VOLUME 81, NO. 2 JUNE 2006

THE QUARTERLY REVIEW *of* BIOLOGY

SMALL MOLECULE INTERACTIONS WERE CENTRAL TO THE ORIGIN OF LIFE

ROBERT SHAPIRO

Department of Chemistry, New York University
New York, New York 10003–6688 USA

E-MAIL: RS2@NYU.EDU

KEYWORDS

replicator, RNA world, clay minerals, plausibility, free energy, coupling, reaction cycles, metabolic network, "driver" reaction

ABSTRACT

Many scientists believe life began with the spontaneous formation of a replicator. This idea has been supported by "prebiotic" syntheses carried out by chemists using modern apparatus and purified reagents. The probability that such reactions would take place spontaneously on the early Earth is minute. These points are illustrated here by considering the often cited oligomerization of activated RNA components by clay minerals. A more likely alternative for the origin of life is one in which a collection of small organic molecules multiply their numbers through catalyzed reaction cycles, driven by a flow of available free energy. Although a number of possible systems of this type have been discussed, no experimental demonstration has been made. The inclusion of a "driver" reaction, directly coupled to the energy source, may lead to a solution.

INTRODUCTION

NUCLEIC ACIDS and proteins dominate the activities of modern cells and constitute much of their biomass. The most widely held theories regarding the origin of life have usually assumed that life began with the spontaneous appearance of a large molecule of this type. RNA, in particular, has been favored because it combines the genetic capabilities of nucleic acids with the catalytic abilities of a protein. The idea that life began with an RNA world, with "RNA molecules performing the catalytic activities necessary to assemble themselves from a nucleotide soup" (Gilbert 1986:618), has diffused into many technical articles and reviews (e.g., Bada and

The Quarterly Review of Biology, June 2006, Vol. 81, No. 2

0033-5770/2006/8102-0001$15.00

Lazcano 2002; Spirin 2002; Russell 2003; Franchi and Gallori 2004) and several textbooks (e.g., Voet and Voet 2004). Support for RNA at the start of life has been based in part on the elegance with which this assumption solves "the chicken or the egg" problem. Other evidence has been circumstantial: RNA plays many vital and central roles in the cell today (Herbert 2004). Further, nucleotide-containing cofactors are present in many cellular metabolic reactions. These nucleotide residues, which usually play no central role in the reactions, may be considered "molecular fossils" of life based on RNA (McGinness and Joyce 2003:5). However, such "fossils" may not reflect the era when life began, but could simply mark a later stage when RNA first entered ongoing metabolism (Dyson 1999).

This hypothesis does not explain how the first self-replicating RNA molecule was formed. The chemical obstacles to such an event (Cairns-Smith 1982; Shapiro 1984, 1986, 2000; de Duve 1991) appear so formidable that many scientists in the field have been convinced that "the de novo appearance of oligonucleotides on the primitive earth would have been a near miracle" (Joyce and Orgel 1999:68). For this reason, many advocates of a replicator-first origin of life have suggested that life began with a pre-RNA world (Bartel and Unrau 1999), where a simpler macromolecule exercised the genetic and catalytic functions (Joyce and Orgel 1999; Bada 2004). The genetic and catalytic properties of such systems remain largely undemonstrated, however, and the obstacles to their spontaneous abiotic synthesis are nearly as formidable as those for RNA (Shapiro 2000). Further, no trace of their earlier existence in the form of cofactors or other "molecular fossils" remains in living cells today. A recent comprehensive review summarized the situation: "The idea that RNA was 'invented' by a simpler genetic system is now a popular one, but no convincing precursor system has been described" (Orgel 2004a:116). Therefore, the RNA-first hypothesis continues to appear in many accounts of the origin of life. Philosopher and physicist Paul Davies (1995:21) has succinctly outlined "three philosophical positions concerning the origin of life: (i) it was a miracle; (ii) it was a stupendously improbable accident; and (iii) it was an inevitable consequence of the outworking of the laws of chemistry and physics, given the right conditions." The first alternative clearly falls outside the realm of science. The classification of a particular hypothesis under the third alternative depends upon its description of the "right conditions." The specifications should be few in number and not invoke unusual events. An extended series of specific reactions carried out under strictly defined experimental circumstances would belong in the second alternative. My purpose in this review is to illustrate why the RNA-first theory belongs to the second class and to present a more plausible and often overlooked alternative that fits the third category.

The experimental evidence offered to support the abiotic formation of RNA has generally been drawn from a practice called prebiotic synthesis. Publications in this area usually contain detailed descriptions of the chemistry presented, but seldom consider the likelihood of this chemistry in the context of the early Earth. In the next sections, I shall attempt to assess the plausibility of the example that is most frequently cited in support of the RNA-first theory: the clay-catalyzed synthesis of oligonucleotides from activated RNA monomers (Ferris 2002; Huang and Ferris 2003). The remainder of this review will be used to discuss an alternative: life began within a mixture of simple organic molecules, with possible participation by minerals, under the influence of a flow of available free energy. A number of suggestions for systems of this type have been made (an overview can be found in Fry 2000), but no successful demonstration has been carried out. Thoughtful critiques have been made regarding this concept, and a number of misconceptions exist concerning its nature. I shall describe the general requirements for this type of process without proposing a detailed solution. A new suggestion will be made, however, which may aid the design of future experimental demonstrations. The free energy source must be coupled to a specific "driver" reaction. This linkage can then provide the thermodynamic driving force for the conversion of the initial unorganized mixture into an organized, self-regulated metabolic network.

"Prebiotic" Synthesis in the Study of the Origin of Life: Questions of Plausibility

In a well-known experiment published by Miller (1953), an electrical spark was passed through a mixture of gases then believed to represent the composition of the atmosphere of the early Earth. The appearance of amino acids among the products raised public and scientific interest in the origin-of-life question. The experiment inspired a host of investigations that attempted to simulate other chemical reactions that may have taken place before life began. In a recent review by Orgel (2004a:2), a limited set of restrictions, for which it was claimed "there is fairly general agreement," was put forward concerning experiments of this type:

> It must be plausible, at least to the proposers of a prebiotic synthesis, that the starting materials for a synthesis could have been present in adequate amounts at the site of synthesis.
>
> Reactions must occur in water or in the absence of a solvent.
>
> The yield of a product must be 'significant,' at least in the view of the proposers of the synthesis.

As this outline may place too much power in the hands of those who propose a particular scheme, I would suggest that it be supplemented by a stipulation from an earlier work: "Quantitative equilibrium and kinetic data must be accumulated for many reactions and close attention must be paid to the geological evidence in order to define reasonable primitive earth conditions" (Miller and Orgel 1974:2). With this data in hand, some effort could then be made by objective referees to assess the plausibility of the proposed reaction sequence as a spontaneous event in natural surroundings, rather than as a controlled procedure carried out by chemists in a laboratory.

Questions of plausibility are often dismissed in the origin-of-life literature because vast amounts of time and space were available on the early Earth, and life's origin was an event that needed only a single occurrence. This argument, as presented in one geology text, declared, "How many times 10,000 trials of many such random events could have occurred within a period of 3.3 billion years? One's imagination boggles at the thought of trying to calculate so great a number. No one familiar with statistics rejects the idea of chance chemical combinations simply because there wasn't enough time. There was a huge abundance of it" (Flint 1973:119). This analysis fails, however, when a particular result requires not a single stroke of luck, but a series of improbable events. As a simple analogy, consider the phrase "to be or not to be." With advance knowledge of the desired result, I was able to type this message on my word processor in five seconds in one trial with no error. In doing this, I had to make the correct selection from 50 keys 18 times in a row, a probability of 1 in 50^{18}, or about 1 in 4×10^{30}. Current estimates of the age of the universe converge on 13.7 billion years, or about 4×10^{17} seconds. We can see that if I had sat at my word processor faithfully since the formation of the universe, but struck the keys at random, the chance of that short phrase appearing, while not zero, would be insignificantly small.

Prebiotic syntheses conducted in the laboratory often involve multistep procedures, with purified reagents and very different conditions permitted at each new step (Zubay 2000). The extensive purification procedures and changes of locale that would be needed to produce comparable results on the early Earth are seldom discussed, but must be taken into account when attempting to judge the plausibility of the entire sequence.

Clay-Catalyzed Oligonucleotide Synthesis: An Assessment of Plausibility

An extensive series of studies on the polymerization of activated RNA monomers has been carried out by Ferris and his collaborators (for reviews, see Ferris 2002; Ertem 2004). A recent publication from this group concluded with the statement: "The facile synthesis of relatively large amounts of RNA oligomers provides a convenient route to the proposed RNA world. The 35–40 mers formed are both sufficiently long to exhibit fidelity in

replication as well as catalytic activity" (Huang and Ferris 2003:1459). The first review cited above had stated this more succinctly: "The generation of RNAs with chain lengths greater than 40 mers would have been long enough to initiate the first life on Earth" (Ferris 2002:330).

The chemistry involved in their favored route is illustrated in Figure 1. The activated RNA monomer, adenosine 5'-phosphoro-1-methyladeninium (1-MeadpA), was simply allowed to adsorb onto montmorillonite clay. After one to three days, the products, labeled with ^{32}P for purposes of analysis, were eluted with pyrophosphate and fractionated by gel electrophoresis. A well-separated "ladder" of oligomers of differing lengths was produced, extending to a maximum of about 40. Additional experiments involved the successful oligomerization of similarly activated uridine 5'-phosphate and a mixture of the activated uridine and adenine derivatives. This synthesis was considered by the authors to be "vastly improved" (Huang and Ferris 2003:1458) when compared to an earlier one that had used imidazole rather than 1-methyladenine as the activating group (Ferris 2002). The earlier procedure required the use of a synthetic DNA primer and the daily addition of fresh activated monomer over a period of two weeks.

Ferris's assessment has been reflected in multiple sources. For example, a recent review stated: "The potential importance of mineral-assisted catalysis is demonstrated by the montmorillonite-promoted polymerization of activated adenosine and uridine derivatives producing 25–50 mer oligonucleotides, the general length range considered necessary for primitive biochemical functions" (Bada 2004:7). In a summary of origin-of-life research, one textbook concluded: "Reactions among the molecules in the reducing atmosphere of the prebiotic Earth are thought to have formed the simple precursors from which biological molecules developed. Eventually, in reactions that may have been catalyzed by minerals such as clays, polypeptides and polynucleotides formed" (Voet and Voet 2004:36). Similar sentiments found their way into a recent article in the *New York Times:* "It is not too easy to see how the first RNA molecules could have come into existence. But a clay called montmorillonite, formed from weathered volcanic ash and familiar in many households as cat litter, has the interesting property of catalyzing the formation of RNA from its subunits" (Wade 2003:F11).

The simplicity of the one-step procedure gives it elegance. It appears that a chemical, in the presence of a common mineral, affords an RNA molecule with sufficient length of nucleotides to be at the threshold of life. My purpose in this review, however, will be to describe the string of implausibilities inherent in this procedure that would justify use of Davies's classification, "a stupendously improbable accident" for this origin of life mechanism. The implausibilities will be framed in a series of questions in six separate sections. At the start of each new question, we will assume that previous difficulties have been overcome by chance and proceed to a new problem.

1. DO NATURAL CLAYS CATALYZE THIS REACTION?

The attractiveness of this oligonucleotide synthesis rests in part in the ready availability of the catalyst. Montmorillonite is a layered clay mineral rich in silicate and aluminum oxide bonds. It is widely distributed in deposits on the contemporary Earth. If the polymerization of RNA subunits was a common property of this native mineral, the case for RNA at the start of life would be greatly enhanced. However, the "[c]atalytic activity of native montmorillonites before being converted to their homoionic forms is very poor" (Ertem 2004:567). The native clays contain bound polyvalent cations, such as Cu^{2+}, Fe^{3+}, and Zn^{2+}, that interfere with phosphorylation reactions. This handicap was overcome in the synthetic experiments by titrating the clays to a monoionic form, generally sodium, before they were used. Even after this step, the activity of the montmorillonite depended strongly on its physical source, with samples from Wyoming yielding the best results (Ferris et al. 1989; Ertem 2004). Eventually the experimenters settled on Volclay, a commercially-processed Wyoming montmorillonite provided by the American Colloid Company.

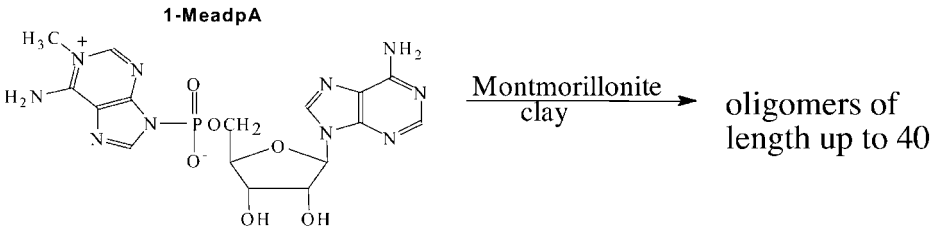

FIGURE 1. MONTMORILLONITE-CATALYZED POLYMERIZATION OF ADENOSINE 5'-PHOSPHORO-1-METHYLADENINIUM (1-MEADPA)

Oligomers of length up to 40 were produced by this reaction (Huang and Ferris 2003). Only the D-enantiomer of 1-MeadpA was employed, so the products would be expected to be homochiral. Earlier studies had indicated that both 3'-5' and 2'-5' linkages would be present in the product, with the 3'-5' links predominant (Prabahar and Ferris 1997).

Further purification steps were applied to obtain the catalyst used for the "prebiotic" formation of RNA. It is conceivable, of course, that native samples with the appropriate properties existed on the early Earth. This idea cannot be tested directly, but the present planet could be searched more thoroughly to see if native unprocessed clays with the necessary catalytic ability can be found. If they were absent, or restricted to a few very limited locales, the plausibility of this reaction on the early Earth would be greatly diminished.

One solution to this problem would be to replace montmorillonite with a different catalyst. Another research group has achieved a lesser degree of oligomerization (to 11 mers) of phosphorimidazole-activated nucleotides in partially frozen aqueous solutions in the presence of certain divalent metal cations (Kanavarioti et al. 2001). Apart from the shortness of the product, this substitution still limits the number of suitable locales for reaction on the early Earth, and introduces new implausibilities in place of those that have been removed.

2. WAS 1-METHYLADENINE AVAILABLE ON THE EARLY EARTH IN SUFFICIENT QUANTITIES TO SERVE AS AN ACTIVATING GROUP FOR OLIGONUCLEOTIDE SYNTHESIS?

Nucleotides do not self-condense in aqueous solution unless they are converted to an activated form containing a high-energy bond. The reagents of choice in the synthesis were the phosphoramidates formed from 1-methyladenine, selected after a comparison of activating groups which could support nucleotide polymerization (Prabahar and Ferris 1997). This study had set aside some alternative activators with the statement: "It appears unlikely that 4-aminopyridines or imidazoles were present in high concentrations on the primitive earth so alternative bases were investigated" (Prohabar and Ferris 1997:4332). By contrast, "[p]urines, in particular 1-methyladenine, were proposed as more plausible prebiotic activating groups that may have been present on the primitive Earth" (Huang and Ferris 2003:1458). The authors argued earlier: "Since adenine and other purines are required for the prebiotic synthesis of RNA-like monomers, it is likely that adenine or its derivatives would be present in sufficient amounts on the primitive earth to also serve as phosphate-activating groups. 1-methyladenine is one of the reaction products of methylamine with adenine while the corresponding 5'-phosphoro-1-methyladeninium derivative could have been formed by the reaction of 1-methyladenine with ATP" (Prabahar and Ferris 1997:4337).

The element of circular reasoning in the first statement can be illustrated by an apt quotation from Orgel (2004b:361).

> It is possible that the RNA world was the very first organized 'biochemical world'. If this is the case, a supply of the components of RNA, nucleotides, must have accumulated *de novo* on the primitive earth. It is also possible that RNA is not a prebiotic molecule, because the RNA

> world was preceded by some other biological world, and the synthesis of nucleotides was totally dependent on the 'enzymes' of that earlier world."

Prabahar and Ferris (1997) appear to have presumed that the first alternative must be correct and ignored the second one, in which neither adenine nor 1-methyladenine were present in quantity on the prebiotic Earth.

Does existing chemical evidence support the possibility that significant amounts of 1-methyladenine were formed abiotically? The suggested path for its formation involves the conversion of adenine to N^6-methyladenine, and the rearrangement of the latter substance to 1-methyladenine (Figure 2) (Levy and Miller 1999). The question of the prebiotic availability of adenine is controversial and will not be discussed here. The reader is invited to read Shapiro (1995) and Orgel (2004b) to draw conclusions about the plausibility of the suggested mechanisms. As an example of the suggestions for abiotic adenine synthesis that have been put forward, however, I will cite one pathway taken from a textbook intended for use at the college level (Zubay 2000:240). In this account, the initials AICN, AICI, and DAMN represent chemical intermediates in the route from hydrogen cyanide to adenine.

> Formation of hydrogen cyanide in the atmosphere would probably have proceeded best in a region of high volcanic activity . . . The hydrogen cyanide formed in the atmosphere would be expected to rain down because of its high solubility in water. A freshwater pond on a mountainside might have served as a convenient catch basin for the hydrogen cyanide; here the compound could have become concentrated in the winter months by partial freezing. In a cold concentrated solution, DAMN would be expected to form slowly over a period of many months. The conversion of DAMN to AICN might occur in the spring after a thaw as a stream containing DAMN irradiated by the sunlight flows down the mountainside to a second location, where a fraction of the AICN might be converted to AICA by contact with a clay at somewhat elevated temperatures (e.g., 75°C). Following this, the stream now containing both AICN and AICA could continue to flow until it reached warmer waters containing ammonium formate. The slow evaporation of this pool to dryness over the summer months should result in the efficient conversion of the two imidazoles into the corresponding purines.

While no single reaction or location in this sequence violates the possibilities of chemistry or geology, the need for them to occur in an exact order creates an implausibility comparable to that involved in generating a particular English sentence by hitting word processor keys at random.

In this article, I will bypass the question of adenine availability and proceed to the question of 1-methyladenine. Miller and Levy (1999) produced *N6*-methyladenine from adenine in yields of up to 50% by heating adenine with 20 molal methylamine/methylamine hydrochloride in a sealed ampoule at 100°C (Figure 2). The relevance of this extraordinarily elevated amine concentration to any plausible environment on the early Earth was justified by a brief reference to a "drying-lagoon or drying-beach environment" and the comment that "[a]mine hydrochlorides are very soluble compounds and would be among the last compounds to precipitate under drying conditions in a lagoon" (Levy and Miller 1999:631). The authors noted that the equilibrium between adenine and *N6*-methyladenine in the presence of equal amounts of ammonia and methylamine favored the latter by 4.5 to 1, and that a host of primary and secondary amines as well as amino acids were able to individually react with adenine under these conditions. Rates and yields were cited, rather than equilibria, so that the final product composition of a simultaneous reaction of adenine involving these substances could not be determined. If reaction times were extended sufficiently, however, it appears likely the entire set would be converted by the 100-fold slower deamination of adenine (and presumably *N6*-substituted adenines) to hypoxanthine.

The synthetic usefulness of this procedure may be considerable, but its relevance to

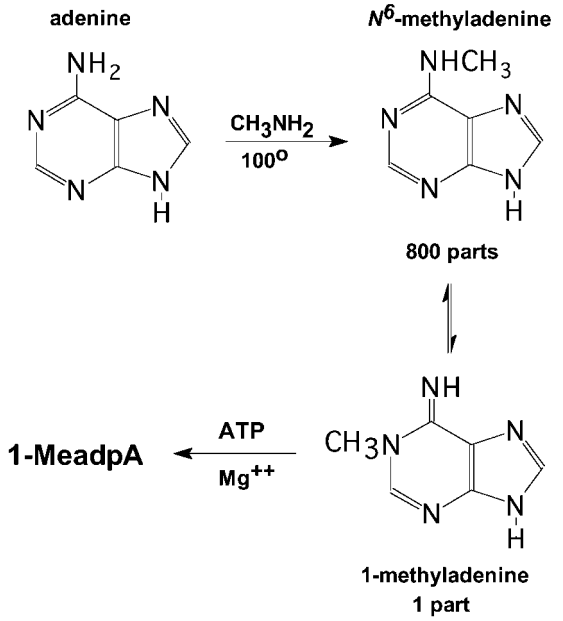

FIGURE 2. THE CONVERSION OF ADENINE TO 1-MEADPA

Adenine, heated with concentrated methylamine solution in a sealed tube, affords *N6*-methyladenine. Subsequent rearrangement provides small amounts of 1-methyladenine (Levy and Miller 1999). Huang and Ferris (2003) suggested that reaction of 1-methyladenine with ATP would afford 1-MeadpA, citing the general procedure of Lohrmann (1977). This reaction has not been carried out experimentally.

natural geochemical processes is extremely questionable. A heated, sealed ampoule can hardly be compared to an open body of water, which is subject to loss by seepage, dilution by rain, and loss of volatile components (such as amines) by evaporation. An extensive critique of the drying lagoon model as applied to a different reaction has been published (Robertson and Miller 1995; Shapiro 1999, 2002), including the observation that no drying lagoon with the properties necessary for the described reactions appears to exist on the contemporary Earth. The publications of those who propose the drying lagoon model imply, however, that a series of such lagoons existed on the Early earth, each carrying out a different reaction involving a very limited set of components.

Despite the implausibility of the above scenario, the reaction can be carried out in a laboratory. The possibility that some unusual set of geological circumstances may have brought together adenine and a temporary high concentration of methylamine on a few occasions during the history of this planet cannot be excluded by logic. Pockets containing adenine and *N6*-methyladenine may have been sequestered, like diamonds or rubies, in isolated locations on the Earth. These two compounds would not have been alone of course. Analyses of the Murchison meteorite testify to the abiotic occurrence of a suite of aliphatic amines and amino acids in that body (Cronin and Chang 1993). As there is no reason why the reaction site should contain only methylamine, a mixture of *N6*-methyladenines would be expected in those locations. However, a different substance, 1-methyladenine, was the one required for the activation of nucleotides. Levy and Miller (1999) mentioned briefly that 1-methyladenine could be formed from *N6*-methyladenine by the reversal of the well-known Dimroth rearrangement (Figure 2) (Macon and Wolfenden 1968). They added, however, that the equilibrium of the interconversion favored the latter by a ratio of 800 to 1 at neutral pH. Further details were not provided.

Thus, the prebiotic preparation of 1-methyladenine suggested by Levy and Miller (1999) and cited as justification for its availability on the early Earth would have produced that substance mixed with *N6*-methyladenine, in a 1:800 ratio. Unspecified amounts of *N6*-alkyl- and dialkyladenines as well as other *N6*-substituted adenines and, in some cases, the products of deamination or reverse Dimroth rearrangement of these substances also would have been present. The procedure for 1-MeadpA published by Prabahar and Ferris (1997), on the other hand, utilized 0.33 mmol of 1-methyladenine obtained from a commercial source, with none of the other above mentioned substances present. As the presence of *N6,N6*-dimethyladenine, for example, inhibits the clay-catalyzed oligomerization (Ertem 2004), it seems clear that some sequence of steps achieving the separation, purification, and concentration of 1-methyladenine was essential before oligomerization could succeed on the early Earth. The plausibility of such unmentioned purifications was described clearly by Cairns-Smith (1982:46):

> In organic chemistry it is often the work-up rather than the reaction that causes most of the trouble. Think about the techniques that are used: pH adjustments, solvent extractions, chromatography, evaporations to dryness, recrystallisations, filtrations and so on. Now you can say that such things might have taken place fortuitously under primitive geological conditions. Each individual operation can be imagined—a transfer of a solution, a washing of a precipitate, an evaporation, and so on. But very many such operations would have to take place consistently and in the right order. In a typical work-up procedure there are subtle things that make the difference between success and mess—how long to wait, say, after the pH adjustment before filtering. Practical organic chemistry is not easy. Very much has to be engineered. It is not sensible to suppose that an uninformed geochemistry would fortuitously be expert in such things.

The need for such a purification sequence introduces another large element of implausibility to the prebiotic relevance of the procedure of Huang and Ferris (2003).

3. HOW WOULD 1-METHYLADENINE BE CONVERTED TO 1-MEADPA ON THE EARLY EARTH?

Another item needed for the preparation of 1-MeadpA would be the 5'-phosphate, or polyphosphate of adenosine. The multiple improbabilities involved in the prebiotic conversion of adenine to adenosine monophosphates have also been described in the literature (Cairns-Smith 1982; Shapiro 1984, 1988, 1995; Orgel 2004a). For the sake of simplicity, I have restricted my analysis to the adenine nucleotide case, but comparable difficulties appear to be present for the abiotic synthesis of the other RNA nucleotides as well. The review by Orgel (2004a) summarized the situation: "The inevitable conclusion of this survey of nucleotide synthesis is that there is at present no convincing, prebiotic total synthesis of any of the nucleotides" (p 108).

To facilitate the incorporation of adenine nucleotides into oligomeric products, 1-methyladenine must be converted to 1-MeadpA. Prabahar and Ferris (1997) prepared this substance by allowing 1-methyladenine to react with 5'-AMP in the presence of a water-soluble carbodiimide (Sheehan and Cruickshank 1968). The preparation of that reagent involved the use of organic solvents and other features that would discourage any claim that it is prebiotic. Prahabar and Ferris cited, but did not use, a procedure of Lohrmann (1977) (Figure 2) to support the possibility that 1-MeadpA would be present on the early Earth. In Lohrmann's procedure, ATP was allowed to react with a variety of amines under conditions he considered prebiotic. 1-methyladenine was not one of the amines investigated, however. Its closest structural analog was imidazole, so we will examine his synthesis of adenosine-5'phosphorimidazolide (ImpA).

In that work, ImpA was prepared to a maximum yield of 17.7% from ATP, imidazole, and $Mg^{2+}$by allowing them to evaporate together on a cover glass over a number of days.

Careful control of the variables was necessary for the best results. Because there was little reaction in dilute solution, 0.625 M imidazole was used to obtain the highest yield. Lohrmann (1977) commented: "Such highly concentrated solutions may be prebiotically feasible in situations where more dilute solutions evaporate in contact with an atmosphere of low humidity. It is important, however, that the humidity is not too low, since little reaction occurs if the mixtures evaporate quickly and leave behind dry residues" (p 153). Humidity control was essential; the maximal yield was found at 70% humidity, with sharp decreases at humidities of 58% and 81%. If the air was water-saturated, the reaction failed. Control of pH was also essential since "it is known that ImpA decomposes at pH's below 8.5 in presence of Mg^{2+} salts to give high yields of AppA" (p 148). The need for careful control of the experimental variables adds another limitation. Sites on the prebiotic Earth that lacked the appropriate pH and humidity would be disqualified.

The use of high imidazole concentrations and the other features of the above reaction have motivated Ertem (2004) to comment that "activated monomers with 95% purity have only been synthesized using relatively elaborate synthetic procedures . . .they are, at best, only marginally plausible as prebiotic molecules" (p 567).

If one attempt to connect the various steps in the prebiotic sequence of reactions proposed as a source of 1-MeadpA, another requirement can be stipulated: multiple transfers of intermediates between various specific locations. Two different routes for adenine transport would be required. In one, the adenine, which presumably arose in a cyanide-rich environment, would be moved to an alternate site where ribose had been generated for the purpose of nucleoside synthesis (Shapiro 1988). Another larger portion of the adenine would avoid ribose and other sugars, as the presence of 1-substituted adenosines in RNA would disrupt Watson-Crick base pairing. It would proceed, instead, to an amine-rich environment for conversion to a mixture of *N6*- and 1-alkyl adenines. Additional transfers would be needed to enable purification. In the next step, the purified 1-methlyadenine would be moved to a locale whose pH, humidity, and salt content best approximated those of Lohrmann's open watch glass. A different transport mechanism would also be required to ensure the simultaneous arrival of the nucleotide triphosphate components of RNA. After the synthesis of the nucleotide phosphoramidates had been carried out, a final shift of the mixture to the location of a functioning montmorillonite clay would be necessary.

Transfers of material commonly take place on Earth today, and certainly must have taken place in the past. The synchronization and precision of the above moves, however, bring to mind the analogy of a military marching band. The likelihood that such transfers would have taken place spontaneously can hardly be quantified, but would appear insignificant.

4. WOULD LONG OLIGOMERS BE PRODUCED IN AN ABIOTIC POLYMERIZATION ON THE EARLY EARTH?

The oligomerization experiments conducted by Prabahar and Ferris (1997) and Huang and Ferris (2003) contained the chromatographically purified bifunctional molecule 1-Meadpa, or the analogous activated uridine nucleotide, 1-Meadpu. In an authentic abiotic simulation, the bifunctional molecules needed to construct such oligomers would be accompanied by related monofunctional ones that would terminate growing chains.

The chain termination problem has already been discussed with reference to the possible prebiotic case of amino acid oligomerization (Shapiro 2000). Using meteorite analyses as a sample of the actual mixtures produced by prebiotic organic processes, it could be seen that terminators, such as carboxylic acids and amines, were present in greater concentration than chain-extending amino acids, and that no long peptides would be expected. In fact, no polypeptides have been found, and the bulk of high-molecular weight material detected in meteorites has been described as a structurally heterogeneous material comprised essentially of cross-linked condensed homo- and heterocyclic

ring systems (Cronin and Chang 1993). A recent detailed study of this material using solid state nuclear magnetic resonance spectroscopy confirmed this conclusion, declaring that "[t]he insoluble fraction is thought to be composed of a variety of aromatic ring systems cross-linked by short methylene chains, esters, ethers, sulfides, and biphenyl groups" (Cody et al. 2002:1851). The origin of this material was unclear, and the authors speculated that it might have originated in the Solar Nebula, or perhaps formed later during the hydrothermal alteration of the parent body of the meteorite.

Such insoluble heterogeneous material (though not necessarily of the same composition) has often been produced as a contaminant in model prebiotic simulations, such as spark discharge experiments. In the original experiments of this type performed by Stanley Miller, it was observed that "[t]he compounds identified account for 15% of the carbon added to the apparatus. A substantial quantity of polymer or tar was found, perhaps a cyanide, aldehyde, or mixed cyanide-aldehyde polymer" (Miller and Orgel 1974:84). This yellow-brown insoluble material could be removed from the apparatus only by oxidation with a hot dichromate-sulfuric acid mixture (Miller 1955).

Colored, polymeric materials are the unwanted byproducts of a variety of organic processes, for example, the Maillard reaction of sugars with amino acids (Paulsen and Pfleughauft 1980). Their composition varies with the identity of the reaction used to produce them, but the word "tars" has often been used to describe such dark insoluble byproducts. Their structures differ markedly from those of the linear polymers with regular backbones, produced by loss of water from a select group of monomers, which dominate life processes today. Chemical considerations, as well as experimental observations of the types described above, indicate that polypeptides, polynucleotides, and other structurally well-defined linear polymers would be very unlikely products of abiotic organic chemistry.

In the case of a genetic polymer linked by phosphodiester bonds, it is expected that a host of simple alcohols would be present and terminate nascent chains. Methanol, ethanol, ethylene glycol, serine, glycolic acid, and an armada of other aliphatic and aromatic alcohols immediately come to mind. Ethylene glycol, in particular, has a particular affinity for the interlayer regions of montmorillonite (Cairns-Smith 2001) and should act as an avid competitor in any phosphorylation process. In addition, a phosphorylation process that converted nucleosides to nucleotides should operate on these simpler alcohols as well. The resulting alcohol phosphates could be converted to phosphoramidates by the same processes that activated the RNA monomers, and would compete for the hydroxyl groups of growing oligonucleotide chains.

A related type of interference would be provided by the presence of nucleotide 3'-phosphoramidates (in addition to the 5'-substituted ones such as 1-MeadpA) in the oligomerization mixture. Reimann and Zubay (1999) had attempted to identify a specific prebiotic phosphorylation reaction that would avoid this complication: "Since we are primarily interested in the 5'-isomer of the nucleoside monophosphate conditions were sought under which formation of the 5'-NMP was favored over the 3'-NMP" (p 240). Their best result was 68% 5'-AMP to 32% 3'-AMP, which was produced by careful control of temperature, reaction time, and the concentration of catalytic urea. If this proportion were to carry through the activation process into the polymerization mixture, a highly irregular product containing 5'→5' links and 2'-3' cyclic phosphates, in addition to normal 5'→3' bonds, would be expected. Albert Eschenmoser has commented that "the chances for autocatalytic replication of irregular polymeric information carriers would appear to be drastically reduced on mechanistic grounds" (Eschenmoser and Kisakürek 1996:1253).

5. IS A MIXTURE OF RNA 40-MERS, FORMED IN AN UNDIRECTED ABIOTIC POLYMERIZATION, LIKELY TO CONTAIN A FUNCTIONAL RNA-DEPENDENT RNA POLYMERASE?

In his recent comprehensive review, Orgel (2004a:112) concluded: "If one believes that

the RNA World was the first organized biological world, one must postulate that a library of RNA strands with different sequences formed spontaneously on the primitive Earth and that this family of sequences included catalysts able to support self-replication of RNA. The idea of an RNA that performed some of the functions of an RNA polymerase is, therefore, an essential feature of the *de novo* RNA world hypothesis." Such a substance should have the ability to copy itself with good fidelity and specificity from a pool of activated RNA monomers, leading to an exponential increase of its numbers. The pool of "daughters" produced would be subject to mutation and competition with one another, initiating Darwinian evolution. Huang and Ferris (2003) cited Szostak and Ellington (1993) and Joyce and Orgel (1999) in support of the idea that a 40-unit RNA molecule could meet these requirements. Szostak and Ellington had noted (1993:513) that "Somewhat longer random regions, of 30–60 nucleotides, allow the formation of more complex structures and motifs that may be essential for function," but these were regions considered as part of longer oligonucleotides that they subjected to in vitro selection procedures. Joyce and Orgel (1999:11) made the more positive, if guarded, claim that "a triple stem-loop structure, containing 40–60 nucleotides, offers a reasonable hope of functioning as a replicase ribozyme." Later in the same article, however, the authors stated clearly:

> In the above discussion we have tried mightily to present the most optimistic view possible for the emergence of an RNA replicase ribozyme from a soup of random polynucleotides. It must be admitted, however, that we do not consider this model to be very plausible. Our discussion has focused on a straw man: the myth of a small RNA molecule that arises de novo and can replicate efficiently and with high fidelity under plausible prebiotic conditions. Not only is such a notion unrealistic in light of our current understanding of prebiotic chemistry, but it should strain the credulity of even an optimist's view of RNA's catalytic potential.

We should note that natural ribozymes of that length function in processes of self-cleavage and phosphodiester bond hydrolysis (Doudna and Cech 2002), but such substances would hinder rather than promote the chances of forming longer functional RNA molecules.

The latter goal has been approached by elaborate in vitro selection procedures carried out in modern laboratories by skilled molecular biologists (Joyce 2004). A recent achievement of this biotechnology produced a 189-nucleotide ribozyme polymerase that was capable of extending a primer by 14 nucleotides (Johnston et al. 2001). The ribozyme was constructed by an extensive selection procedure that started with a pool of 10^{15} RNA molecules. The sequences in that starting pool were only random in part, with extensive sections dictated by the results of previous selection experiments. Fidelities in the range of 97-98% were observed, with the substrates limited to the normal nucleotide triphosphate building blocks of RNA. In modern cells, this limited offering of building blocks is produced by metabolism, and the same limitation was achieved in the in vitro experiments by the action of the experimenters (Shapiro 2000). An abiotic process capable of producing activated derivatives of RNA monomers would be under no such constraints, and a host of similarly competing analogs would also be expected to be present in the polymerization mixture. An abiotic process that produced adenosine should also produce nucleosides derived from other D- and L-pyranose and furanose sugars, carrying a variety of heterocyclic bases, attached by α- or β-linkages (Shapiro 1984; Joyce 2002). This mixture would be expected to afford a comparable mixture of activated nucleotides. A very elaborate and discriminatory polymerase, with capabilities beyond those of any known protein or ribozyme polymerase, would be needed to accurately select the correct four activated nucleotides from this array.

6. COULD ANY PRODUCT OF ABIOTIC POLYMERIZATION OF RACEMIC OLIGONUCLEOTIDES EXHIBIT BOTH CATALYTIC AND GENETIC FUNCTIONS?

The procedure of Huang and Ferris (2003) used derivatives of natural D-ribonucleotides. Presumably any process that led to the formation of D-nucleotides on the early Earth would have produced L-nucleotides as well. In earlier work, the same group had investigated montmorillonitecatalyzed oligomerization of a mixture of D- and L-adenosine nucleotides (Joshi et al. 2000). The dimer fraction was analyzed, and both homochiral (D,D and L,L) and heterochiral (D,L and L,D) dimers were found in a 6:4 ratio. Greater homochiral selection was observed in a control reaction with no clay present. The authors speculated that greater selectivity might be observed in longer oligomers formed under clay catalysis, but no experimental evidence was offered to support this idea. In the absence of such evidence, the most prudent assumption is that each new nucleotide added to the growing oligomer would have a 40% chance of having the opposite chirality as the previous one. For any specified 40-mer sequence, about 2^{40}, or 10^{12} stereoisomers should exist. The most probable would be the all D- or all L-stereoismers, but each should exist to the extent of about 1 in 10^9 parts in the mixture. The remainder would be D- and L-hybrids.

The biochemical function of an oligomer of well-defined sequence but mixed chirality is unclear. Some experiments have been performed on the oligomerization of activated D-GMP derivatives on oligo(dC) templates, which contain L- and D-dC residues (Kozlov et al. 1998). The presence of a few L-dC residues in the template could be tolerated without forming a complete block to replication. When the activated monomer was racemic as well, however, the results were far less satisfactory, a problem termed "enantiomeric cross inhibition" (Joyce et al. 1984).

The effect of the substitution of an L-nucleotide for a D-residue (or vice versa) on a functional ribozyme polymerase should be even more harmful, however. Such a substitution, at least in sensitive locations, should produce a pronounced change in secondary structure resulting in reduction or loss of activity. In a mostly D-replicase that contains L-residues in specified locations, its secondary structure, and hence its function as a ribozyme polymerase, should depend upon its chirality, residue by residue. A ribozyme polymerase of mixed chirality at the start of life would then be burdened by a formidable number of tasks: (1) it must select only its own sequence, or a closely related one, for copying, out of a forest of alternatives; (2) as I discussed above, it must copy its own sequence precisely, rejecting numerous malfunctional alternative components (modern DNA polymerases are assisted by a number of repair enzymes in this process); and (3) the nucleotide selected for incorporation must not only have the correct base and sugar, but also the appropriate chirality. This task would again seem formidable for a single catalyst of any size, let alone a small oligonucleotide chain.

An additional complication exists. Another source of stereochemical variance lies in the lack of complete regioselectivity in the formation of 3',5'-linked oligonucleotides among the products. Although the regioselectivity for the natural 3'-5' internucleotide bond is favorable (86%) with 1-methyladenine in the activating group (see Figure 1) (Prabahar and Ferris 1997), and 2',5' linkages do not necessarily block template-directed replication (Ertem and Ferris 1996), the possibility still exists that the substitution of a 2',5' linkage in a vulnerable position of a ribozyme might alter its secondary structure in a way that diminished or abolished its activity.

Joyce and Orgel (1999) have considered RNA 40-mers and attempted to estimate the weight of random RNA oligomers that would contain two identical copies of a single self-replicating RNA molecule (one molecule would be needed for the template, the other as the ribozyme catalyst). Allowing for the possibility that many possible 40-mers with the right properties might exist, they still concluded that "a much larger library, consisting of 10^{48} RNAs and weighing 10^{28} grams would be required. This amount is comparable to the mass of the earth" (Joyce and Orgel 1999:61). The stereochemical complications

discussed here serve to render that implausible event even more unlikely, by vastly increasing the number of different 40-mers.

Prebiotic Biopolymers Are Extremely Implausible

The above considerations permit the abiotic synthesis of a functional RNA replicator on the early Earth only as an extremely improbable event, one that fits the second category listed by Davies (1995). Gifted experimentalists, in multiyear efforts that used the cutting edge tools of biotechnology, have not yet prepared such a molecule (Joyce 2004). As neither chemists nor laboratories were present on the early Earth, an immense stroke of luck would be needed to achieve this result. Jacques Monod (1971:144-146) has embraced this possibility:

> The present structure of the biosphere far from excludes the possibility that the decisive event occurred *only once*. Which would mean that its *a priori* probability was virtually zero . . . If it was unique, as may perhaps have been the appearance of life itself, then before it did appear its chances of doing so were infinitely slender. The universe was not pregnant with life nor the biosphere with man. Our number came up in the Monte Carlo game."

However, a more plausible origin theory would better suit the spirit of science, in that it could be tested experimentally. Christian de Duve (1991:112) expressed this position with eloquence: "life arose through the succession of an enormous number of small steps, almost each of which, given the conditions at the time, had a high probability of happening. This assumption simply amounts to a rejection of improbabilities so incommensurably high that they can only be called miracles, phenomena that fall outside the scope of scientific inquiry."

An Alternative Position: Life Began in a Mixture of Small Organic Molecules

As we have seen, replicator theories face difficulties in fundamental chemistry that place them in the category of extremely improbable events. The appearance of such a substance would require the combination of diverse chemicals in a long sequence of reactions in a very specific order, interspersed with a number of complicated separations, purifications, and changes of physical location (Cairns-Smith 1982). If we wish a more plausible origin of life, then we must work with the assumption that life began, somehow, among one of the mixtures of simple organic molecules that are produced by abiotic processes. The only natural examples that we have in our hands at this time are the components of meteorites that have fallen to Earth. Through spectroscopy, we have also gained partial lists of the organic molecules in interstellar space and interplanetary dust clouds.

For a mixture of this type to move in the direction of life, a process of self-organization would be necessary. This process would enhance the concentration of certain components of the mixture, either at the expense of others, or by new synthesis from raw materials, such as carbon monoxide or carbon dioxide. An external source of free energy would be needed to drive these changes, which otherwise would involve negative overall entropy. This approach has often been called "metabolism first," as the absence of a genetic polymer has been equated with the lack of any mechanism of heredity. However, a transformed mixture of this type can be considered to hold hereditary information, which would be represented by the identity and concentration of its constituents. The term "compositional genome" has been coined to describe this system in which genetic information is not stored in a list, as in DNA, but is represented by the presence or absence of organic components (Segré and Lancet 2000). Evolution would be represented by changes in the composition of the system and in the reactions used to sustain it, in response to changes in the surrounding environment. Growth of the system would take place through the acquisition or synthesis of additional quantities of the key components, and reproduction would occur when physical forces split the enlarged system into two or more fragments. When competi-

tion for material and energy was established between these fragments, then Darwinian natural selection would enter the picture as a driving force for further evolution.

These concepts are not new. They were initiated in the ideas of Alexander Oparin in the first half of the twentieth century, and a limited but accurate scheme of this type was presented by Eakin (1963). A summary of the ideas of a number of contributors, including Gunther Wächtershäuser, Harold Morowitz, Christian de Duve, Stuart Kauffmann, Freeman Dyson, Michael Russell, Doron Lancet, and others can be found in Fry (2000). These participants disagree, however, on specific details of the type of energy to be utilized, the source of the organic raw materials, the identity of responsive chemical system, and the most suitable location on the early Earth for chemical self-organization. There is general agreement that some barrier or mechanism is needed to protect the evolving entity from dispersal by diffusion, but the contributors differ on the nature of that barrier. I can provide only a brief account of the more prominent suggestions in this space, and the listed references should be consulted for more details. Some of the important variables are summarized below.

(1) The energy source: A variety of possibilities were available on the early Earth (Chang 1993; Deamer 1997). Morowitz (1999) has considered solar radiation and chemical redox energy to be the most significant, and favored the latter because of the difficulty of harnessing solar radiation effectively. The redox energy is derived primarily from encounters between the effluents of a reduced mantle and the more oxidized atmosphere and lithosphere produced by the loss of hydrogen to space, after the photochemical decomposition of water. Oxidized and reduced species are brought into contact by volcanism and other geological processes, producing a supply of available free energy (Smith and Morowitz 2004).

(2) The location: Historically, the suggestions have varied from Darwin's "warm little pond" to the global ocean as a gigantic "prebiotic soup." Two sites have received special attention recently: hydrothermal deep sea vents (Holm 1992) and mounds (Martin and Russell 2002) and the ocean-atmosphere interface (Chang 1993; Donaldson et al. 2004).

(3) The diffusion barrier: Much attention has been directed toward primitive amphiphile vescicles, as they self-assemble from simple components and have an obvious ancestral connection with the more complex membranes that enclose modern cells (for review, see Segré et al. 2001; Monnard and Deamer 2002; Hanczyc et al. 2003). Other prominent alternatives that would limit loss by diffusion have included electrostatic forces at mineral surfaces (Wächtershäuser 1992), iron sulfide membranes (Russell et al. 1994), and aerosols at the ocean-atmosphere interface (Donaldson et al. 2004).

(4) The source of organic materials: Possible sources cited have included mineral-catalyzed hydrothermal synthesis, atmospheric syntheses driven by radiation or electric discharges, and delivery from outer space. For a summary of these alternatives, see Cleaves and Chalmers (2004).

(5) The reactive chemical system: Some scientists have attempted to specify key components, but not the entire reactive system (e.g., de Duve 1991, 2003; Weber 2001, 2002). Others have specified complete chemical cycles. Modified versions of the reductive citric acid cycle, a carbon-fixation pathway that is used by several organisms today, have frequently been invoked (Wächtershäuser 1992; Morowitz 1999; Cody et al. 2001; Lindahl 2004; Smith and Morowitz 2004).

Why Should a System of This Type Work At All?

The general form of a primitive metabolic cycle that emerges from such descriptions is illustrated in Figure 3, where each letter within a circle (or "node") represents one or more carbon-containing compounds. In this illustration (and the following one), I have not attempted to specify a particular set of chemicals that would initiate a process of self-organization, but rather to supply a framework into which many of the suggested schemes could fit. The number of nodes illustrated here is arbitrary. We will assume that the clockwise direction around the cycle is associated with a net decrease in free energy,

and should proceed spontaneously. To assure that the reactions will occur at a reasonable rate, it is assumed that the components of the cycle will cross-catalyze the reactions of the cycle. For example, product E may serve as a catalyst for the B to C conversion. This behavior has been implied by the use of phrases such as "autocatalytic self-organizing cycles" to describe such systems. Further, several of the steps involve the incorporation of a simple and abundant carbon-containing compound, such as carbon dioxide, into the cycle. In this illustration, carbon fixation has been indicated for the conversion of A to B and of C to D. In some schemes, the number of carbon atoms added in one turn of the cycle equals the number present in the compound designated as the starting point. The material within the cycle will then double with each turn of the cycle, leading to its rapid growth.

These assumptions have been challenged, however. Orgel (2000:12504) has pointed out that "each step of a proposed cycle must proceed at a reasonable rate, and that this will often depend on the availability of a suitable catalyst." He questions why, fortuitously, a participant in a particular metabolic cycle should catalyze other reactions in that cycle, or why a particular mineral should catalyze the suite of reactions in a cycle rather than other processes that would disrupt the cycle. One such process is indicated in Figure 3 by the diversion of E to form "tar" (insoluble materials that precipitate and are thus withdrawn from the cycle; such materials have been described under the subheading 4. WOULD LONG OLIGOMERS BE PRODUCED IN AN ABIOTIC POLYMERIZATION ON THE EARLY EARTH?). Pross (2004) has noted that no experimental evidence exists to support the spontaneous formation of such a metabolic cycle. In the case of the reductive citric acid cycle, efforts to "ignite" the cycle from a mixture of its components have not yet succeeded, and the suggestion has been made that a long induction period, determined by chance, may have been necessary (Wächtershäuser 1992; Cody 2004). Pross has also noted that the spontaneous emergence of a far-from-equilibrium organized metabolic system, in the absence of any identifiable driving force, would be contrary to the Second Law of Thermodynamics.

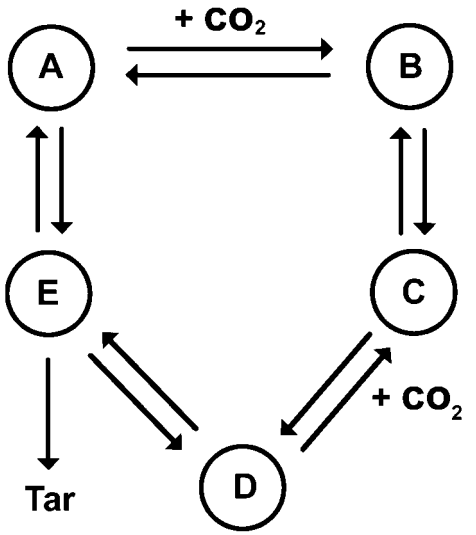

FIGURE 3. A SIMPLE AUTOCATALYTIC CYCLE
The chemicals involved are represented by the letters A–E. It is presumed that the members of the cycle catalyze the reactions in which they are not directly involved. The carbon gained by carbon dioxide fixation within the cycle must exceed that lost through irreversible reactions (indicated as "tar") and diffusion, in order for the cycle to exhibit growth.

A MODIFIED METABOLIC SCHEME

I will argue that these objections can be met in principle by the introduction of a limited number of assumptions: (1) A thermodynamically favorable, irreversible "driver" reaction that is directly coupled to an external source of available free energy can occur in a plausible abiotic setting. (2) A multistep reversible pathway is possible that converts the product of the driver reaction back to the starting material, completing a cycle. (3) The cycle functions at a "profit" within its environment; the gain of carbon by the cycle exceeds its loss by all mechanisms. These stipulations will be discussed individually.

(1) The "driver" reaction: This process is indicated by the reaction of A to form B in the modified cycle illustrated in Figure 4. This reaction is "coupled" to the conversion of factor X, external to the cycle, to another form Y. This latter conversion is accompanied by a highly favorable free energy change, rep-

resented in Figure 4 by the release of heat. The term "coupled" signifies that neither change occurs at an appreciable rate without the other. A need for a flow of external energy to enable the self-organization process has been mentioned by a number of writers, but the need for specific coupling has often been overlooked. The inclusion of a reaction of this type within the cycle provides a driving force and satisfies the Second Law of Thermodynamics.

Coupled reactions are quite common in the biochemistry of modern cells, where they allow a thermodynamically favorable reaction to drive an unfavorable one (Voet and Voet 2004:59). For example, if the reaction of A to B has an unfavorable free energy, it will not occur spontaneously. The coupling of this reaction to the highly favorable process of X→Y allows the combined reaction A + X → B + Y to take place; the conversion of X to Y drives the A to B conversion.

For such coupling to take place, the two transformations must share a common intermediate, which we will call Z. The two reactions can then be written as A→ B +Z and Z + X → Y. When they are combined, the result is A + X → B + Y. An example of such coupling is the conversion of a fatty acid and Coenzyme A (CoA) to the acyl-CoA, which contains a high energy thioester bond. This transformation, catalyzed by an acyl-CoA synthetase, is coupled to the hydrolysis of ATP to AMP and inorganic pyrophosphate (PPi). PPi is subsequently hydrolyzed to two phosphates (Pi) by inorganic pyrophosphate. The overall process can be written as:

Fatty acid + CoA + ATP → Acyl-CoA + AMP + 2Pi.

The thermodynamically unfavorable formation of a thioester bond has then been compensated by the hydrolysis of the two high energy phosphoanhydride bonds of ATP. The shared intermediate, an acyladenylate formed by the initial reaction of the fatty acid with ATP, does not appear in the overall reaction (Voet and Voet 2004:915).

Of course, the absence of both ATP as an energy source and enzymes to bring the reactants into intimate contact would be expected on the prebiotic Earth for reasons discussed above (see also Shapiro 1995). Naturally available materials might serve as surrogates, however. For example, pyrite (FeS_2) formation from FeS and H_2S could serve as a robust source of free energy and reducing power (Wächtershäuser 1990, 1992; Cody 2004). A mineral surface could attract charged organic substrates and bring them into proximity for reaction (Wächtershäuser 1988). For coupling to occur, however, an intermediate involved in the release of free energy would also have to appear in the mechanism of the organic conversion that is being driven. In this particular system, a bound sulfide or disulfide could satisfy the requirement.

In the above paragraph, we have pictured the conversion of A to B and of X to Y as oxidation-reduction reactions, with X and Y

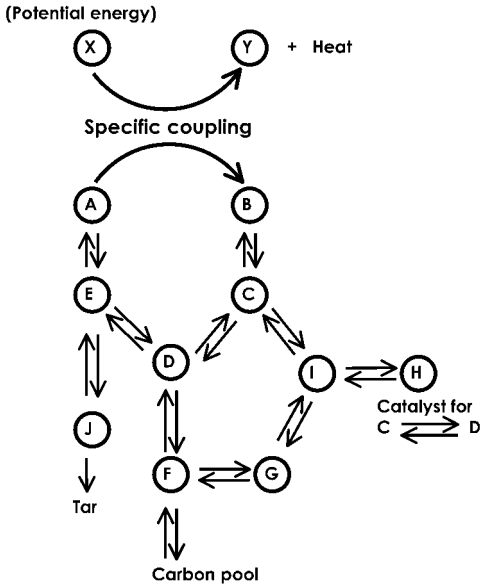

FIGURE 4. A METABOLIC NETWORK ORGANIZED BY A "DRIVER" REACTION

The thermodynamically favorable transformation of X to Y, external to a cycle, is coupled to the conversion of A to B (the "driver" reaction). This joint reaction moves the cycle represented by letters A–E in a clockwise direction. The movement of material from side reactions (indicated by equilibria involving chemicals F–J) and from an external carbon source ("carbon pool") into the cycle is favored by the release of the potential energy in the X to Y transformation.

as common inorganic substances. However, the change of X to Y could also signify the release of energy stored in a pH or concentration gradient or present in a form of radiation. The key requirement is that the available free energy stored in X performs work on the organic chemical system, in this case by driving the conversion of A to B through a suitable interactive mechanism.

(2) The closure of a cycle: The external free energy source favors the conversion of A to B. We will presume that B, however, can undergo a variety of transformations (e.g., dehydration, aldol condensation, hydrolysis, tautomerization, oxidation, and reduction) in accord with our general experience of organic chemistry (Weber 2001, 2002). If B could revert directly to A, then the A-B system would simply be acting as a catalyst for energy release. If the system is to be capable of growth and chemical evolution, however, an indirect and reversible pathway for the return of B to A must exist to complete a cycle. This pathway is indicated in Figure 3 by the route B→C→D→E→A. (If B could not return to A by any path, then the transformation would represent a dead end, or at best, would supply material for use in a different cycle.) The continual energy-driven depletion of A should move the central cycle in the direction B→C→D→E→A. Each of the intermediates may be connected by reversible reactions to many other substances, represented in part in Figure 4 by nodes F through J. The continual energy-driven depletion of A, however, would pull material from the side branches into the central cycle. The concentrations of the cycle participants should therefore increase at the expense of these other substances. Wächtershäuser (1992, 1994), for example, described his initial system as "pyrite-pulled" metabolism, suggesting that the reductive redox energy available from the synthesis of iron pyrites would be sufficient to move a number of other reactions in a particular direction.

(3) Growth of the network: The term network will be used to represent the cycle and the side branches that can feed material into it. In practice, loss of network components through a variety of mechanisms will be inevitable. Such mechanisms could include the incorporation of network molecules into insoluble materials as we discussed above (this is indicated in Figure 4 by the path E → J → tar), the irreversible formation of unreactive substances such as alkanes (Weber 2002), and loss of substances by simple diffusion. If the network is to grow, then these losses must be compensated by the continual addition of carbon-containing material to the system. A number of proposals for autocatalytic cycles have included carbon dioxide fixation within the central cycle for this purpose. The additional carbon needed for growth could, in principle, be added at any node of the network, however. In Figure 4, a "carbon pool" has been linked arbitrarily at node F. The energy released by the "driver" reaction would pull this material into the central cycle. The ultimate carbon source could be carbon monoxide (fixed by mineral reactions) (Cody 2004), formaldehyde (produced by a variety of processes) (Weber 2002), carbon dioxide fixation by minerals (Wächtershäuser 1992; Freund et al. 2001), or some other resource. The central requirement for network survival and growth is that the system show a "profit;" the rate of carbon input must exceed its rate of loss by the mechanisms listed above.

Further Evolution of the Network

The diversity of organic chemistry, with its harvest of competing, interconnected reactions (Weber 2001, 2002), becomes an asset rather than a liability in the case of the energy-driven system that I have illustrated above. The existence of side reaction paths can provide the network with the capacity of reacting to circumstances. Let us suppose, for example, that some environmental change slows the conversion of C to D, so that this limits the overall rate of the cycle. C will accumulate and be diverted into side paths, perhaps activating an alternative pathway to D, which is illustrated in Figure 4 as proceeding through nodes I, G, and F. Another possibility is that a different path would produce an alternative substance, H, which catalyzes the C to D transformation. An element of regulation would thereby enter the network. When the concentration of C is high, H would be

produced, and catalyze the C to D conversion until the excess C dissipates. At that point, the concentration of H would drop as well.

Through such circumstances, a variety of pathways and catalysts might be discovered and added to the network. The term "grafting" has been chosen by Wächtershäuser (Lindahl 2004) to describe such additions. Many examples of monomer catalysis have been reported in the literature. Bar-Nun et al. (1994), for example, have documented that free amino acids, or combinations of them, exhibit limited β-galactosidase activity. In one experiment, they reported that the formation of a catalytic dimer could be induced by the presence of the substrate (Kochavi et al. 1997). Amines catalyze many transformations in sugar chemistry, with rate enhancements of 1,200 observed for pyruvaldehyde synthesis (Weber 2001). We mentioned above the idea that catalysis could be enhanced in some systems by the alignment of the reaction participants and catalyst along a mineral surface (Wächtershäuser 1992). Rate enhancement through alignment might also occur at an aqueous-lipid interface (Segré et al. 2001). Even with this assistance, the rate enhancements observed would be orders of magnitude less than those achieved by modern enzymes. This point does not appear to be critical, however, as the role of these catalysts would be to enhance the capability of the network, rather than to make its existence possible.

Suggestions for Further Experiments

I have provided only a framework, and not a specific recipe, to illustrate how a coupled free-energy source could initiate the process of self-organization in a complex mixture of organic monomers. A detailed discussion about the possible further development of one system of this type that utilizes the general features of Wächtershäuser's surface theory (Wächtershäuser 1992), and cites specific reactions, can be found in Lindahl (2004). My hope is that many systems of this type may be possible, and have operated not only on the early Earth but in diverse planetary environments. In that case, life may be abundant in the universe, and local geochemical circumstances may have determined which specific system was the winner on the early Earth. In the long run, this question will be answered by the exploration of space, but we may get some guidance from appropriate laboratory experiments. The first priority in such experiments should be to characterize any systems that will self-organize when coupled to appropriate free energy sources. The principal initial task will be the identification of candidate driver reactions. Such reactions need not be given some great burden to carry, such as carbon dioxide fixation. In fact, transformations that accomplish the reverse could be considered, such as the decarboxylation of pyruvate to yield an acetyl thioester with reduction of an oxidized cofactor. This transformation is prominent in biochemistry today, and supported by elaborate cellular machinery. In establishing a self-organizing cycle, the loss of a carbon atom to the environment could be considered an investment, needed to provide the thermodynamic driving force. If B is to revert to A, of course, it would be necessary to return the lost carbon in an energetically permissible reaction elsewhere in the central cycle. Once a plausible driver reaction has been found, there should be no need to specify the remainder of the system in advance. If the materials and coupled energy source needed for that reaction were brought together, perhaps with an input of simple carbon-containing compounds to provide for growth, the metabolic network should then establish itself, and its identity could be determined by simple analysis. Nature will be instructing us, rather than we attempting to impose our schemes onto it. The information that we would gather by observing such a system could enable us to identify the subsequent steps, such as compartment formation, in the energy-driven self-organization process that leads to life.

ACKNOWLEDGMENTS

I wish to thank Dr Christian de Duve, of the Christian de Duve Institute of Cellular Pathology, and Professors Suse Broyde, James Canary, Nicholas Geacintov, and Mark Tuckerman of New York University for their advice concerning an early version of this manuscript, as well as Ms Ella Ewart and Mr Frank H Green for technical assistance with the figures.

REFERENCES

Bada J L. 2004. How life began on Earth: a status report. *Earth and Planetary Science Letters* 226:1–15.

Bada J L, Lazcano A. 2002. Some like it hot, but not the first biomolecules. *Science* 296:1982–1983.

Bar-Nun A, Kochavi E, Bar-Nun S. 1994. Assemblies of free amino acids as possible prebiotic catalysts. *Journal of Molecular Evolution* 39:116–122.

Bartel D P, Unrau P J. 1999. Constructing an RNA World. *Trends In Biochemical Sciences* 24:M9–M13.

Cairns-Smith A G. 1982. *Genetic Takeover and the Mineral Origins of Life.* Cambridge and New York: Cambridge University Press.

Cairns-Smith A G. 2002. The origin of life: clays. Pages 169–192 in *Frontiers of Life,* Volume 1, edited by D Baltimore et al. San Diego (CA): Academic Press.

Chang S. 1993. Prebiotic synthesis in planetary environments. Pages 259–299 in *The Chemistry of Life's Origins,* edited by J M Greenberg et al. Dordrecht (The Netherlands): Kluwer Academic Publishers.

Cleaves H J, II, Chalmers J H. 2004. Extremophiles may be irrelevant to the origin of life. *Astrobiology* 4:1–9.

Cody G D. 2004. Transition metal sulfides and the origins of metabolism. *Annual Review of Earth and Planetary Sciences* 32:569–599.

Cody G D, Alexander C M O'D, Tera, F. 2002. Solid-state (^{1}H and ^{13}C) nuclear magnetic resonance spectroscopy of insoluble organic residue in the Murchison meteorite: a self-consistent quantitative analysis. *Geochimica et Cosmochimica Acta* 66:1851–1865.

Cody G D, Boctor N Z, Hazen R M, Brandes J A, Morowitz H J, Yoder H S, Jr. 2001. Geochemical roots of autotropic carbon fixation: hydrothermal experiments in the system citric acid, H_2O-($\pm$FeS)-($\pm$NiS). *Geochimica et Cosmochimica Acta* 65:3557–3576.

Cronin J R, Chang S. 1993. Organic matter in meteorites: molecular and isotopic analyses of the Murchison meteorite. Pages 209–258 in *The Chemistry of Life's Origins,* edited by J M Greenberg et al. Dordrecht (The Netherlands): Kluwer Academic Publishers.

de Duve C. 1991. *Blueprint for a Cell: The Nature and Origin of Life.* Burlington (NC): N Patterson.

de Duve C. 2003. A research proposal on the origin of life. *Origins of Life and Evolution of the Biosphere* 33:559–574.

Deamer D W. 1997. The first living systems: a bioenergetic perspective. *Microbiology and Molecular Biology Reviews* 61:239–261.

Davies P. 1995. *Are We Alone?: Philosophical Implications of the Discovery of Extraterrestrial Life.* New York. BasicBooks.

Donaldson D J, Tervahattu H, Tuck A F, Vaida V. 2004. Organic aerosols and the origin of life: an hypothesis. *Origins of Life and Evolution of the Biosphere* 34:57–67.

Doudna J A, Cech T R. 2002. The chemical repertoire of natural ribozymes. *Nature* 418:222–228.

Dyson F. 1999. *Origins of Life.* Revised Edition. Cambridge and New York: Cambridge University Press.

Eakin R E. 1963. An approach to the evolution of metabolism. *Proceedings of the National Academy of Sciences USA* 49:360–366.

Ertem G. 2004. Montmorillonite, oligonucleotides, RNA and origin of life. *Origins of Life and Evolution of the Biosphere* 34:549–570.

Ertem G, Ferris J P. 1996. Synthesis of RNA oligomers on heterogeneous templates. *Nature* 379:238–240.

Eschenmoser A, Kisakürek M V. 1996. Chemistry and the origin of life. *Helvetica Chimica Acta* 79:1249–1259.

Ferris J P. 2002. Montmorillonite catalysis of 30–50 mer oligonucleotides: laboratory demonstration of potential steps in the origin of the RNA world. *Origins of Life and Evolution of the Biosphere* 32:311–332.

Ferris J P, Ertem G, Agarwal V. 1989. Mineral catalysis of the formation of dimers of 5'-AMP in aqueous solution: the possible role of montmorillonite clays in the prebiotic synthesis of RNA. *Origins of Life and Evolution of the Biosphere* 19:165–178.

Flint R F. 1973. *The Earth and Its History.* New York: Norton.

Franchi M, Gallori E. 2004. Origin, persistence and biological activity of genetic material in prebiotic habitats. *Origins of Life and Evolution of the Biosphere* 34:133–141.

Freund F, Staple A, Scoville J. 2001. Organic protomolecule assembly in igneous minerals. *Proceedings of the National Academy of Sciences USA* 98:2142–2147.

Fry I. 2000. *The Emergence of Life on Earth: A Scientific Overview.* New Brunswick (NJ): Rutgers University Press.

Gilbert W. 1986. Origins of life: the RNA world. *Nature* 319:618.

Hanczyc M M, Fujikawa S M, Szostak J W. 2003. Experimental models of primitive cellular compartments: encapsulation, growth, and division. *Science* 302:618–622.

Herbert A. 2004. The four Rs of RNA-directed evolution. *Nature Genetics* 36:19–25.

Holm N G. 1992. Why are hydrothermal systems proposed as plausible environments for the origin of life? *Origins of Life and Evolution of the Biosphere* 22:5–14.

Huang W, Ferris J P. 2003. Synthesis of 35–40 mers of RNA oligomers from unblocked monomers. A sim-

ple approach to the RNA world. *Chemical Communications:* 1458–1459.

Johnston W K, Unrau P J, Lawrence M S, Glasner M E, Bartel D P. 2001. RNA-catalyzed RNA polymerization: accurate and general RNA-templated primer extension. *Science* 292:1319–1325.

Joyce G F. 2002. The antiquity of RNA-based evolution. *Nature* 418:214–221.

Joyce G F. 2004. Directed evolution of nucleic acid enzymes. *Annual Review of Biochemistry* 73:791–836.

Joyce G F, Orgel L E. 1999. Prospects for understanding the origin of the RNA world. Pages 49–77 in *The RNA World*, Second Edition, edited by R F Gesteland et al. Cold Spring Harbor (New York): Cold Spring Harbor Laboratory Press.

Joyce G F, Visser G M, van Boeckel C A, van Boom J H, Orgel L E, van Westrenen J. 1984. Chiral selection in poly(C)-directed synthesis of oligo(G). *Nature* 310:602–604.

Kanavarioti A, Monnard P-A, Deamer D W. 2001. Eutectic phases in ice facilitate nonenzymatic nucleic acid synthesis. *Astrobiology* 1:271–281.

Kochavi E, Bar-Nun A, Fleminger G. 1997. Substrate-directed formation of small biocatalysts under prebiotic conditions. *Journal of Molecular Evolution* 45:342–51.

Kozlov I A, Pitsch S, Orgel L E. 1998. Oligomerization of activated D- and L-guanosine mononucleotides on templates containing D- and L-deoxycytidylate residues. *Proceedings of the National Academy of Sciences USA* 95:13448–13452.

Levy M, Miller S L. 1999. The prebiotic synthesis of modified purines and their potential role in the RNA world. *Journal of Molecular Evolution* 48:631–637.

Lindahl P A. 2004. Stepwise evolution of nonliving to living chemical systems. *Origins of Life and Evolution of the Biosphere* 34:371–389.

Lohrmann R. 1977. Formation of nucleoside 5'-phosphoramidates under potentially prebiological conditions. *Journal of Molecular Evolution* 10:137–154.

Macon J B, Wolfenden R. 1968. 1-Methyladenosine. Dimroth rearrangement and reversible reduction. *Biochemistry* 7:3453–3458.

Martin W, Russell M J. 2002. On the origins of cells: a hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes to nucleated cells. *Philosophical Transactions of the Royal Society of London B* 385:59–85.

McGinness K E, Joyce G F. 2003. In search of an RNA replicase ribozyme. *Chemistry and Biology* 10:5–14.

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

Miller S L. 1955. Production of some organic compounds under possible primitive Earth conditions. *Journal of the American Chemical Society* 77:2351–2361.

Miller S L, Orgel L E. 1974. *The Origins of Life on the Earth.* Englewood Cliffs (NJ): Prentice Hall.

Monnard P-A, Deamer D W. 2002. Membrane self-assembly processes: steps toward the first cellular life. *Anatomical Record* 268:196–207.

Monod J. 1971. *Chance and Necessity: An Essay on the Natural Philosophy of Modern Biology.* New York: Knopf.

Morowitz H J. 1999. A theory of biochemical organization, metabolic pathways, and evolution. *Complexity* 4:39–53.

Orgel L E. 2000. Self-organizing biochemical cycles. *Proceedings of the National Academy of Sciences USA* 97:12503–12507.

Orgel L E. 2004a. Prebiotic chemistry and the origin of the RNA world. *Critical Reviews in Biochemistry and Molecular Biology* 39:99–123.

Orgel L E. 2004b. Prebiotic adenine revisited: eutectics and photochemistry. *Origins of Life and the Evolution of the Biosphere* 34:361–369.

Paulsen H, Pflughaupt K-W. 1980. Glycosylamines. Pages 881–927 in *The Carbohydrates: Chemistry and Biochemistry,* Second Edition, Volume 1b, edited by W Pigman and D Horton. New York: Academic Press.

Prabahar K J, Ferris J P. 1997. Adenine derivatives as phosphate-activating groups for the regioselective formation of 3',5'-linked oligoadenylates on montmorillonite: possible phosphate-activating groups for the prebiotic synthesis of RNA. *Journal of the American Chemical Society* 119:4330–4337.

Pross A. 2004. Causation and the origin of life: metabolism or replication first? *Origins of Life and Evolution of the Biosphere* 34:307–321.

Reimann R, Zubay G. 1999. Nucleoside phosphorylation: a feasible step in the prebiotic pathway to RNA. *Origins of Life and Evolution of the Biosphere* 29:229–247.

Robertson M P, Miller S L. 1995. An efficient prebiotic synthesis of cytosine and uracil. *Nature* 375:772–774.

Russell M J. 2003. The importance of being alkaline. *Science* 302:580–581.

Russell M J, Daniel R M, Hall A J, Sherringham J A. 1994. A hydrothermally precipitated catalytic iron sulphide membrane as a first step toward life. *Journal of Molecular Evolution* 39:231–243.

Segré D, Ben-Eli D, Deamer D W, Lancet D. 2001. The lipid world. *Origins of Life and Evolution of the Biosphere* 31:119–145.

Segré D, Lancet D. 2000. Composing life. *EMBO Reports* 1:217–222.

Shapiro R. 1984. The improbability of prebiotic nucleic acid synthesis. *Origins of Life* 14:565–70.

Shapiro R. 1986. *Origins. A Skeptic's Guide to the Creation of Life on Earth.* New York: Summit.

Shapiro R. 1988. Prebiotic ribose synthesis: a critical analysis. *Origins of Life* 18:71–85.

Shapiro R. 1995. The prebiotic role of adenine: a critical analysis. *Origins of Life and Evolution of the Biosphere* 25:83–98.

Shapiro R. 2000. Hypothesis paper: a replicator was not involved in the origin of life. *IUBMB Life* 49:173–176.

Shapiro R. 2002. Comments on 'Concentration by Evaporation and the Prebiotic Synthesis of Cytosine.' *Origins of Life and Evolution of the Biosphere* 32:275–278.

Sheehan J C, Cruickshank P A. 1968. 1-Ethyl-3-(3-dimethylamino)propylcarbodiimide hydrochloride and methiodide. *Organic Syntheses* 48:83–86.

Smith E, Morowitz H J. 2004. Universality in intermediary metabolism. *Proceedings of the National Academy of Sciences USA* 101:13168–13173.

Spirin A S. 2002. Omnipotent RNA. *FEBS Letters* 530: 4–8.

Szostak J W, Ellington A D. 1993. In vitro selection of functional RNA sequences. Pages 511–533 in *The RNA World: The Nature of Modern RNA Suggests a Prebiotic RNA,* edited by R F Gesteland and J F Atkins. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press.

Voet D, Voet J G. 2004. *Biochemistry.* Third Edition. Hoboken (NJ): Wiley.

Wächtershäuser G. 1988. Before enzymes and templates: theory of surface metabolism. *Microbiological Reviews* 52:452–484.

Wächtershäuser G. 1990. Evolution of the first metabolic cycles. *Proceedings of the National Academy of Sciences USA* 87:200–204.

Wächtershäuser G. 1992. Groundworks for an evolutionary biochemistry: the iron-sulphur world. *Progress in Biophysics and Molecular Biology* 58:85–201.

Wächtershäuser G. 1994. Life in a ligand sphere. *Proceedings of the National Academy of Sciences USA* 91:4283–4287.

Wade N. 2003. How did life begin? *The New York Times* 11 November:F11.

Weber A L. 2001. The sugar model: catalytic flow reactor dynamics of pyruvaldehyde synthesis from triose catalyzed by poly-l-lysine contained in a dialyzer. *Origins of Life and Evolution in the Biosphere* 31:231–240.

Weber A L. 2002. Chemical constraints governing the origin of metabolism: the thermodynamic landscape of carbon group transformations under mild aqueous conditions. *Origins of Life and the Evolution of the Biosphere* 32:333–357.

Zubay G. 2000. *Origins of Life on the Earth and in the Cosmos.* Second Edition. San Diego (CA): Academic Press.