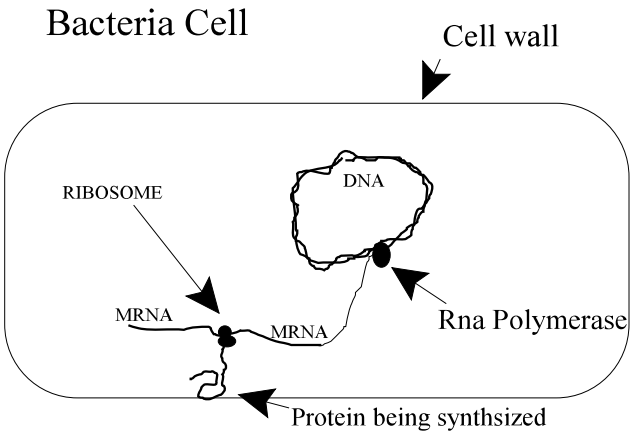
Chapter 9: Prebiotic Synthesis of RNA, DNA and Peptides

Naturalistic theories concerning life's origin began to take shape in 1953. Watson and Crick unraveled the structure of DNA, and Stanley Miller performed an experiment showing that amino acids can be produced in a spark chamber. Most scientists of the day assumed that the mystery of life's origin would be solved in a few years.

The early pioneers in this field realized that a complete living organism, like the bacteria in figure 9.1, could not spontaneously appear in a spark chamber or in any other environment governed by purely naturalistic laws. The pioneers needed the first form of life to be simpler than any living thing that is present on earth today.

Figure 9.1: Information Transfer in a Bacteria Cell



Initial theories hypothesized that the first living thing was a protein. These theories seemed reasonable at the time because many of the building blocks of proteins, amino acids, are easily synthesized under plausible prebiotic conditions. Because proteins regulate and control almost all of the activities necessary for life, the living protein theory quickly gained widespread acceptance, but soon scientists realized that there was a major flaw with the protein theory.

Proteins cannot self replicate, so the first living protein would not be able to reproduce itself, and without replication there can be no natural selection; therefore, the first living protein would have no way to evolve.

This issue led to the demise of the protein theory. In its place, emerged the RNA theory. This theory gained substantial momentum when it was found that just like proteins some RNA molecules can catalyze chemical reactions. Recently this theory has also fallen out of favor because it has its own set of problems which will be discussed later. Today the most popular theory involves a self replicating pre-RNA molecule.

Self replicating molecules are probably not the best theory to pursue, because such molecules cannot reproduce for any length of time without running into serious problems with the second law. Nevertheless, many researchers in the origins field are absolutely sure that the first living thing was a self replicating chemical, and their point of view is understandable. There is simply no chance that a complete bacterium spontaneously formed from the chemicals in a puddle four billion years ago. In many ways, a self replicating molecule that violates the second law is a better choice.

Nevertheless, the second law should not be casually dismissed because its existence explains why investigators have not been able to create a self replicating molecule in the lab. Unless a self replicator has the knowledge and ability to harness the power of sunlight (or some other abundant energy source) and use this energy to drive its own replication, then its lifetime will be short lived and its existence forbidden by the laws of physics.

The origin of self replication requires a solution to five problems:

- Chemical evolution must create a protein, an RNA molecule or an RNA like molecule.

- This molecule must possess the molecular knowledge that enables self replication.

- It must also be able to implement this knowledge.

- The molecule must be able to harness an energy source to do useful work.

- The first self replicator must be able to synthesize any chemicals lacking from its surroundings that it needs to self replicate.

Experiments investigating the origin of life have for the most part ignored the last two issues. This is understandable because until a molecule that can at least replicate itself for a little while can be found, there is no need to try to find one that can replicate itself indefinitely. This chapter will investigate the prebiotic synthesis of RNA and proteins. The next chapter will investigate self replication.The pioneers in chemical evolution expected to show that the primordial ocean was full of biological molecules. These researchers suggested that the early atmosphere contained no free oxygen, and that under these conditions, the required biological precursors should be plentiful. The remainder of this chapter will evaluate the validity of this hypothesis.

It is extremely difficult to synthesize biological molecules under plausible prebiotic conditions, and today this difficulty has led most to conclude that the primitive ocean contained a very limited supply of biological precursors. This finding does not mean that the primordial soup did not exist. It does mean that the primitive ocean was not the primordial soup because any relevant molecules in it would be too dilute.[4,11,18]

It is possible to imagine environments that will concentrate biological precursors, but this leads to further problems. It limits the soup in such a way that the conditions necessary for its existence rarely exist and leads to the perhaps alarming conclusion that even given 5 billion years the soup may not have existed.

Zero Tries

The goal of this chapter and the next is to show that given 5 billion years and an almost unlimited source of energy, the probability of creating a protein or an RNA molecule is vanishingly small. Furthermore, the probability that the molecule so created contains the knowledge needed to self replicate is also vanishingly small. The chance of success is given by multiplying these two vanishingly small numbers. The trapped scientist in figures 9.2 and 9.3 helps illustrate this concept. With zero tries, even a short combination eludes the scientist (figure 9.2), and unfortunately, the required combination for self replication whether protein or RNA is quite long (figure 9.3).

Figure 9.2: Zero Tries

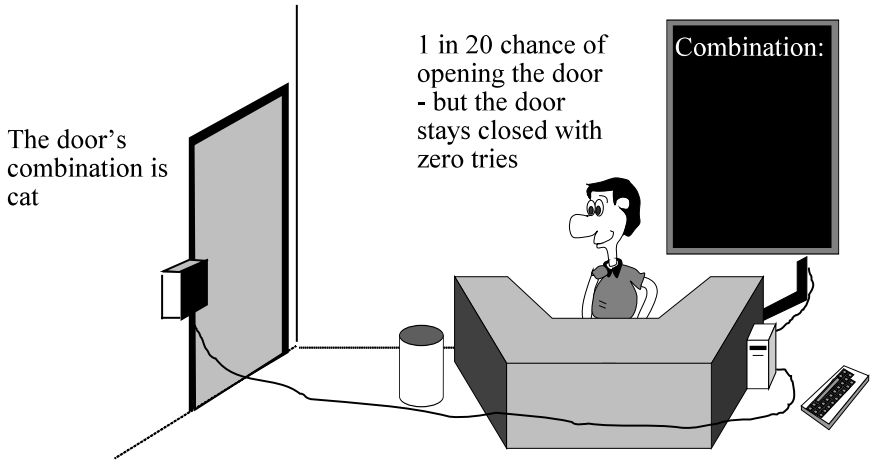
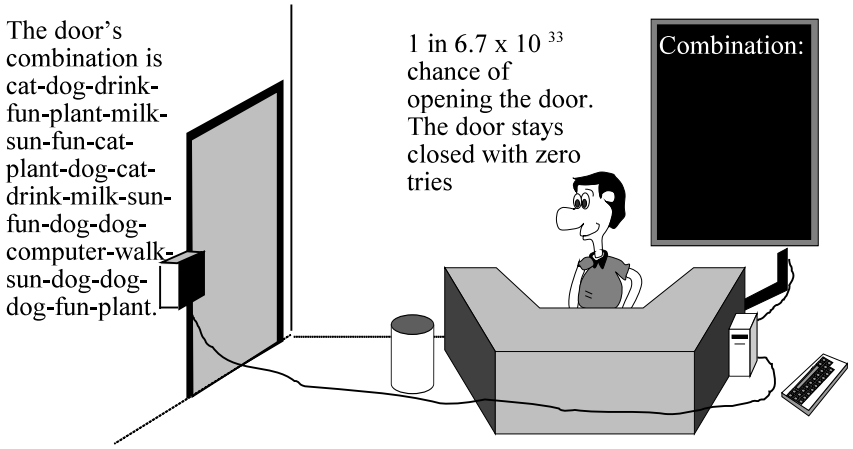


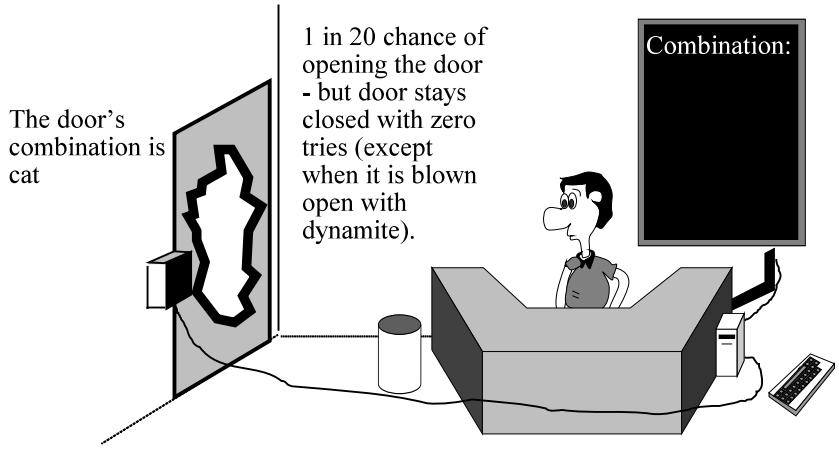
Figure 9.3: Zero Tries with Long Combination



Investigator Interference

In figure 9.2, the scientist is not cooperating. He refuses to play the game. He enters no words into the computer, so he accumulates no tries. The researches are very unhappy with these results. So they blast the door with dynamite. This of course opens the door (figure 9.4). The researchers then conclude that given 5 billion years the door will open. Figure 9.4 is an obvious example of investigator interference.

Figure 9.4: Investigator Interference



The concept of investigator interference was first introduced by Thaxton et al. in The Mystery of life's Origin: Reassessing Current Theories. In this book, the authors suggest that some interference is warranted. Scientists cannot conduct experiments that last for one billion years. So interference is useful in that it speeds up the process of evolution, and to be fair, the interference is a great learning tool because it allows scientists to rule out extremely unlikely scenarios. Thaxton also concludes that in many cases the interference is excessive.

While the interference is a good idea because it helps scientists learn, it can also be very misleading. The scientist did not open the door in figure 9.4. The dynamite opened the door. Any conclusion that given time, the scientist will open the door is completely unfounded. This chapter will introduce many examples of interference. Readers should use their own judgement as to whether the degree of interference is acceptable or excessive using the following criteria: if the artificial conditions generated in the lab might happen in nature given 5 billion years, then the interference is acceptable. Otherwise, it is excessive. Proteins will be considered first followed by RNA.

Protein Synthesis

Synthesizing proteins under prebiotic conditions is not as straight forward as many would have predicted. Ten (maybe 12) of the amino acids are relatively easy to create. Both L and D isomers are created, and two amino acids alanine and glycine almost always dominate the mixture. Despite these issues, creating amino acids is not that difficult. It is forcing the amino acids to form peptide bonds that is difficult.

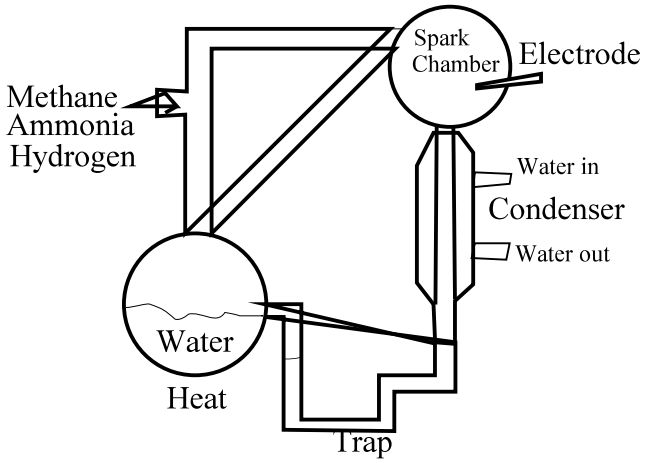
Miller's Experiment

Figure 9.5 illustrates the Miller spark chamber. The water in the flask is boiling. The atmosphere above the water and in the spark chamber is controlled. In this example, hydrogen, methane and ammonia are introduced. The electrode is charged to a very high voltage, and it creates an electric spark. This spark is an energy source. It allows the chemicals in the chamber to react and form new chemicals. The condenser removes the chemicals from the spark chamber, and they accumulate in the trap. Life uses 20 amino acids. Miller's chamber can create between 0 and 10 of the 20 (the number created depends on the gases used in the atmosphere).

The chamber also creates many other chemicals. Other scientists have repeated this experiment with alternative energy sources like UV light and heat. These experiments demonstrate that many amino acids are easy to create. Miller's chamber is a non-equilibrium system cleverly designed and optimized to create nonvolatile organic compounds like amino acids.

Whether or not this experiment is representative of the conditions on the early earth is questionable. Many scientists today do not believe that ammonia, hydrogen and methane were present in the earth's early atmosphere, and without at least one of these, no amino acids are produced by the spark chamber.

Figure 9.5: Miller's Spark Chamber



Thermal Proteins or Protenoids

Since water inhibits the formation of peptide bonds, the first step to create a peptide often involves removing water. Fox successfully created chains of amino acids by heating a purified concentration of amino acids to 150 degrees Celsius for about 14 hours. At this temperature, water and other volatile compounds vaporize. This is important because when a peptide bond forms, a single water molecule is also produced. The heat drives this molecule off forcing the reaction forward because without water it cannot go backwards.

Fox obtained very long chains when he included high concentrations of the amino acids, glutamate, aspartic acid and lysine. Fox called the amino acid chains formed by heating, protenoids. They are also called thermal proteins. They are different from normal proteins in two important ways. Thermal proteins contain both D and L isomers, and the peptide bonds that form are very unusual. The side chains associated with lysine, glutamate and aspartate form over ½ of the peptide bonds.[1] This second feature has led most origin of life researchers to drop protenoids as a viable candidate for the first living protein. Stanley Miller in particular has criticized thermal proteins as unlikely candidates because the conditions necessary to form them probably rarely exist. The temperature has to fall within a narrow range (150-180 deg C), and if the heating lasts too long (more than a day), then the thermal proteins are destroyed.[2] Furthermore, given that amino acids will not form thermal proteins without a very high concentration of aspartate, glutamate, or lysine leads to another question. How do proteins with reasonable concentrations of these 3 amino acids form in the soup?

Given 5 billion years, a few thermal proteins may have had a chance to form. In this respect, thermal proteins are unique. While they are not biological precursors (due to the unconventional peptide bonds), they do at least have a chance of existing.

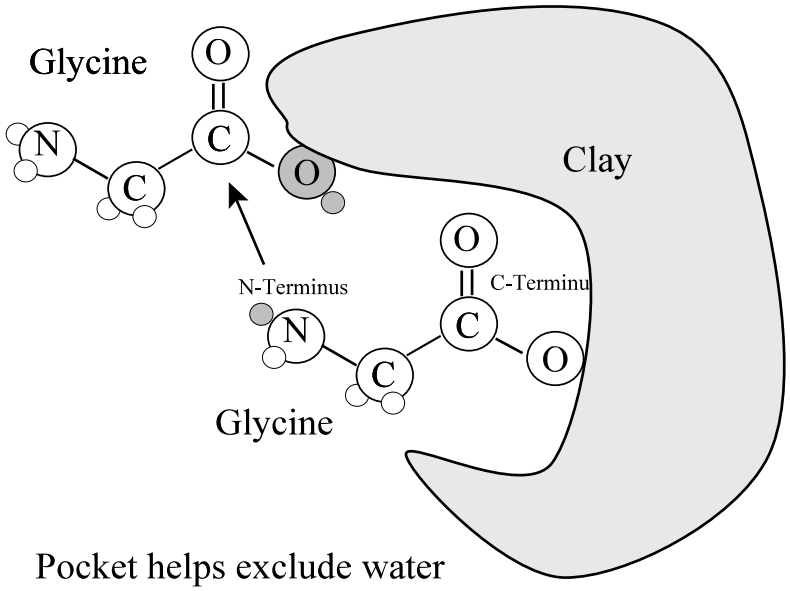
Short Peptides Chains in Water

Short peptide chains have been produced in water. Usually a catalyst like clay or some other mineral like pyrite is required. The minerals promote the formation of peptide bonds.

Peptide bond formation can also be induced by sodium chloride (salt) in the presence of copper.[24] These salt-induced peptide bonds are generally limited to very short chains. Nevertheless, the reaction is of interest because salt and copper are very common, the reaction takes place in water, and at least for a few amino acids, there seems to be a slight preference for peptides composed of L-amino acids. These short peptides may then interact with other minerals like clay to form longer chains.[25] Clay can form pockets that may help exclude water. Figure 9.6 shows how a mineral like clay may help a peptide bond form. The C-terminus of each glycine molecule interacts with the clay substrate. The arrow shows how the N-terminus of one glycine attacks the C-terminus of the other. This forms a peptide bond. Peptide chains of up to 10 amino acids have been created in the lab using these techniques.

Salt, copper and clay are very common. These minerals would have been present in or near the primordial soup and facilitated amino acids joining together into short peptides. While the resulting peptides would be too short to be biological precursors (< 10 amino acids), some of these peptides almost certainly existed on earth before life. A few readers might assume that given 2 billion years these processes would naturally create longer peptide chains. This is unlikely because the soup must have contained chemicals other than the amino acids used by life. Specifically, formic acid, amines, formaldehyde, and non-biological amino acids must have been present, and all of these would interfere with proper chain growth. Furthermore, many destructive processes would destroy any growing chains. So given 2 billion years, maybe a few peptides greater than 10 amino acids evolved, but the quantity would have been very limited. Because the starting point for these experiments model evolution in a test tube, one week may already correspond to a billion years of evolution in nature. This technique certainly did not fill up the primitive oceans with peptides.

Figure 9.6: Clay and Peptide Bond Formation

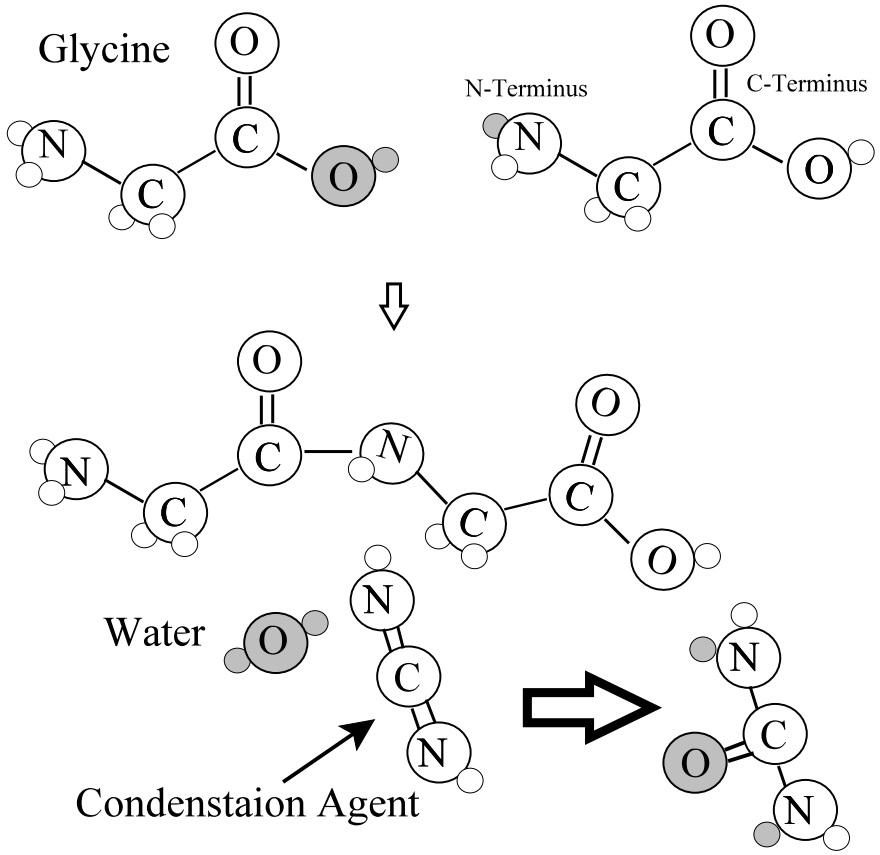


Long Peptide Chains in Water

The last class of experiments will consider extreme investigator interference. In these experiments, a chemical that is either not found in the Miller type spark chamber or is very rare is added to the solution in high concentrations. The chemical is almost always a condensation agent. These chemicals contain double bonds that can absorb water. The reaction is sometimes carried out in a set of sequential steps specifically designed to elongate the peptide chain. Figure 9.7 shows how this process works. The condensation agents must be used in high concentrations because they are not stable in the presence of water. Condensation agents do not have the ability to differentiate between the water molecule released when two amino acids combine, and the water molecules already present in the soup. Thus, finding a condensation agent in the primordial soup is like finding a dry sponge in the ocean.

Many authors have claimed that experiments that use condensation agents are relevant to the origin of life. Nevertheless, the justification for adding a condensation agent that has no plausible route for prebiotic synthesis to the mixture is questionable, and adding it in the excessive quantities required for peptide chain growth is certainly not justified. The sequential washing steps which are controlled by the investigator add to the already excessive interference. The results of these experiments are not relevant to the origin of life.

Figure 9.7: Peptide Bonds With Condensation Agents



RNA Synthesis

When researchers moved from the living protein theory to the living RNA theory, they unknowingly took a giant step backwards. The decision to switch was made for two reasons: 1) Conceptually, RNA should be able to replicate itself much more easily than a protein and 2) RNA can under special circumstances regulate a few chemical reactions. This second finding solidified RNA as the natural choice for the first living organism. The hope was that it might be able to regulate its own synthesis and thus be a very effective self replicating molecule. Nevertheless, the living RNA theory has created a whole new set of problems that need a solution. Many are much more difficult than the problems created by the living protein theory.

1) The building blocks for RNA are harder to synthesize under plausible prebiotic conditions than amino acids. In fact, cytosine has never been synthesized. Cytosine is also absent from meteorites.

2) Unlike amino acids, two of the building blocks required for RNA (cytosine and ribose) are not stable and have very short lifetimes. It is unlikely that these molecules existed in the soup.

3) Just like proteins, the building blocks for RNA do not form RNA molecules unless water is excluded. Given the short lifetimes of many of the RNA subunits, the high temperatures required to drive off water just accelerate decomposition.

RNA Building Block Synthesis

Creating several amino acids is easy. The hard part is coercing the amino acids to link together in a chain to form a protein. RNA proves much more difficult because even the building blocks are hard to synthesize. Furthermore, once they are created, they do not last long. This makes it difficult to understand how the necessary building blocks achieved a suitable concentration for further reactions. Several key building blocks will now be considered.

Adenine and Cytosine

Adenine has been synthesized in the lab from concentrated solutions of hydrogen cyanide and ammonia. While this process works in the lab, it is not clear how the necessary conditions to create adenine would arise in nature.

To synthesize significant quantities of adenine, a concentrated solution of hydrogen cyanide and ammonia is required. Concentrating hydrogen cyanide and ammonia under plausible conditions is problematic. Hydrogen cyanide is a very reactive chemical. In low concentrations, it reacts with water to form many products that are not adenine. These side reactions use up the hydrogen cyanide and lower its concentration. To make the process more difficult, one of the most abundant chemicals produced in the early atmosphere was undoubtably formaldehyde and "Formaldehyde reacts spontaneously with hydrogen cyanide to form cyanohydrin, a well known reaction that has vexed workers in the field of prebiotic chemistry relying on the unencumbered availability of HCN in high concentration to form a plethora of evolved molecules."[21] Ammonia is equally problematic because it decays rapidly when exposed to sunlight,[15] and it boils at sub-freezing temperatures. So while some adenine might be formed under plausible conditions, very little is produced. The high concentrations of ammonia and hydrogen cyanide required to make adenine do not represent plausible prebiotic conditions.[3]

Because adenine has been found in meteorites, there is evidence that it is produced by nature in space.[3] Nevertheless, based on the above discussion, adenine was certainly a very rare chemical 4 billion years ago.

Cytosine is much more problematic than adenine. It has never been produced under any plausible prebiotic conditions, even in minute quantities. It is not found in meteorites, so it is not easily synthesized in space. Cytosine is not stable in water. Its lifetime depends on the temperature. At 100 degrees Celsius, cytosine decomposes in 19 days. At room temperature, the decomposition is 340 years. These observations have led Miller and several other researchers to suggest that Cytosine was not found in the first self replicating molecule. [5,6]

Ribose

Ribose is the most troublesome subunit. It can only be synthesized in small quantities under plausible prebiotic conditions, and its lifetime in water is extremely short (73 minutes at 100 degrees Celsius, and 44 years at 0 degrees Celsius). Given that it is hard to synthesize in large quantities and that it decays rapidly once it is produced, it is difficult to see how a reasonable concentration of ribose ever existed in the soup. Many scientists including Miller have suggested that the first RNA molecules probably did not include ribose (thus the term pre-RNA).[7]

Ribose presents another difficulty. Just like amino acids, sugars have isomers (mirror images). It has been found experimentally[8,9] that these isomers interfere with self replication. The interference is severe because it terminates the growing chain. So when the first living prebiotic RNA tries to replicate, it must do so in an environment enriched in one isomer of ribose. No mechanism for such an enrichment has been proposed by researchers.

Many scientists have decided that the problems with ribose are so severe that the molecule should be excluded as a possible building block. Since ribose is just the glue that holds RNA together, other chemicals should be able to take its place.[8]

Finally, ribose is a reducing sugar. This means that it will react very quickly with amino acids, and the resulting polymer will fall out of solution. Any ribose in the soup will quickly be eliminated by reactions with the amino acids in the soup.

A Pre-RNA World?

The most recent origin of life theory involves a pre-RNA living molecule. This molecule probably lacked cytosine and ribose. Because such a molecule no longer exists in life, it is hard to address all the possible candidates. How can one possibly test an hypothesis phrased as follows: We believe that some chemical (but we don't know what it was) at one time lived on earth, and this chemical was capable of self replication. We are confident that one day we will find it, and prove our hypothesis correct.

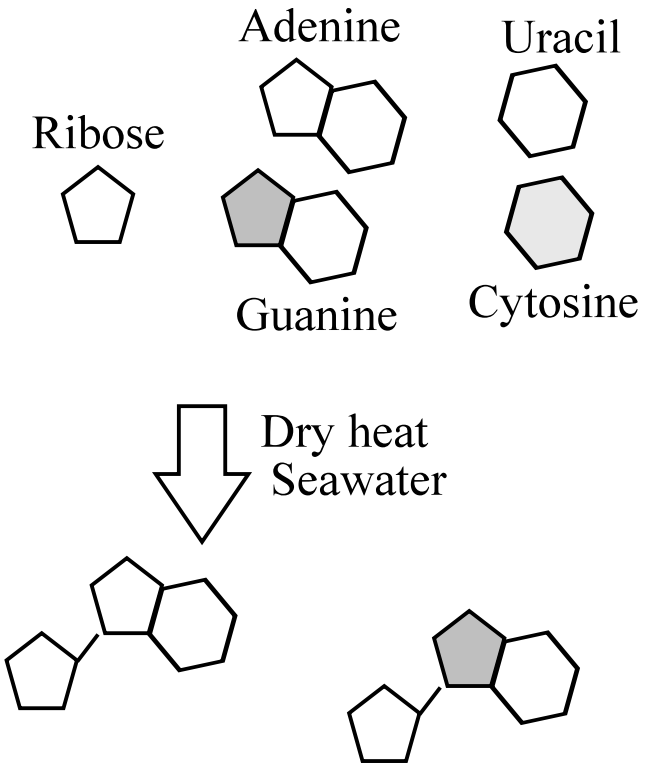
Assembling the Building Blocks

The building blocks for RNA are called nucleotides. A nucleotide consists of 1 phosphate group attached to a ribose which in turn is attached to one of the four bases, uracil, cytosine, adenine or guanine.

In the case of proteins, the amino acid is the smallest building block. No condensation reactions are required to create amino acids. In contrast, two condensation reactions are required to create a nucleotide . One to attach the phosphate to ribose and one to attach one of the bases (adenine, guanine, uracil, and cytosine) to the ribose (figure 8.10).

Under optimal conditions, adenine and guanine can be attached to ribose in the lab. The procedure involves dry heat and sea water. Nucleotides that use cytosine and uracil have no plausible mechanisms to attach the base to ribose.[10]

Figure 9.8: Condensation Reaction (Subunits of RNA)

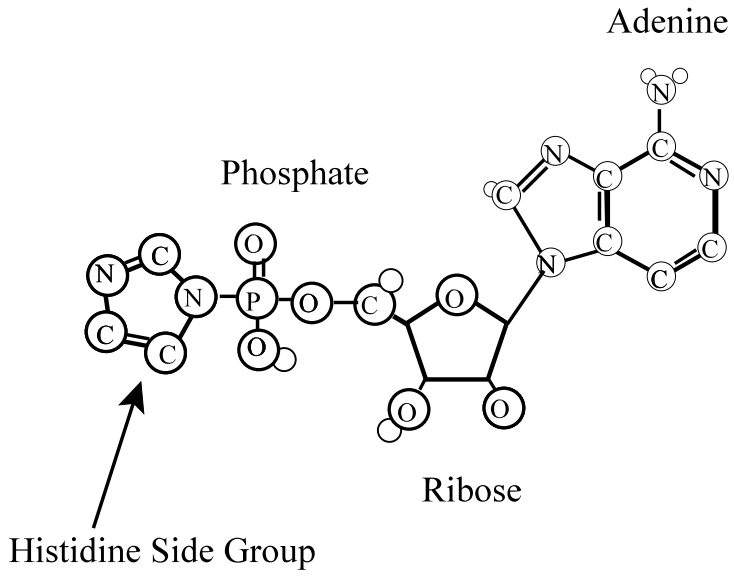


Uracil and Cytosine do not participate.

Activated Monomers

Condensing agents are popular for peptide synthesis. They are also effective for RNA synthesis, but in the case of RNA, the nucleotides are usually directly activated before being added to the mixture. In the presence of clay, such activated monomers have been shown to form chains greater than 50 nucleotides long.[16] The most popular activation agent is impA and is shown in figure 9.9. To form impG, impC, and impU replace the adenine with the appropriate base. The source of these activated molecules is unknown. They would have not been present in the soup, so when researchers add them to test tubes in high concentration and then claim that their experiment models the origin of the life, their claims are without merit.

Figure 9.9: ImpA



Review of Investigator Interference

Investigator interference will now be summarized. As mentioned earlier, investigators do not have 5 billion years to observe experiments, so some interference is necessary.

Interference Strategy #1: Eliminate the Undesirable Chemicals

If chemical A and chemical B react to form chemical P, then this chemical reaction can be written as A+B—> P. Suppose that Miller's water trap contains 3 chemicals, A, B and C. The possible reactions involving the chemicals are as follows: A + B–>P and A +C->D.

Unfortunately, the second reaction is favored. So after a few days all of the chemicals in the flask are D, but the researcher desires chemical P. So instead of using the contents of the flask to create P, he orders A and B from his chemical supplier. He mixes these two chemicals while applying heat, and the product is P. This process is how organic chemists make chemicals. They control the chemicals that they start with, and this influences the products that they get. Applying this technique to origin of life scenarios is questionable because it is not clear how nature can exclude the undesirable chemicals.

In chapter 7, figure 7.1 shows that one of the functions that enzymes perform is to eliminate undesired reactions. They accomplish this by speeding up the desired reactions. When investigators manipulate the chemicals in their system to create a desired product, they are mimicking this particular mode of enzyme action. They are using their knowledge of chemistry because the required molecular knowledge is not present in the system.

Examples of cross reaction elimination:

- Fox's thermal proteins. He did not include carboxylic acids or other organic components (like aldehydes) that might terminate a growing protein chain.

- The most extreme examples of cross reaction elimination involve RNA. The reason is that ribose is included freely, but no amino acids are included. This is not a plausible condition. Amino acids react very quickly with sugars like ribose to create very long chain polymers. Anyone who has baked cookies or toasted a piece of bread is familiar with this reaction. Browning is caused when amino acids (especially lysine) react with sugar. This reaction would make any sugar present in the primordial soup unavailable for RNA formation.[11,12]

Interference Strategy #2: Concentrating Volatile Chemicals

Concentrated formaldehyde is critical for the synthesis of ribose. Concentrated hydrogen cyanide and ammonia are critical for the synthesis of adenine. It is not clear how these chemicals could ever be present in high concentrations on the early earth.[13] How does one concentrate a chemical that boils at sub-freezing temperatures in a small puddle? This is a difficult problem.

Interference Strategy #3: The Use of Condensation Agents or Activated Monomers

Condensation agents help form many of the bonds that are necessary in biological precursors, whether RNA or protein. Condensation agents remove water and by doing so promote the formation of large biological molecules. Condensation agents were discussed for proteins, but they have also been used to successfully join RNA nucleotides into short chains. RNA synthesis usually just skips this step, and instead researchers usually just add an activated monomer like impA, impG, impC, or impU.

There are no plausible synthesis mechanisms for the condensation agents or the activated monomers. If they are created in Miller's spark experiment or if they exist in meteorites, then the amount present is minuscule. How some investigators can add these chemicals to reactions in massive qunatities, and still think that they are modeling plausible prebiotic conditions is certainly an unsolved mystery.

Nevertheless, the motivation for using these techniques is clear. Without these techniques, the biological precursors are limited to a size that is too small to be biologically active.[16] Given that condensation agents and activated monomers are often coupled with carefully timed washes designed to grow the protein or RNA molecule, the analogy to blowing up the door in figure 9.4 definitely applies.

Interference Strategy #4: Controlling the Energy Sources

In most experiments, destructive energy sources are eliminated by the investigator. For example, if the trap in Miller's spark chamber is illuminated with UV light, many of the products will be destroyed.[4]

Interference Strategy #5: Substituting Human Knowledge

This is the most subtle form of interference, and the most common. In systems that lack the required molecular knowledge, it is very easy for researchers to unintentionally add knowledge to the system through the design of their experiment.

The carefully controlled sequential washes that accompany many RNA and protein chain elongation experiments are a perfect example. Often a growing RNA or protein molecule is attached to a stationary substrate, activated nucleotides or amino acids are added, and a rinse is applied after the desired chemical bond forms. This form of interference is present in most prebiotic experiments, and sometimes it goes unnoticed.

Conclusion:

The goal of this chapter was to show that the precursors to life whether RNA or proteins are extremely difficult to create. Maybe one or two such molecules are expected given optimal conditions and 5 billion years. The design inference based on this conclusion alone is very strong. The inference will be strengthened in the next chapter. The next chapter will show that the knowledge required for self replication is very large. If the entire ocean is packed tight with either proteins or RNA, then the odds that one of the molecules can self replicate is still zero. Several thousand bits of knowledge are required, and zero tries (or almost zero) will never allow chance to find a solution.

Many investigators researching the origin of life are disappointed with their progress, and this shows in the scientific literature. Today, it is acceptable to publish an article that is critical of the origin of life paradigm as such articles do get published.

Any publication suggesting the possibility of design is either rejected or starts a witch hunt in which the editor who approves the article is the target. The first step in any scientific revolution is to realize that there is a problem with the current theory, and for many scientists this realization has already taken place. Joyce and Orgel summarize the situation as follows:

"In our initial discussion of the RNA World we will accept The Molecular Biologist's Dream: "Once upon a time there was a prebiotic pool of Beta-D-nucelotides . . . . . We will now consider what would have to happen to make the dream come true. This discussion triggers the Prebiotic Chemist's Nightmare: how to make any kind of self replication system from the intractable mixtures that are formed in the experiments designed to simulate the chemistry of the primitive earth." [20]

References:

1) Temussi, et al., "Structural Characterization of Thermal Prebiotic Polypeptides," Journal of Molecular Evolution, p105-110, 1976.
2) Miller, Orgel, The Origins of Life on Earth, Prentice Hall, 1974.
3) Shapiro, "The Prebiotic Role of Adenine: A Critical Analyis," Origins of Life and the Evolution of the Biosphere," 25:83-98, 1995.
4) Thaxton, Bradley, Olsen, The Mystery of Life's Origin: Reassessing Current Theories, Philosophical Library, 1984.
5) Levy, Miller, "The Stability of the RNA bases: Implications for the Origin of Life," PNAS, 95: 7933-7937, 1998.
6) Shapiro, "Prebiotic Cytosine Synthesis: A Critical Analysis and Implications for the Origin of Life," PNAS, 96: 4396-4401, 1999.
7) Larralde, Robertson, Miller, "Rates of decomposition of Ribose and other Sugars: Implications for chemical Evolution," PNAS, 92: 8158-8160, 1995.
8) Joyce, Schwartz, Miller, Orgel, "The Case for an Ancestral Genetic System Involving Simple Analogues of the Nucleotides," PNAS, 84: 4398-4401.1989.
9) Joyce, Visser, Boeckel, Boom, Orgel, Westrenen "Chiral Selection in Poly (C) Directed Synthesis of Oligo (G)," Letters to Nature, 310: 602-604,1984.
10) Fuller, Sanchez, Orgel, "Studies in Prebiotic Synthesis. V11 Solid State Synthsis of Purine Nucleosides," Journal of Molecular Evolution, 1:249-257, 1972.
11) Nissenbaum, Kenyon, Oro, "On the Possible Role of Organic Melanoidin Polymers as Matrices for Prebiotic Activity," Journal of Molecular Evolution. 6:253-270, 1975.
12) Thaxton, Bradley, Olsen, The Mystery of Life's Origin: Reassessing Current Theories, Philosophical Library, pp 60-61,1984.
13) Thaxton, Bradley, Olsen, The Mystery of Life's Origin: Reassessing Current Theories, Philosophical Library, p 64,1984.
14) Ferris, "Prebiotic Synthesis: Problems and Challenges," Cold Spring Harbor on Quantitative Biology, Vol L11: 29-34, 1987.
15) Thaxton, Bradley, Olsen, The Mystery of Life's Origin: Reassessing Current Theories, Philosophical Library, pp 43-44,1984.
16) Ferris, "Montmorillonite Catalysis of 30-50 Mer Oligonucleotides: Laboratory Demonstartion of the Potential Steps in the Origins of the RNA world," Origins of Life and Evolution of the Biosphere, 32:311-332, 2002.
17) Osterberg, Orgel, Lohrmann, "Further Studies of Urea Catalyzed Phosphorylation Reactions," Journal of Molecular Evolution, 2:231-234, 1973.
18) Fox, Dose, Molecular Evolution and the origin of Life, Freeman and Co., 1972.
19) Thaxton, Bradley, Olsen, The Mystery of Life's Origin: Reassessing Current Theories, Philosophical Library, pp 66,1984.

References (continued)

20) Joyce and Orgel, The RNA World, Gesteland, Cech, Atkins, Cold Spring Harbor, "Prospects for Understanding the Origins of the RNA World," p50, 1999.
21) Mojzsis, Krishnamurthy, Arrhenius, The RNA World, Gesteland, Cech, Atkins, Cold Spring Harbor, "Constraints on Molecular Evolution," p20-21, 1999.
21) Fox, Dose, Molecular Evolution and the origin of Life, Freeman and Company, p37, 1972.
22) Shapiro, Origins: A Skeptics Guide to the Creation of Life on Earth, 1986.
23) Overman, A Case Against Accident and Self-Organization, 1997.
24) Plankensteiner, Reiner, and Rode, Stereoselective Differentiation in the Salt-induced Peptide Formation Reaction and Its Relevance for the Origin of life, Peptides, 2004.
25) Bujdak, Eder, Yongyai, Faybikova, and Rode, Investigation on the Mechanism of Peptide Chain Prolongation on Montmorillonite, Journal of Inorganic Biochemistry, 1996.

Readers who wish to read more about chemical evolution and its problems should try to find Thaxton's book, The Mystery of Life's Origin: Reassessing Current Theories, in their local university library. Unfortunately, the book is out of print. His book was the primary reference for this chapter. The two papers by Shapiro, reference 3 and 6, are also excellent resources.

Chapter 10: Self Replicating Molecules and Systems

This chapter will attempt to quantify the amount of molecular knowledge needed for self replication. Both proteins and RNA will be considered. While many researchers have theorized that one of these molecules emerged as the first self replicator, origin theories stand a much better chance if both are involved. While RNA can perform some of the functions normally performed by proteins, proteins are much more efficient. Amino acids have many functional groups available in their side chains, and these functional groups impart to proteins a versatility than RNA cannot possibly possess. To understand why a system comprised of both is better, consider how numbers and letters are used in the following two sentences.

- The number is 4,900,555,015 dollars.
- The number is four billion nine hundred million five hundred fifty five thousand and fifteen dollars.

Often numbers communicate numerical concepts better than words. The first sentence is much easier to understand. Forcing RNA to do the job of a protein is clumsy. It is analogous to writing out a very large number using words to represent the numbers. Just because it is possible, does not mean that it is the easiest or best way to accomplish the task. RNA is good at storing information. Proteins are good at regulating chemical reactions. The first system of replicating molecules was probably a combination of both, and a good model for such a system is alive and well today in the simplest bacteria. Nevertheless, because chemical evolution does not explain the spontaneous emergence of bacteria from the primordial soup something simpler needs to be considered. The goal of this chapter is to show that something simpler does not work because simple systems cannot self replicate.

A Self Replicating Peptide

In 1996, an article was published in Nature in which David Lee reports to have found a self replicating peptide.[1] The title of the article is appropriately "A Self Replicating Peptide." Unfortunately, the investigator interference required for self replication is perhaps the most extreme in the history of origins research.

The peptide of interest contains 32 amino acids. The sequence is as follows:

arg-met-lys-gln-lys-glu-glu-lys-val-tyr-glu-lys-lys-ser-lys-val-ala-cys-leu-glu-tyr-glu-val-ala-arg-leu-lys-lys-leu-val-gly-glu.

The peptide does not self replicate using amino acids. Instead it uses a pool of two peptides, one is 17 amino acids long and the other is 15 amino acids long. The amino acid sequences of these two peptides are shown below. Notice that if a peptide bond forms between ala (last amino acid on right in the peptide with 17 amino acids) and cys (first amino acid on left in the peptide with 15 amino acids) then a replica of the self replicating peptide results.

arg-met-lys-gln-lys-glu-glu-lys-val-tyr-glu-lys-lys-ser-lys-val-ala

cys-leu-glu-tyr-glu-val-ala-arg-leu-lys-lys-leu-val-gly-glu

Because the peptide with 32 amino acids facilitates the formation of this single peptide bond, Lee claims that this peptide can self replicate. But is this really true? To self replicate, this peptide requires a pool of two peptides. One of these peptides has the same amino acid sequence as the first 15 amino acids in the self replicating peptide, and the other has the same amino acid sequence as the next 17 amino acids. Where do these peptides come from? In this case, they are supplied by the investigator.

Chapter 9 discussed the difficulties of creating peptide chains under plausible prebiotic conditions. Due to the difficulties, peptides with more than six amino acids are expected to be very rare chemicals. Peptides composed of 15 to 17 amino acids will be much more scarce. Yet to self replicate, this peptide requires an abundant supply of both, and not just any peptide. One of these peptides must be identical to the first half of the self replicator, and the other peptide must be identical the second half of the self replicator.

This last requirement is particularly troublesome. Suppose the self replicator comes into contact with two random peptide chains. One is 15 amino acids long and the other is 17. How often will the two smaller peptides be an exact replica of the self replicator? Answer 1 time in every $4x10^{41}$ tries (assuming that every amino acid has a 1 in 20 chance of occurring at each position). Given the low concentration of peptides in the primordial soup, the probability for such an encounter is zero.

The interference does not stop here. It is critical that the first amino acid in the peptide with 15 amino acids be a cysteine. Cysteine has chemical properties that facilitate peptide bond formation, and to make sure that the interference sets the record, the alanine (last amino acid on right in the peptide with 17 amino acids) must be chemically altered to make it much more susceptible to attack by the sulfur atom in cysteine's side chain.

Finally, the self replicating peptide contains eight lysines. Lysine is instrumental in its self replication as its charge plays a role in aligning the two small peptides. Lysine is one of the amino acids that has yet to be synthesized under plausible prebiotic conditions. So even if lysine was present in the soup, its concentration would have been negligible.

Every possible strategy of interference is employed by this investigator to promote replication. This mixture of peptides has almost no chance of existing on the primitive earth. Even if it did, as soon as the supply of 15 and 17 amino acid peptides runs out, the replication stops. Despite all of this interference, the claim of self replication is not valid. Self replication involves a system that can duplicate all of its components. In this system, the self replicating peptide is supplied with one peptide containing 15 amino acids and one with 17 amino acids. A true self replicating molecule could generate these two smaller peptides from the amino acids in the primordial soup.

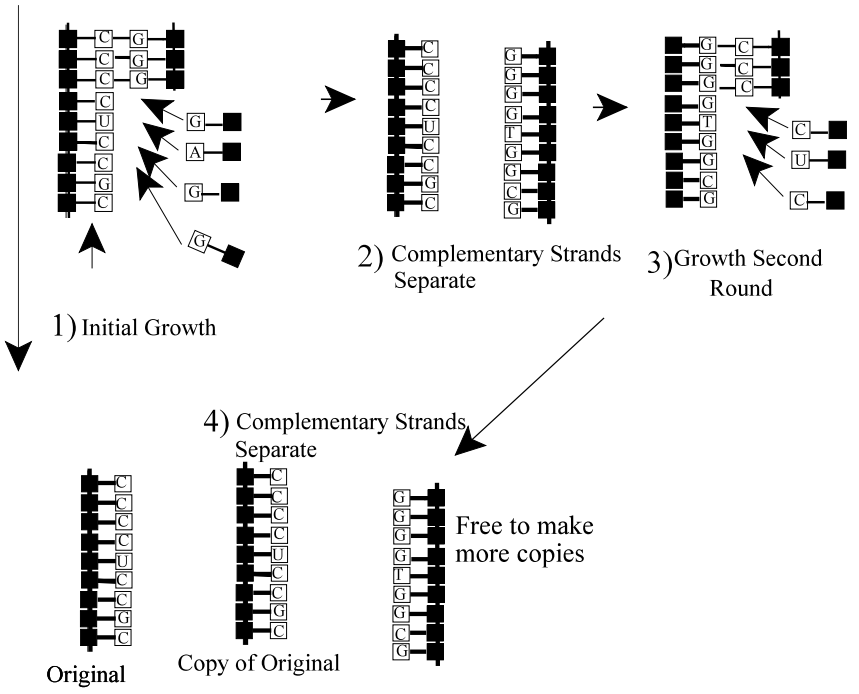
The authors of this paper tried to use dynamite to blow up the door in figure 9.4, but the door withstood the blast and did not open. So the authors just claimed that it opened.

Proteins do not self replicate, and this explains why most scientists rejected the self replicating protein hypothesis in favor of the self replicating RNA hypothesis.

RNA Self Replication

Conceptually RNA should be able to self replicate without the help of proteins. This is shown in figure 10.1. The original strand serves as a template. New base pairs arrive and form weak bonds with their complement. Adenine can form a bond with Uracil, and Guanine can form a bond with Cytosine. After one replication, two complementary strands exist. Another round of replication is necessary to duplicate the original strand. The complement to the original strand is also free to make more copies.

Figure 10.1: Conceptual Model for RNA Self Replication



On paper, this model is great. Nevertheless, it does not work in the lab. The problems were described by Joyce and Orgel as follows:

- Most strands of RNA are unsuitable templates. The original RNA molecule that serves as the template must contain a very high concentration of cytosine to make process 1 in figure 10.1 viable.[2,3] This situation is unlikely to be met because as discussed earlier cytosine has no plausible prebiotic synthesis pathway and it decays rapidly. Nevertheless, the original strand depicted in figure 10.1 meets the high C requirement.

- The chain will not grow correctly unless a very specific activation agent is used to activate the nucleotides. The activation agent of choice is not ATP (GTP, CTP or UTP). While life uses these, if these activation agents are used without proteins the phosphate groups usually attach to the wrong carbon atom in ribose.[2,3] ImpA, impG, impU and impC are the activation agents of choice. These activation agents contain the same side group as the amino acid histidine, which is one of the three amino acids that have not been synthesized in prebiotic experiments. Thus, it is unlikely that these activation agents were present in the primordial soup.

- The complement of the original chain will have a high G content. This is inevitable due to the requirement for high C in the original chain. This is problematic because RNA with a high concentration of guanine tends to fold up in such a way that it cannot be an effective template for replication.[2,3] Thus, the second round of growth in figure 10.1 does not happen.

- If different isomers of ribose are present, these isomers will terminate the growing chain.[2,3]

Joyce and Orgel comment that "In light of the available evidence, it seems unlikely that a pair of complementary sequences can be found each of which facilitates the synthesis of the other . . . " [3]

Just to add to the difficulties, if too many steps form in the replication ladder (complementary bonds between base pairs), then the strands will never separate.[4] Furthermore, figure 10.1 is oversimplified in that it does not show that in order for the RNA strands to grow, an RNA enzyme is required to catalyze the reaction. Because a growing chain cannot catalyze its own replication, two identical RNA molecules must arise simultaneously in the soup. Each capable of replicating the other.

A pattern is beginning to emerge for the RNA world. The RNA world is a speculative world without proteins where RNA is the most important molecule. RNA regulates all chemical reactions and contains all of the molecular knowledge for life. The pattern that is emerging is that perhaps this world is too speculative in that it may have never existed.

Again Joyce and Orgel put it best: "Scientists interested in the origins of life seem to be divided neatly into two classes. The first, usually but not always molecular biologists, believe that RNA must have been the first replicating molecule and that chemists are exaggerating the difficulties of nucleotide synthesis . . . The second group of scientists are much more pessimistic. They believe that the de nova appearance of oligonucleotides on the primitive earth would have been a near miracle. The author's subscribe to this latter view. Time will tell which is correct." [3]

One last point, RNA replication in the lab makes use of extensive investigator interference. Chemicals like amino acids, aldehydes, and sugars (other than ribose) are arbitrarily excluded. Very specific activation agents are used to encourage replication (ImpA for adenine, ImpG for guanine, ImpC for cytosine, and ImpU for uracil). The concentration of the chemicals (especially cytosine and ribose) is billions and billions of orders of magnitude higher than what one would expect under plausible prebiotic conditions.

Dynamite is being used to blow the door open in figure 9.4, and the door is just too solid. It remains closed and the scientist remains trapped. Fortunately, many scientists understand this, and they no longer claim that the door is open.

How Much Knowledge is Required to Create a Ribozyme

RNA molecules capable of facilitating chemical reactions do exist. Because such RNA molecules perform a role traditionally carried out only by protein enzymes, they are called ribozymes. Ribozymes have been shown to facilitate the creation of both peptide bonds in proteins, and the bonds between phosphate and ribose in RNA. This discovery is very significant in that it means RNA can both store and implement knowledge. It also explains the popularity of RNA as the first living molecule.

Bartel carried out a very relevant experiment. In this experiment. 65 ribozymes were isolated from a pool of $1x10^{15}$ RNA molecules. All ribozymes isolated contained 200 bases. This result allows for a direct calculation of the knowledge in ribozymes. If 65 sequences have some minimal enzymatic activity out of a pool containing 10^{15} random sequences, then one in every 15 trillion sequences is a ribozyme. Thus the molecular knowledge is as follows: knowledge = 3.32 x log (15 trillion) or 44 bits. Note that knowledge and not information is used because the 65 ribozymes were not yet optimized. The experiment also subjected the ribozymes to several rounds of selection in which only the best were chosen. Selection dramatically improved their catalytic efficiency.

Given the extreme difficulties associated with synthesizing an RNA molecule containing 200 or more bases, it is unlikely that even one such molecule ever existed on the primitive earth, and 15 trillion are needed to just get 65 functional ribozymes. Furthermore, ribozymes are not self replicators. The knowledge required for self replication is certainly many orders of magnitude more than the 44 bits required for a marginally functional ribozyme. Finally, the 44 bits calculated above are for evolution in a test tube where all competing side reactions are eliminated. If the primordial soup contains free amino acids, aldehydes, and undesirable isomers of ribose, then the 44 bits will increase by a factor of at least 10 and probably more.

Molecular Knowledge in the Primordial Soup

In chapter 5, the difficulties with creating a functional protein in the primordial soup were explored. A similar analysis will now be undertaken for RNA. Because of the scarcity of the RNA subunits (especially ribose and cytosine), the information content of any RNA molecule that evolves in the soup is expected to be very high.

If the soup existed, its exact composition is unknown. Nevertheless, several generalizations are possible. Ribose and cytosine should be extremely rare (see chapter 9). Furthermore, ribose will react with any free amino acids in the soup forming an insoluble polymer. Adenine can be synthesized in the lab, but not under plausible conditions with high yield. Even phosphate will be scarce if inorganic salt is present in the soup.[7]

While the concentration of cytosine and ribose in the soup is probably zero, applying information theory to this situation is not productive because infinite information, implies zero chance for success. So instead this section will make some very favorable assumptions concerning the composition of the soup. The assumptions are not realistic. They are made for educational purposes only.

Favorable Assumptions:

1) All phosphate, sugar and base molecules in the soup exist only as activated nucleotides. That is any adenine in the soup is assumed to be attached to ribose or another sugar. All sugars either have a high energy phosphate group attached or they are attached to some other activating agent.

2) No amino acids are found in the soup. While these are easily synthesized in prebiotic experiments, they must be excluded as they react quickly with ribose and other aldehydes, removing ribose from solution and preventing more ribose from forming. Amino acids and ribose cannot coexist in the soup.

3) No aldehydes exist in the soup. While these are required for the synthesis of ribose and other sugars, they cannot be allowed to persist. Aldehydes react with the four biological bases. These reactions will interfere with the formation of RNA.

Given this starting point, what is the probability that an RNA molecule will emerge from the soup? Assume the following:

- Every time an activated nucleotide attacks a ribose, it has a 50% chance of attacking the wrong carbon atom. This results in premature chain termination.[2,3]

- Half of the ribose present is the wrong isomer, this also results in premature chain termination.[2,3]

- 3/4 of the bases attached to the ribose are not biological. That is adenine, guanine, cytosine, and uracil are only used in 1/4 of the activated nucleotides. The most common base is likely ammonia or some other simple amine.

- 3/4 of the activated nucleotides use a sugar other than ribose or deoxyribose. This also results in premature chain termination. Given that ribose is usually only a minor product in any prebiotic experiment that synthesizes simple sugars, this is a very generous assumption.

Even with these most favorable assumptions that ignore all competing side reactions, every nucleotide added to the RNA chain still contributes a minimum of 6 bits of primordial information (for every 64 nucleotides added to the chain, only 1 is expected to be biologically relevant, and this corresponds to 3.32 x log (64/1) = 6 bits of primordial information). This is three times the value calculated for amino acids in chapter 5.

Thus, a 200 base pair random RNA sequence contains 6 x 200 = 1200 bits of primordial information, and as explained in chapter 5, primordial information can be related to a probability because it is a form of knowledge - the knowledge to exclude chemicals found in the soup that are not used by life today.

Thus, a 200 base pair random RNA sequence has a 1 in 2^{1200} chance of emerging in the primordial soup. Given that only 65 out of 15 trillion will exhibit any ribozyme functionality, the odds are staggering - 1 time in 3.9×10^{372} tries. Furthermore, this calculation is only for a ribozyme capable of regulating a simple chemical reaction. The odds of a self replicating ribozyme emerging are certainly much smaller.

In summary, the probability of creating a 200 base ribozyme is extremely small because so few random sequences contain the required knowledge, but given that no 200 base RNA molecules existed on the primitive earth, the odds are no longer almost zero, but instead almost zero multiplied by zero.

Finally, as noted in chapter 5, using information theory to calculate the odds has some drawbacks. Information theory only takes into account the concentration of the various chemicals. It does not have the ability to deal with chemical properties that may make certain reactions more probably, and this can skew the results in favor of evolution or against it. In the case of RNA, a very strong argument can be made that the skewing is strongly in favor of evolution. This is because the above calculation excluded amino acids and aldehydes from the soup. Thus, the information calculated above represents RNA that evolves in a test tube, not the real world.

Self Replication and Perpetual Motion

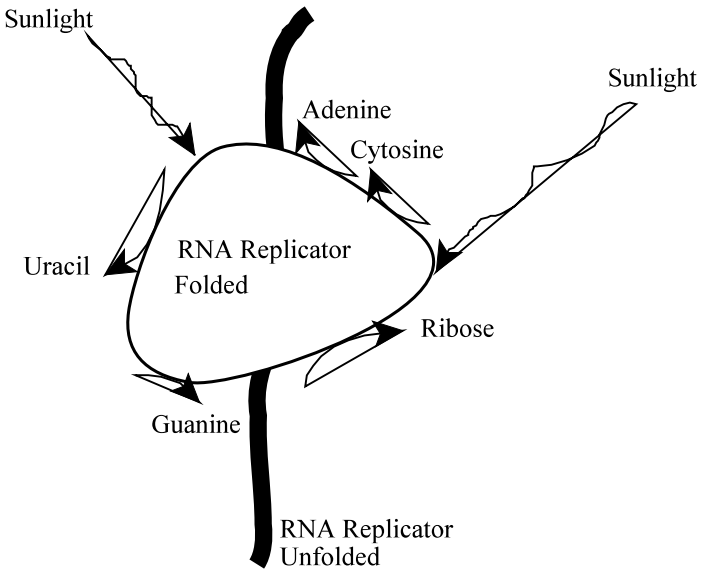
Researchers today are actively seeking and finding new ribozymes. Many are artificially engineered and others arise from random sequences. Many of these researchers believe that in time they will find a self replicating RNA molecule. Others like Joyce and Orgel who are at the forefront of the research disagree.

In chapter 7, several techniques used by life to circumvent the second law of thermodynamics were discussed. Unless a self replicating RNA molecule has the capability to implement some of these same techniques, its existence can be ruled out on purely theoretical grounds.

Based on fundamental laws of physics, science can state with certainty that if a self replicating RNA molecule is found, the molecule will only be able to replicate in a test tube. It will require a continuous supply of activated nucleotides to drive its replication. While this might work in the test tube, it would certainly not work in the primordial soup. Activated nucleotides in the soup would not last for more than a few days. Since their decay would dominate any conceivable path for prebiotic synthesis, activated nucleotides in the soup would be very rare and probably non-existent.

Given the difficulty associated with the prebiotic synthesis of ribose, adenine, and cytosine, the concentration of these critical molecules in the soup would be extremely low (probably negligible). This means that the first successful self replicating RNA molecule must be able to direct the synthesis of adenine, cytosine, ribose, uracil and guanine. If it cannot do this, it will not be able to replicate in the soup. Furthermore, it must be able to activate all of these nucleotides. So this special RNA molecule must know how to tap a plentiful energy source and use it to drive many different chemical reactions. If it cannot perform all of these functions, then it is a perpetual motion machine, and its very existence is limited to biology textbooks.

Figure 10.2: A Self Replicating RNA Molecule



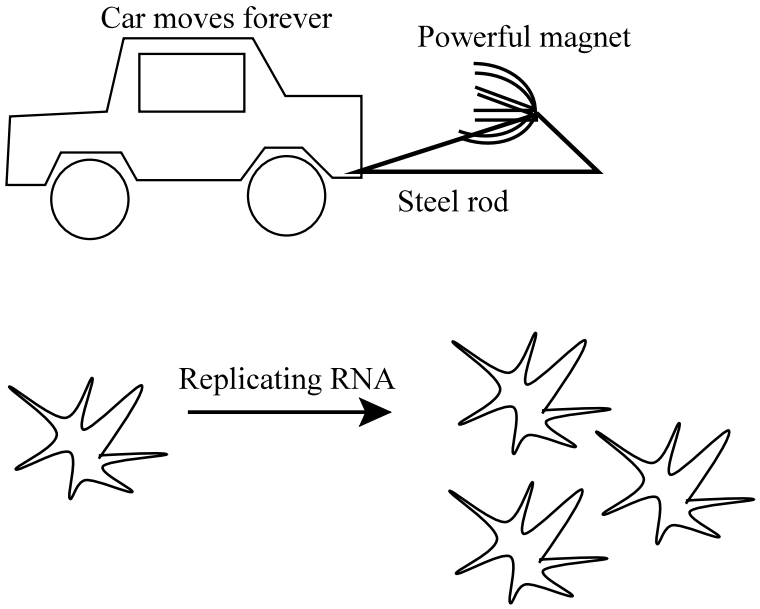
In figure 10.2, the RNA molecule can exist in two states, folded and unfolded. When folded, it catalyzed RNA replication, and the formation of adenine, ribose, cytosine, uracil, and guanine. It also must create activated nucleotides (not shown). When unfolded, it serves as a template for replication. The folded version must also know how to replicate the unfolded version.

This particular ribozyme taps into sunlight as an energy source using a primitive form of photosynthesis. Other self replicating RNA molecules could potentially oxidize a chemical like methane, hydrogen, or sulfur to generate the required energy.

Figure 10.2 is what is required of a "living molecule." Anything less is not alive. This figure was constructed with due consideration to the second law. Any RNA molecule that does not possess all of the capabilities shown in figure 10.2 is a perpetual motion machine. It may replicate in the lab as long as it is supplied with activated nucleotides, but it will not replicate in the soup. Thus, it only exists in textbooks, and there is no need to wait to see if researchers can locate it.

Inventors have been trying to invent perpetual motion machines for at least 2000 years. They have all failed. Nevertheless, many have been issued patents by various governments throughout the world. Two examples of perpetual motion are shown in figure 10.3. Both examples are equally absurd. While many scientists apparently only recognize the absurdity of the first picture, nature can recognize both, and it does not allow either to exist.

Figure 10.3: Perpetual Motion Machines



The first picture in figure 10.3 is a clear violation of energy conservation. It does not work because the force that the magnet exerts on the car is exactly cancelled by the force that the car exerts on the magnet. The magnet does not cause the car to move. The second violation is more subtle only because it violates a different law of nature. When a self replicating molecule replicates, the replication decreases the entropy of the universe. The second law is violated. To get around this problem, any real self replicator must know how and be able to couple its replication to a plentiful energy source. If it is unable to do this, then it is a special type of perpetual motion machine, and it only exists on paper and in the imagination of researchers.

References:

1) Lee, Granja, Martinez, Severin, Ghadiri, “A Self Replicating Peptide,” Letters to Nature, 382:525-528, 1996.
2) Joyce, Visser, Boeckel, Boom, Orgel, Westrenen, “Chiral Selection in Poly (C) Directed Synthesis of Oligo (G),” Letters to Nature, 310: 602-604, 1984.
3) Joyce and Orgel, The RNA World, Gesteland, Cech, Atkins, Cold Spring Harbor, “Origin of the RNA World,” 1999.
4) Bartel, The RNA World, Gesteland, Cech, Atkins, Cold Spring Harbor, “Recreating an RNA Replicase,” 1999.
5) Ekland, Szostak, Bartel, “Structurally Complex and Highly Active RNA Ligase Derived from Random RNA Sequences, ” Science, 1995.
6) Bartel and Szostak, “Isolation of New Ribozymes from a Large Pool of Random Sequences,” Science, 261:1411-1418, 1993.
7) Thaxton, Bradley, Olsen, The Mystery of Life’s Origin: Reassessing Current Theories, Philosophical Libraries, 1984.
8) Orgel, Self-organizing Biochemical Cycles, Salk Institute of Biological Studies, 99:12503-12507, 2000.
9) Green, Szostak, “Selection of Ribozyme that Functions as a Superior Template in Self Copying Reaction,” Science, 258:1910-1915, 1992.