molecules, their synthesis, and regulation.

DNA is divided into short sections that code for certain specific proteins or code for the activation or inhibition of various chemical activities within the cell. These sections are often referred to as *structural* or *regulatory genes* and together they contain the information necessary to construct the organism. This information can be passed on to offspring. Protein synthesis occurs at specialized combinations of RNA and proteins called *ribosomes*. Each group of three base pairs that has been *transcribed* is called a *codon*. Each codon is associated with one of the twenty amino acids that make up the proteins of life or they code for a stop signal that ends translation. Because there are four different bases that make up DNA and there are three bases in each codon, there are a possible 64 different codes. However, there are only 20 amino acids used by all living things. Why are there not more amino acids? The genetic code contains some redundancy, a fact readily observed in Figure 2.2. Some amino acids are coded for by several codons, whereas others have only a single codon. This redundancy prevents serious conformational problems because the simple base substitutions often result in the same amino acids being coded.

We can simplify an extremely beautiful and complex process down to the following schematic designated the *Central Dogma of Molecular Biology* by Francis Crick, the co-discoverer of DNA.

$$\text{DNA} \xrightarrow{\textit{transcription}} \text{RNA} \xrightarrow{\textit{translation}} \text{Protein}$$

As seen from the diagram protein synthesis appears to be one directional (i.e., from the genetic information of the DNA to the formation of proteins). Genes code only for proteins, a fact that many students fail to comprehend. Genes code only for proteins!

Do proteins ever code for DNA? There are examples of RNA being back coded (reverse transcription) but there are no examples of transfer of information from proteins back to either genetic molecule (i.e., RNA or DNA).

| | *A* | *G* | *C* | *U* | |
|---|---|---|---|---|---|
| *A* | *lys* | *glu* | *gln* | *stop* | A |
| | *lys* | *glu* | *gln* | *stop* | G |
| | *asn* | *asp* | *his* | *tyr* | C |
| | *asn* | *asp* | *his* | *tyr* | U |
| *G* | *arg* | *gly* | *arg* | *stop* | A |
| | *arg* | *gly* | *arg* | *trp* | G |
| | *ser* | *gly* | *arg* | *cys* | C |
| | *ser* | *gly* | *arg* | *cys* | U |
| *C* | *thr* | *ala* | *pro* | *ser* | A |
| | *thr* | *ala* | *pro* | *ser* | G |
| | *thr* | *ala* | *pro* | *ser* | C |
| | *thr* | *ala* | *pro* | *phe* | U |
| *U* | *ile* | *val* | *leu* | *leu* | A |
| | *met* | *val* | *leu* | *leu* | G |
| | *ile* | *val* | *leu* | *phe* | C |
| | *ile* | *val* | *leu* | *phe* | U |

**Figure 2.2** The genetic code consists of four nucleotides that code for 21 amino acids and stop frames.

---

A one-time cc
oned by Lamar
concept any alte
subsequent gene
to feed on highe
neck genes to of
produce them ai

We bring up t
how life evolvec
lowed by protei
evolved first and
stood today. Tl
examine each m

Casti (1989) i
meet:

1. There is life
2. All life opera
3. Life is very c

Any theory a
earth allowed o
of facts 2 and 3
isms were able t
isms. All living
set of basic cons
and snails, colle
life.

The gene-pro
living organism
lated into the ap
translated withc
addition to the g
be solved before
tional hurdles:

1. *Genetic code*
nucleic acids.
by all living t
2. *Chirality.* All
acids and nuc
tical in comp
determines fu
in Miller-typ
cules. Interes
only left-han
of the respec

A one-time competing concept to evolution by natural selection was that championed by Lamarck and known as the *inheritance of acquired characteristics*. In this concept any alteration or mutation in any cell of an organism could be passed on to subsequent generations. The classic example is of a giraffe constantly reaching its neck to feed on higher leaves, which results in a longer neck, and the passing on of the long neck genes to offspring. The proteins that build longer necks can alter the genes that produce them and the reproductive cells that pass the trait on to offspring.

We bring up this concept at this time because there are two competing models of how life evolved. One of these models suggests that genes or DNA came first, followed by proteins. The other model proposes just the opposite (i.e., that proteins evolved first and then genes). The first model is the basis of natural selection as understood today. The second model is very reminiscent of lamarckian evolution. We examine each model as a means of understanding evolutionary ecology.

Casti (1989) identifies three evident facts about life that a theory of its origin must meet:

1. There is life on earth.
2. All life operates according to the same basic mechanisms.
3. Life is very complicated.

Any theory about the origin of life must show how the early conditions of the earth allowed or promoted the rise of life forms. After life had arisen, explanations of facts 2 and 3 would require logical and plausible paths for how primitive organisms were able to evolve the complicated gene-protein linkage seen in all living organisms. All living organisms use the same genetic code and they all use the same small set of basic constituent molecules. All means all. Bacteria and sequoia trees, cheetahs and snails, college students and amoeba all use the same information molecules of life.

The gene-protein linkup is the fundamental problem in the origin of life. Before a living organism can synthesize proteins the genetic material must be read and translated into the appropriate amino acids. On the other hand the genetic code cannot be translated without special proteins (replicases) that facilitate the copying process. In addition to the gene-protein linkup problem there are a few other problems that must be solved before a model on the origin of life is sufficient. Casti identifies three additional hurdles:

1. *Genetic code/protein structure.* There are thousands of possible amino acids and nucleic acids. So why are there only five nucleic acids and twenty amino acids used by all living things?
2. *Chirality.* All molecules in Nature have a mirror image. That means that all amino acids and nucleic acids have both left- and right-handed molecules which are identical in composition but which twist in opposite directions. The direction of twist determines function or chemical action of the molecule. The molecules produced in Miller-type experiments have equal amounts of right- and left-handed molecules. Interestingly all life forms on earth use only right-handed nucleic acids and only left-handed amino acids. Any origin model should explain why these forms of the respective molecules are used exclusively.

3. *Junk DNA.* With the exception of bacteria and viruses, all DNA contains long sections that do not code for any proteins. These junk segments have to be edited out before a protein can be made. Why is this material even here? If bacterial-like creatures were the first living things to come into existence then we need to understand how this extra DNA was incorporated into higher organism's genome.

## Genes-First Models

Let us consider the genes-first model. The essence of the gene-first model is that the first living things were not real organisms but rather replicators of random origin formed from the chemical constituents found in the primitive oceans. There were no proteins and therefore no early replicase enzymes. The major problem of this model is how the replicators form and how they replicate.

Some researchers have performed experiments that demonstrate that RNA can act as an autocatalyst by cutting out a central portion of itself and then resealing the cut ends. It has been shown that some RNA acts as an enzyme by cutting up RNA molecules that are different from it. This self-catalytic RNA can join several short strands of RNA together into chains under conditions that mimic the early earth.

Six major steps are involved in the gene-first model:

1. Start with a primordial soup that contains randomly constructed proteins and lipids (fatty acids) to be able to construct fragments of cell membranes. Nucleotide units must also be available for the construction of nucleic acids.
2. At least one self-catalytic replicating RNA molecule forms by chance. This molecule is not a gene because no proteins are formed. There is no unique nucleotide sequence. The RNA develops a range of enzymatic activities.
3. The RNA molecule evolves in self-replicating patterns and learns to exert control over proteins. The new proteins are better "enzymes" than the RNA was.
4. A series of interactions that are both complex and cooperative occur between nucleic acids and proteins.
5. DNA eventually appears which gives a stable, error-correcting information molecule.
6. RNA is no longer the premier molecule having been replaced by DNA as the information molecule and by proteins which perform the earlier enzymatic functions more effectively.

The biggest question in this scenario is the emergence of the first replicator. This is a random event and presupposes that a subset of right-handed nucleic acids happened to come together and exert control over the other molecules. How difficult is it to randomly assemble even a small strand of RNA? For this molecule to provide continuity between generations the molecule must replicate with a fairly high level of exactness. Unfortunately if the error rate associated with replication is greater than $1/N$ where $N$ is the number of nucleotide bases in the chain the population will probably go extinct. The "proof-reading" step in replication is performed by enzymes but enzymes would not have been made yet because they require longer—much longer strands of nucleic acids. Small RNA strands would not be long enough to code for enzymes and specifically replicases. If you cannot code for the enzyme, it would be

---

impossible to k
different after j
Siefert et al.
that features sl
history of bact
ria, there are at
same among tl
RNA-level me
RNA) might h
ancestor of the
ceded by proge

## Proteins-Fi

Let us now co
subsequently fc
and that of Fo
tique. The first
Oparin, the au
when certain oi
called *coacerva*
that the enzym
diffused into th
and the drople
"daughter" drc
The critical lin
felt that as m
would become
from random p
not have any h
overall summat

Pri

Oparin's obser
DNA so his m
In the 1960s,
model was bas
included lysine
formed polyme
Fox labeled th
formed million
the Oparin stu
researcher. In t
found. In sumr

contains long sec-
ve to be edited out
bacterial-like crea-
ieed to understand
enome.

t model is that the
of random origin
ins. There were no
blem of this model

that RNA can act
n resealing the cut
ting up RNA mol-
veral short strands
ly earth.

icted proteins and
branes. Nucleotide
ls.
chance. This mole-
unique nucleotide

ns to exert control
RNA was.
ive occur between

cting information

DNA as the infor-
nzymatic functions

rst replicator. This
nucleic acids hap-
es. How difficult is
nolecule to provide
fairly high level of
tion is greater than
pulation will prob-
ned by enzymes but
nger—much longer
enough to code for
nzyme, it would be

---

impossible to keep replication exact, and long strands of nucleic acids would all be different after just a few generations.

Siefert et al. (1997), using the complete genome sequences of five bacteria, suggest that features shared by these genomes must have arisen early in the evolutionary history of bacteria. While gene order is generally not preserved among these bacteria, there are at least 16 gene clusters of two or more genes whose order remains the same among the eubacteria. Many of these clusters are known to be regulated by RNA-level mechanisms in *E. coli*. This suggests that this type of regulation (i.e., RNA) might have arisen very early during evolution, and although the last common ancestor of these specific bacteria might have had a DNA genome, was likely preceded by progenotes with RNA-based genomes.

## Proteins-First Models

Let us now consider the models that presuppose that proteins came first and were subsequently followed by nucleic acids. There are two main models: that of Oparin and that of Fox. Both of these models have received considerable attention and critique. The first model is based on a series of observations and experiments by A. I. Oparin, the author of the original primordial soup recipe. Oparin has shown that when certain oily liquids are mixed with water that they form small droplets that are called *coacervates*. He added an enzyme that converted sugars into starch and found that the enzyme accumulated in the coacervates. When glucose was added the sugar diffused into the droplets where the enzyme proceeded to convert the sugar into starch and the droplet began to grow. At a certain size the droplet split apart and these "daughter" droplets would also grow and split as long as there was enzyme present. The critical link in this whole scenario is that the enzyme has to be present. Oparin felt that as more diverse molecules accumulated inside the droplet "metabolism" would become more diverse and life would begin. The formation of the initial enzyme from random processes is the weak link in this scenario. Furthermore, coacervates do not have any hereditary mechanism so natural selection could not act on them. The overall summation of Oparin's model is as follows:

Primitive cells (coacervates) → Enzymes or proteins → Genes

Oparin's observations and experiments took place before the discovery of genes and DNA so his model says little about inheritance.

In the 1960s, Sidney Fox came up with another model based on proteins first. This model was based on an observation that when amino acids in certain mixtures that included lysine, aspartic acid, or glutamic acid were heated under dry heat, they formed polymers. These polymers were different from anything found in biology, and Fox labeled them *proteinoids*. When the proteinoids were dissolved in water they formed millions of small spheres that had some nonspecific enzymatic capabilities. In the Oparin studies, the enzyme was very specific but added to the mixture by the researcher. In the Fox scenario, nothing was added, but some enzymatic activity was found. In summary, the model is as follows:

Amino acids → Protenoids → Cells → Genes

This model has been attacked by many scientists because the dry heat conditions seem difficult to come by and because, like the Oparin model, there is nothing for natural selection to act on.

## Dual-Origin Models

We have spent some time developing theories that from all appearances seem to be deficient in one or more points. These theories cannot be proved or disproved, but they can be supported with evidence from carefully designed experiments. All the theories presented are based on a "chicken or the egg" scenario or that one molecule came first. An alternative approach might be to consider a dual origin. Both proteins and nucleic acids are needed for life to propagate along the lines that all living things seem to follow today.

It is particularly difficult to see how nucleic acids with enough base pairs could have been made in the earth's early conditions. Remember that DNA and RNA have three major components: bases, phosphate, and a sugar. Using Miller flask-type experiments, researchers have been able to synthesize nucleotide bases in the laboratory but only under much colder environmental conditions. Sugars have been synthesized using formaldehyde under very restricted conditions. Phosphate is a natural component of oceans and rocks. Even if the scenarios of how the components can be made are correct or nearly so, there is the problem of how to get the component parts together. Not only do they have to be put together, they have to be put together in the right sequence every time. Nucleotides tend to dissolve in water, a condition that does not promote life.

Proteins carry a sort of genetic code. The order of the amino acids is directly related to the genes that code for them. The order of the amino acids contains the information found in the genes. This information in the absence of nucleic acid genes may have served as a template for making similar protein molecules. This template would require some structure that would support the protein in a fashion that allows the information (i.e., amino acids sequences) to be read, something akin to a ribosome. The basic gist of these models is as follows:

$$\text{Cells} \rightarrow \text{Proteins} \rightarrow \underset{\text{Much later}}{\rightarrow} \text{RNA} \rightarrow \text{DNA}$$

There are origin of life models that include clays and silica as the support structures that allow translation of proteins to occur. In these models, the crystalline substance would grow through natural abiotic means and any molecule attached through surface charges or otherwise would or could grow as well. We will not spend any more time on this subject, other than to point out that there is evidence or support for these notions from a wide variety of sciences, as summarized by Casti. These include the following ideas:

1. *Biology.* Genes are pure form and not substance. Evolution can act only on this type of replicable form.
2. *Biochemistry.* Nucleic acids, including RNA and DNA, are complex molecules that are fairly difficult to make. They were probably late arrivals geologically and evolutionarily speaking.

---

3. *Construction indu*
can lead to mutua
ical pathways.
4. *Structure of ropes.*
ing the continuity
isms based on a s
entirely different g
5. *History of technol*
in the immediate
them. As such, th
compared with m
ble or made from
very "low tech" c
6. *Chemistry.* The fo
as a primitive gen
7. *Geology.* Inorgani
urally weathering.
molecules that ha

Let us consider a n
quantitative predictio
favor a jump from dis
which represents the
that were found in the
that a primitive organ
in a chain of amino
ered that there were
behaviors in his mode

*a*: 8 to 10
*b*: 60 to 100
*N*: 2,000 to 20,000

What do these valu
world where life was fo
but the value of *a* sug
Miller flask type of ex
more complex ones. In
proteins before the ot
if *a* is less than 4, wh
nucleic acids to go fro
dicted chain size of p
tion error rate.

Figure 2.3 summari
to order is predicted t
the most meaningful c
region labeled Dead Z

3. *Construction industry.* Materials can be added or subtracted during evolution that can lead to mutual dependencies, such as the components of the major biochemical pathways.
4. *Structure of ropes.* Gene fibers can be added or subtracted without adversely affecting the continuity of the gene line. Casti suggests that this may be one way organisms based on a single genetic material could evolve into organisms based on an entirely different genetic material (i.e., proteins to nucleic acids).
5. *History of technology.* Primitive machines have to be made from available resources in the immediate region, and they have to work with little effort being put into them. As such, these primitive machines have a different design and construction compared with more advanced machines, which do not have to be easy to assemble or made from simple parts. In other words, the first organisms were probably very "low tech" compared with the organisms of today.
6. *Chemistry.* The formation of crystals is a low-tech mechanism that may have acted as a primitive genetic code.
7. *Geology.* Inorganic clay crystals are everywhere and continue to form through naturally weathering. Because of their net charge, they can attract and keep various molecules that have the opposite charge.

Let us consider a model proposed by the physicist Freeman Dyson that provides a quantitative prediction about the nature of primitive cellular metabolism that would favor a jump from disorder to order or life. The model has three main parameters: $a$, which represents the number of distinct amino acid or nucleic acid building blocks that were found in the original organism; $b$, the number of distinct chemical reactions that a primitive organism could catalyze; and $N$, the size of the molecular population in a chain of amino or nucleic acids that makes up such a life form. Dyson discovered that there were certain ranges of these parameters that produced interesting behaviors in his model. The ranges of interest were

$a$: 8 to 10
$b$: 60 to 100
$N$: 2,000 to 20,000

What do these values mean in the real world or, more importantly, in the primitive world where life was forming? All life on the earth today uses the same 20 amino acids, but the value of $a$ suggests that life could evolve with as few as eight amino acids. The Miller flask type of experiments produce most of the simple amino acids but not the more complex ones. In other words, 10 or so amino acids could form plenty of diverse proteins before the other amino acids came into being. Alternatively, the model fails if $a$ is less than 4, which implies that there is not enough chemical diversity in four nucleic acids to go from disorder to order. With $b$ in the range of 60 to 100, the predicted chain size of primitive proteins, the model can support a fairly high replication error rate.

Figure 2.3 summarizes aspects of the Dyson model. The transition from disorder to order is predicted to occur in the transitional zone that is highlighted. However, the most meaningful conditions are those near the cusp of the transition zone. The region labeled Dead Zone are states of the model where only disorder is found. This

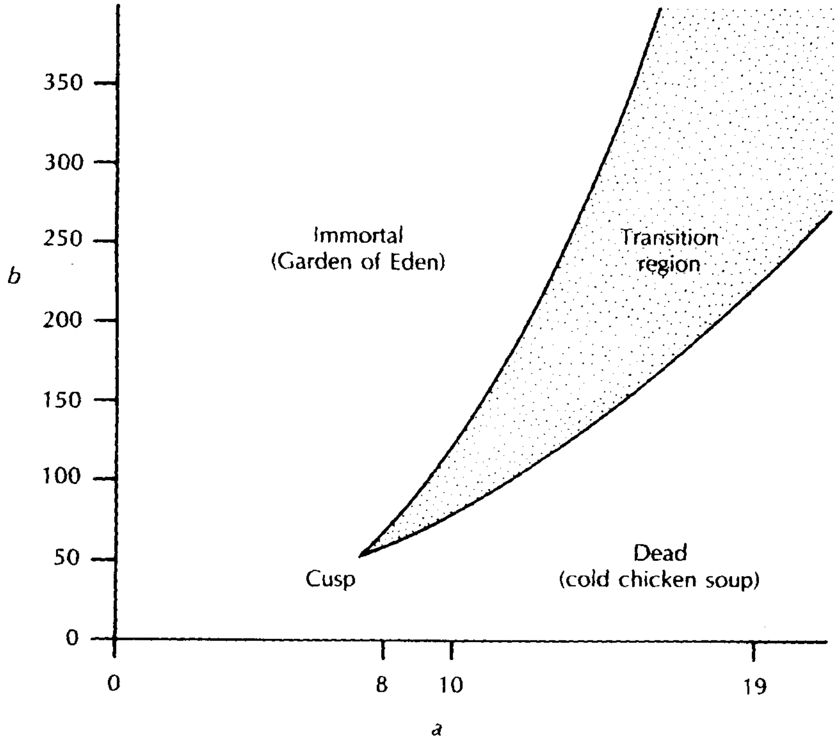


**Figure 2.3** Freeman Dyson model of the origin of life, in which *a* is the number of different amino acids or nucleic acids and *b* is the number of unique catalytic chemical reactions. Dyson predicts that life can originate at the cusp, where the number of amino acids is small and the number of reactions is around 50. (From Figure 4 in Dyson F. *Origins of Life.* Cambridge, UK, Cambridge University Press, 1999; reprinted with the permission of Cambridge University Press.)

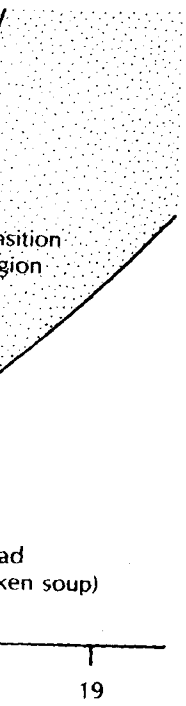
occurs because there is too much chemical diversity and too little catalytic capability. The region labeled "immortal" has too little chemical diversity and too much catalytic activity to allow a disordered state to exist (i.e., no death).

Life is found on the earth and the chemistry and biochemistry of that life is very similar and based on the same basic molecules. It is important to consider how these molecules came into existence because evolution is based on changes to these same molecules. Evolutionary ecology then becomes the study of how these molecules of life are modified by their environment and in turn affect the organism. The most primitive fossils have the appearance of bacteria and it seems likely that they were the first living things or closely related to them. Although the fossils suggest that the form of these early organisms is similar to that of some of today's bacteria, we cannot determine whether their cellular biology was similar, even though it seems likely based on the theoretical considerations given previously.

We have not provided an answer to the origin of life. The exact conditions of the ancient atmosphere, the salinity of the seas, the availability of important molecules are almost impossible to determine. However, because all of life uses the same basic blueprint, it seems logical to make some of the assumptions previously described. Life is more than the chemicals that can be analyzed. Immediately on death, all the chem-

---

icals of life are p
chemicals interac
still a major ques
any other organis



of different amino acids
predicts that life can
of reactions is around 50.
Press, 1999; reprinted

catalytic capability.
d too much catalytic

of that life is very
consider how these
nges to these same
these molecules of
sm. The most prim-
t they were the first
est that the form of
a, we cannot deter-
ems likely based on

t conditions of the
nportant molecules
uses the same basic
usly described. Life
leath, all the chemicals of life are present but the organism is dead. Understanding what makes those chemicals interact and function together in the orchestrated way of living things is still a major question. For microbes, the question has been around longer than for any other organism.