

RNA evolution and the origins of life

Gerald F. Joyce

The evolution of RNA is likely to have played an important role in the very early history of life on Earth but it is doubtful that life began with RNA. Consideration of what came before RNA must take into account relevant information from geochemistry, prebiotic chemistry and nucleic acid biochemistry.

THE question of life's origins is one of the oldest and most difficult in biology. The answer, if it is ever known, will not be a single statement of fact but rather an extended chronology, beginning with the formation of the Earth and ending with the appearance of cellular organisms. The problem is especially difficult because we have no direct evidence of the events that occurred during roughly the first thousand million years of the Earth's history. The oldest rocks that provide clues to life's distant past are 3.6×10^9 years old and by that time cellular life seems already to have been well established¹. Modern organisms are so sophisticated that they furnish little information about what life was like before there was a genetic code and a translation apparatus. Extraterrestrial studies have yet to provide us with an alternative life form for comparison. We are left with only a partial understanding of the origins of life that is based largely on inference and conjecture.

From time to time in the quest for such partial answers, a new experimental fact emerges that forces us to view the problem of the origins of life in a different light. The results of the classical Miller-Urey experiments, in which amino acids were produced by circulating H_2 , CH_4 , NH_3 and H_2O past an electric discharge², for example, make it hard to escape the conclusion that amino acids were present on the prebiotic Earth. The synthesis of purines by polymerization of HCN (ref. 3) similarly suggests that adenine, guanine and other purines were available. The most recent example of such a finding is the discovery of RNA molecules with catalytic activity^{4,5}: in addition to their recognized properties as templates, RNA molecules have enzymatic properties and are thus the only molecules known to function as both genotype and phenotype.

This has led to a revival of interest in the idea that there was a time, before the origin of protein synthesis, when life was based entirely on RNA. The idea is not new: it dates back eighteen years before the discovery of catalytic RNA⁶ and was discussed extensively in the late 1960s⁷⁻⁹. But the discovery of catalytic RNA caused molecular biologists to begin to take the notion of an 'RNA world' seriously¹⁰⁻¹². It is by no means a foregone conclusion that life began with the RNA world, although several lines of evidence suggest that the RNA world existed at some stage in evolutionary development. Here, I shall briefly review the advantages of RNA as the basis for the early evolution of life, before discussing the chemical obstacles to the initiation of an RNA world and suggesting how in principle they might have been overcome.

Why RNA?

A fundamental property of living systems is that they have a chemical basis for the storage of genetic information. RNA, DNA and other nucleic-acid-like compounds are informational macromolecules that have inherent template properties and so lend themselves in a straightforward way to both storing information and replicating it. It has been argued that to focus on nucleic acids is to take too parochial a view: there may have been other information storage systems with template properties that were more readily available on the primitive Earth. Cairns-Smith, for example, has proposed¹³ that life began with self-replicating clays, with the genetic information stored as a distribution of charges along a clay surface and replicated by ionic

interactions between existing and newly-formed surface layers. Other templating systems have been proposed based on organic compounds other than nucleic acids^{14,15} or a combination of minerals and associated organics¹⁶. While keeping an open mind about these possibilities, enthusiasm has to be tempered until there is experimental evidence that such an entity can both replicate and code for specific catalytic behaviours.

Could it be that the first genetic molecule did not have template properties and replicated itself in some other way? There was a time when it was widely believed that life began with self-replicating proteins¹⁷. After all, polypeptides are capable of a wide range of catalytic behaviours and it is very likely that they were produced abiotically on the primitive Earth¹⁸. The amino-acid sequence of abiotic polypeptides would have been tremendously heterogeneous, so we must postulate that some mechanism existed to ensure that a particularly useful sequence would facilitate the production of additional copies of itself. Dyson first suggested the possibility of a self-sustaining network of proteins in which the components mutually catalyse the synthesis of each other from monomeric starting materials¹⁹. Kaufmann has argued that, given a large enough assemblage of random polypeptides, the emergence of a self-sustaining network is almost inevitable²⁰. This conclusion seems to me to rest on a highly over-optimistic estimate of the probability that a random polypeptide can catalyse peptide-bond formation with any significant degree of sequence specificity. Any experimental evidence to the contrary would be welcome.

The evidence that life was based on RNA at some stage in evolutionary history is circumstantial but, I believe, persuasive. None of the following points is convincing by itself but together they constitute a body of evidence that cannot be dismissed as coincidental. First, the template properties of RNA enormously simplify the task of self-replication. RNA genomes are known to exist in certain viruses and Orgel has shown that RNA templates are able to direct the synthesis of complementary oligomers in a chemical system²¹. Second, RNA has catalytic properties, as exemplified by the self-splicing activity of certain group I and group II introns^{4,22}, the transfer RNA processing activity of the RNA component of RNase P (ref. 5), and the self-cleaving activity of 'hammerheads' and related RNA structures^{23,24}. These behaviours, although of doubtful prebiotic significance in themselves, indicate that life based on RNA would have included function as well as informational properties. Third, RNA plays an important part in modern biological systems, appearing most conspicuously in those cellular processes that are presumed to be among the most ancient. RNA primes DNA synthesis during DNA replication, carries the genetic message from DNA to the translation apparatus, gathers and positions amino acids in accordance with the codon structure of messenger RNA and functions as an integral part of the ribosome. RNA also has a central role in the spliceosome apparatus^{25,26} and is an essential component of the ribonucleoprotein enzyme that directs telomere synthesis in eukaryotic cells²⁷. Fourth, most of the biological coenzymes are nucleotides or compounds that could be derived from nucleotides (Table 1), a curious fact that may indicate that RNA and RNA enzymes were present before the development of protein synthesis²⁸. Fifth, histidine, which is important in enzyme catalysis as either

a nucleophile or general acid-base catalyst, is produced by a very unusual biosynthetic pathway that begins with phosphoribosyl pyrophosphate and adenosine triphosphate (ATP). The functional imidazole moiety, which now operates as part of a histidine residue within a protein enzyme may once have operated as part of a purine base within an RNA enzyme²⁸. Sixth, in contemporary metabolism, deoxyribonucleotides are synthesized by enzymatic reduction of ribonucleotides rather than by a biosynthetic pathway comparable to that used to synthesize ribonucleotides. The most widely distributed reaction system makes use of three nucleotide-derived components: two molecules of flavin adenine dinucleotide (FAD) associated with thioredoxin, and one molecule of nicotinamide adenine dinucleotide phosphate (NADP). Seventh, deoxythymidylic acid is formed by 5-methylation of deoxyuridylic acid, again suggesting the primacy of ribonucleosides over deoxyribonucleosides.

The most important feature of RNA is that it combines genotype and phenotype in a single molecule, so that replication of RNA enables darwinian evolution to occur at the molecular level. Genetic variation is introduced as a result of mutational errors and RNA-catalysed recombination events. This variation is in turn expressed as a range of phenotypes. Because of differences in their chemical behaviour, some RNAs will replicate more efficiently than others and will grow to dominate the population until a new variant with even greater selective advantage arises²⁹. This is the RNA world.

There are several objections to the notion that life began with RNA, that fall into two general categories. First, RNA is biochemically inept, particularly when compared to proteins whose catalytic prowess is so apparent in modern biochemistry. With the discovery of catalytic RNA, this argument has lost some force but it can still be said that RNA does not seem as versatile a catalyst as protein is. Second, RNA is not a plausible prebiotic

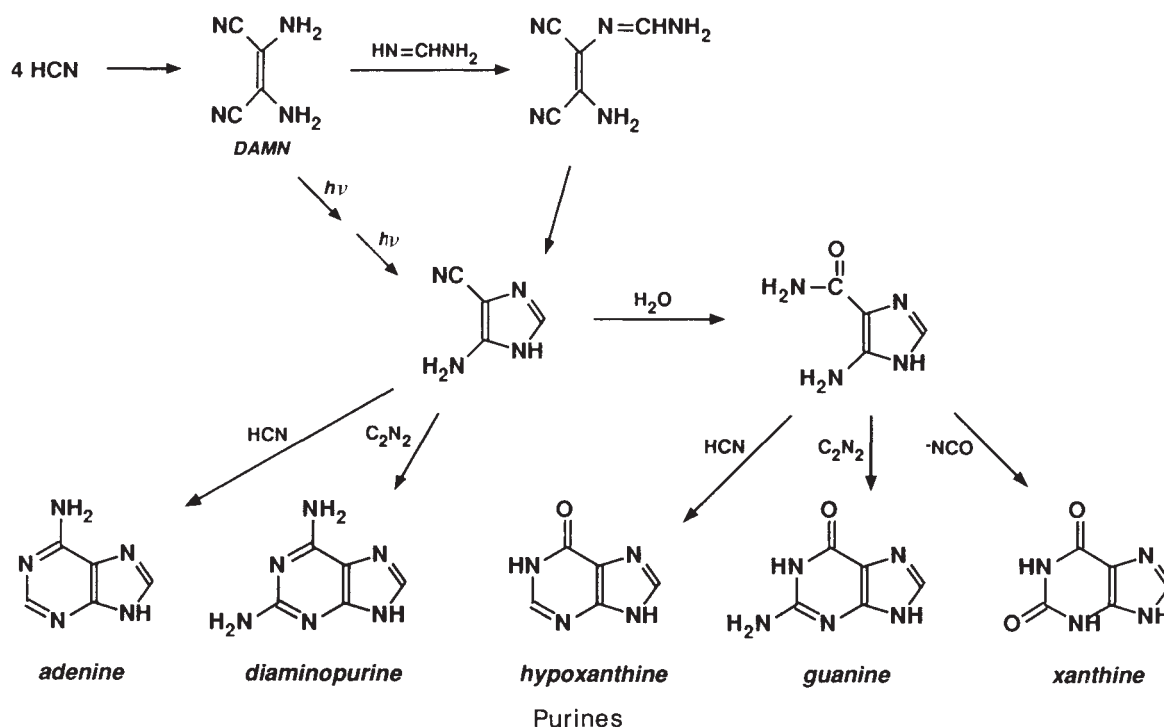


FIG. 1 Prebiotic synthesis of purines by self-condensation of HCN. HCN undergoes a tetramerization reaction to produce diaminomaleonitrile (DAMN). The rate-determining step is the attack on HCN by CN^- to form a dimer,

rapidly converted to a trimer and then to the tetramer, DAMN. This reacts with either formamidine or ultraviolet light to produce 4-aminoimidazole-5-carboxamide, which in turn leads to the formation of various purines.

TABLE 1 Coenzymes that resemble nucleotides

Coenzyme	Base	Sugar-phosphate	Substituent
NAD	Adenine	Ribose 5'-phosphate; ribose 5'-phosphate	Nicotinamide
NADP	Adenine	Ribose 2', 5'-bisphosphate; ribose 5'-phosphate	Nicotinamide
FAD	Adenine	Ribose 5'-phosphate; ribitol 5'-phosphate	Isoalloxazine
FMN	—	Ribitol 5'-phosphate	Isoalloxazine
TPP	Pyrimidine	Pyrophosphate	Hydroxyethylthiazole
THF	2-Amino-4-hydroxy-6-methylpteridine	—	<i>p</i> -aminobenzoic acid; glutamic acid
Coenzyme A	Adenine	Ribose 3'-phosphate, 5'-diphosphate	Pantothenic acid; β -mercaptoethylamine
Coenzyme B ₁₂	Adenine; 5,6-dimethyl-benzimidazole	5'-Deoxyribose; α -ribose 3'-phosphate	Cobalamin
S-adenosylmethionine	Adenine	Ribose	Methionine
ATP	Adenine	Ribose 5'-triphosphate	—
UDP-glucose	Uracil	Ribose 5'-diphosphate	Glucose
CDP-diacylglycerol	Cytosine	Ribose 5'-diphosphate	Diacylglycerol

The carrier molecules S-adenosylmethionine, ATP, UDP-glucose and CDP-diacylglycerol are included with the traditional coenzymes. A number of other nucleoside diphosphate sugars are known to act as carriers in contemporary metabolism. FMN, flavin mononucleotide; TPP, thiamine pyrophosphate; THF, tetrahydrofolate; UDP, uridine diphosphate; CDP, cytidine diphosphate.

molecule because it is unlikely to have been produced in significant quantities on the primitive Earth³⁰. This is a more serious criticism that goes to the heart of the question of whether RNA was the first genetic molecule or whether it took on the role of genetic material at a later stage in evolutionary history. Next, therefore, I consider how difficult it would have been to synthesize RNA in the prebiotic environment.

The primordial setting

Perhaps the most surprising aspect of the origins of life on Earth, other than the fact it happened at all, is that it happened so quickly. The Earth was formed from a diffuse cloud of cosmic gas and dust about 4.6×10^9 years ago³¹ and by most accounts, was a very inhospitable place for the first $0.4\text{--}0.6 \times 10^9$ years of its history. It is virtually certain that no useful organic chemistry could have occurred until about $4.0\text{--}4.2 \times 10^9$ years ago—after the Earth's crust had formed³² and the level of meteoric impacts had subsided so that the environment was no longer in constant upheaval³³. On the other hand, the fossil record strongly supports the existence of cellular life by 3.6×10^9 years ago³⁴ and it has been claimed that indirect evidence exists for biological carbon fixation as early as 3.8×10^9 years ago³⁵. The time window for biogenesis, from organic chemistry to a primitive cell, seems to be 'only' about 0.4×10^9 years.

The composition of the prebiotic atmosphere is a problem of fundamental importance to our understanding of the origins of life. Until recently, it was generally thought that the prebiotic atmosphere was strongly reducing, containing mostly H_2 , CH_4 , NH_3 and H_2O ³⁶. As scientific models of planetary formation have become more sophisticated, it has become increasingly clear that we do not really know whether the prebiotic atmosphere was strongly reducing, mildly reducing or even non-reducing, although it seems certain that there was no free oxygen ($<10^{-6}$ atmospheres) until the advent of oxygen-producing photosynthetic bacteria about $2.5\text{--}2.9 \times 10^9$ years ago³⁷. At present, the weight of evidence indicates that the prebiotic atmosphere was not strongly reducing.

The atmosphere was formed primarily by outgassing of volatile components of the primitive mantle. Its initial composition depends on whether outgassing occurred before or after the iron-rich core of the planet had formed: if outgassing occurred before core formation, contact with metallic iron would have ensured a strongly reducing atmosphere, containing mostly H_2 , H_2O , CH_4 and CO ; if the iron had already sunk to the core, the redox state would have been determined by the $\text{Fe}^{2+}/\text{Fe}^{3+}$ ratio in iron minerals, resulting in a less reducing atmosphere that contained H_2O , CO_2 , H_2 and CO but almost no CH_4 ³⁸. It has been argued that even if reduced species such as CH_4 and NH_3 were produced before core formation, they would have been rapidly depleted by photodecomposition³⁹. As time went on, escape of H_2 from the upper atmosphere and outgassing of less reduced species would have tipped the chemical balance in favour of CO and CO_2 over CH_4 and in favour of N_2 over NH_3 . There may have been some mechanism for the continued production of CH_4 and NH_3 , such as the slow re-equilibration of CO_2 and CH_4 at temperatures below 873K (ref. 40), by the arrival of CH_4 and NH_3 on comets and meteorites⁴¹ or by the photochemical reduction of CO_2 in the presence of Fe^{2+} (ref. 42).

Although a strongly reducing environment is most favourable for the synthesis of organic compounds, several prebiotic syntheses have been demonstrated under less strongly reducing conditions. Consider the synthesis of hydrogen cyanide and formaldehyde, which in turn provide a good starting point for the synthesis of more complex materials. Both HCN and H_2CO have been obtained in good yield in spark-discharge experiments under mildly reducing conditions (partial pressure of N_2 of 100 torr and H_2 of 0–400 torr)⁴³. Even in the absence of H_2 , the yield of HCN and H_2CO is about 2% (based on carbon) or about $2\text{--}3 \text{ nmol J}^{-1}$ (based on the energy input). This works out to about $4.5 \times 10^9 \text{ kg yr}^{-1}$ on the primitive Earth, assuming

the energy available from the solar corona discharge was $12.6 \text{ J cm}^{-2} \text{ yr}^{-1}$ (ref. 44). A photochemical source of comparable magnitude has been proposed for both HCN (ref. 45) and H_2CO (ref. 46).

Once produced, HCN would have been hydrolysed and H_2CO would have been destroyed by ultraviolet irradiation, unless they happened to enter a relatively protected microenvironment and managed to combine into more stable compounds. But such combination cannot occur unless the reactants build up to sufficiently high concentrations, which they cannot do unless they are already protected from decomposition. Two possible solutions to this dilemma have been suggested: there may have been a physical mechanism for concentrating the reactants, such as adsorption on a mineral surface⁴⁷ or eutectic freezing in a shallow pond or hail particle⁴⁸; alternatively, the condensation reactions may have been catalysed and thus may have operated at much lower concentrations of reactants. An example is glycolonitrile, which is produced readily from HCN and H_2CO and which enhances the rate of HCN oligomerization by one to two orders of magnitude⁴⁹.

With HCN and H_2CO thus provided, and with the goal of the RNA world in mind, what came next? Assuming that HCN and H_2CO were present in sufficient quantities to kindle HCN oligomerization, several biologically important molecules would

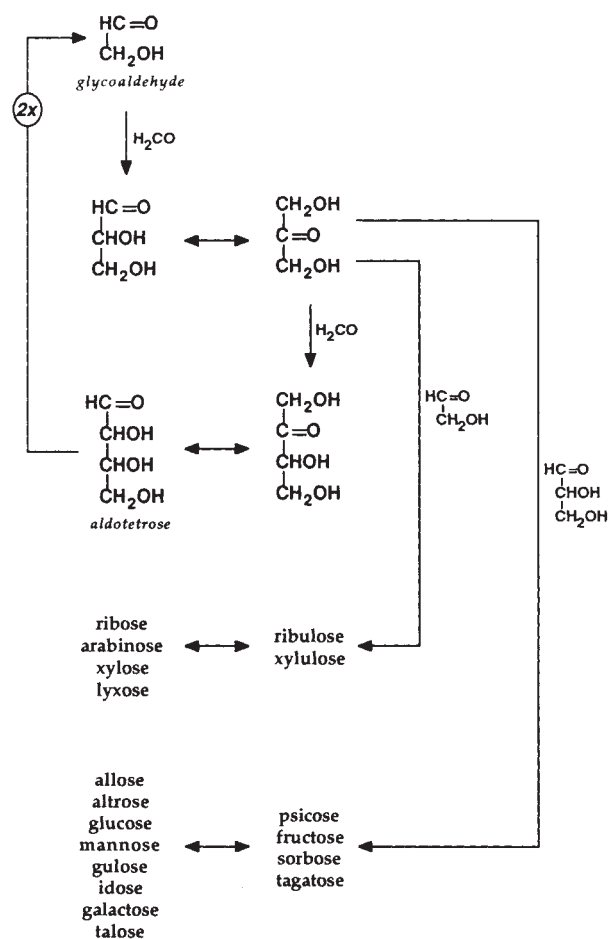


FIG. 2 The formose reaction for the prebiotic synthesis of sugars. Single-headed arrows represent aldol condensations (combination of two aldehydes or ketones to form a β -hydroxyaldehyde or β -hydroxyketone); double-headed arrows represent enolizations (transfer of a proton from an α -carbon to a carbonyl oxygen). The aldotetroses (threose and erythrose) undergo a reverse aldol condensation to yield two molecules of glycoaldehyde. The figure shows D-erythrose as an example of an aldotetrose, but both threose and erythrose can undergo this reaction. Branched-chain and seven-carbon sugars and D- and L-isomers are not shown.

be expected to form (Fig. 1). HCN undergoes a base-catalysed tetramerization reaction to produce diaminomaleonitrile⁵⁰, which participates in several interesting chemical reactions, most notably with formamidine to produce 4-aminoimidazole-5-carbonitrile. This leads to the production of adenine, hypoxanthine, guanine, xanthine and diaminopurine⁵⁰. The 4-aminoimidazole-5-carbonitrile intermediate can also be produced directly from diaminomaleonitrile by irradiating with sunlight, so avoiding the use of formamidine which itself must be derived from HCN and NH₃⁵¹.

Diaminomaleonitrile goes on to form higher oligomers of HCN, which appear as a heterogeneous mixture of compounds. Hydrolysis of HCN oligomers yields glycine, alanine, aspartate, diaminosuccinate and smaller amounts of other amino acids⁵². Amino acids have also been synthesized in spark-discharge experiments performed under mildly reducing conditions⁵³, although the yield and variety of amino acids is substantially lower compared to syntheses performed under strongly reducing conditions. The earliest organisms may have been required to provide their own reducing power, for example using the reduced form of nicotinamide adenine dinucleotide (NAD), NADH, for the biosynthesis of amino acids.

Shifting the focus from HCN to H₂CO, another remarkable polymerization reaction probably took place on the primitive Earth. Accompanying H₂CO would have been smaller amounts of glycoaldehyde, produced either directly or by the condensation of H₂CO in the presence of a catalyst such as calcium carbonate or alumina^{54,55}. Glycoaldehyde begins a cascade of aldol condensations and enolizations that rapidly convert most of the available formaldehyde into trioses, tetroses and higher sugars (Fig. 2). This cascade of reactions, collectively termed the formose (or Butlerow) reaction, proceeds autocatalytically because one molecule of threose or erythrose can be cleaved to give two molecules of glycoaldehyde⁵⁶. From a prebiotic perspective, the formose reaction seems a good way to produce

ribose and other sugars, although unless they were rapidly used to form nucleosides or some other useful compound, they would have been subject to degradation⁵⁷.

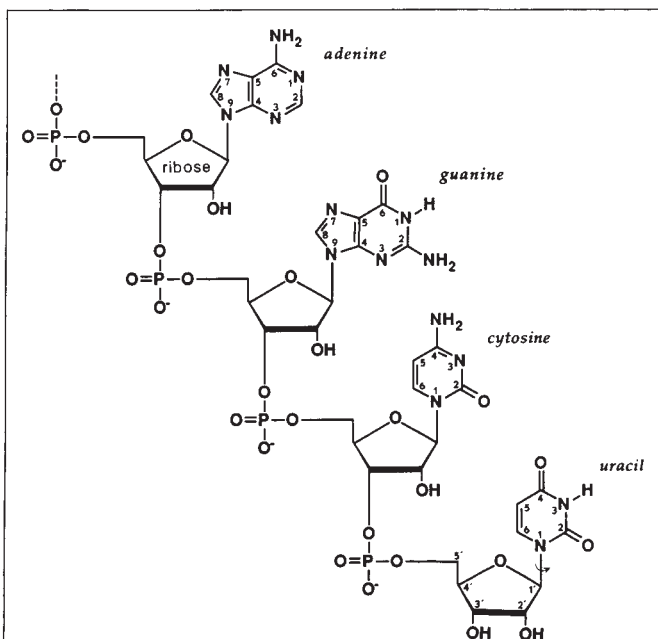
The trouble with RNA

It is a long and difficult road from HCN and H₂CO to the activated nucleoside phosphates that make up RNA (see Box). The self-condensation of HCN to produce purines is such a remarkably simple and efficient reaction that it would be surprising if it did not have some relevance to the early history of life. The prebiotic synthesis of pyrimidines, on the other hand, is more problematic and it is difficult to see how a supply of cytosine or cytosine-containing nucleotides could have been maintained. Cytosine is obtained in about 5% yield when an aqueous solution containing 1.0 M potassium cyanate and 0.1 M cyanoacetylene is incubated at 100 °C for 24 hours⁵⁸. This can hardly be called a prebiotic synthesis because it requires concentrations of reactants that are unlikely to have occurred on the primitive Earth. Cyanoacetylene has been produced by sparking a mixture of CH₄ and N₂ (ref. 58) and presumably could also be obtained from a less reducing mixture. Cyanoacetylene is, however, even more susceptible to hydrolysis than HCN and is unlikely to have built up in significant concentration. There is an alternative route using cyanoacetaldehyde (the first hydrolysis product of cyanoacetylene) and guanidine, although the yields of uracil and, especially, cytosine are low⁵⁹. Furthermore, cytosine would have undergone rapid hydrolysis to uracil (half-life 200 yr at 30 °C)⁶⁰. In contemporary biochemistry, cytidine triphosphate (CTP) is produced by the amination of uridine triphosphate (UTP) in an ATP-dependent reaction.

Another route to the pyrimidines is by acid hydrolysis of HCN oligomers, which can produce 4,5-dihydroxypyrimidine, 5-hydroxyuracil and orotic acid⁶¹. Orotic acid undergoes a photochemical decarboxylation to give uracil⁶², analogous to the enzyme-catalysed decarboxylation that occurs in pyrimidine nucleotide biosynthesis. Uracil can also be produced by condensation of β-alanine and cyanate to give dihydrouracil, followed by photochemical dehydrogenation⁶³. Neither of these pathways for the prebiotic synthesis of pyrimidines is especially convincing.

The greatest problem with the pyrimidines is not their synthesis but the difficulty in attaching them to ribose to form pyrimidine nucleosides. First, consider the synthesis of ribose. The formose reaction may offer a good route for the prebiotic synthesis of sugars but it is unclear how ribose could be singled out from the complex mixture of products (Fig. 2). The five-carbon sugars are produced by condensation of glycoaldehyde and glyceraldehyde, accompanied by comparable amounts of tetroses and hexoses produced by the self-condensation of glycoaldehyde and glyceraldehyde, respectively. Even among the pentoses, there are twelve stereoisomers, eight aldoses and four ketoses, to be reckoned with. Perhaps the chemical properties of ribose led to its selective enrichment by some prebiotic fractionation process. The *cis*-(2,3)-diol structure of ribose and lyxose may have favoured their interaction with a mineral surface where they were relatively protected from hydrolysis compared to arabinose and xylose. The *trans* relationship of the 3-OH and 5-CH₂OPO₃²⁻ groups of ribose-5-phosphate and arabinose-5-phosphate may have rendered them less susceptible to intramolecular cyclization than xylose-5-phosphate and lyxose-5-phosphate⁶⁴. Even if such mechanisms were operating, they would have resulted in a racemic mixture of D- and L-ribose, which still leads to serious difficulties (see below).

Next comes the problem of joining D, L-ribose and a base. When ribose and a purine are heated to dryness in the presence of inorganic salts (Na⁺, Mg²⁺, Ca²⁺, Cl⁻ and SO₄²⁻), a mixture of α- and β-nucleosides is obtained in about 2–10% yield⁶⁵. A comparable reaction between ribose and pyrimidines is very inefficient (<0.1% yield) and no alternative method for the



The primary structure of RNA. Purine and pyrimidine bases are joined to sugar residues (riboses) to form nucleosides. The purines, adenine and guanine, are attached to ribose by a glycosidic bond between C1' of ribose and N9 of the base. The pyrimidines, cytosine and uracil, are attached by a bond between C1' of ribose and N1 of the base. The ribose subunits are joined by a 3', 5' phosphodiester linkage to form the ribose-phosphate backbone. Rotation about the glycosidic bond converts the *anti* conformation to the *syn* conformation. Inversion about C1' would convert the β-conformation to the α-conformation. The numbering schemes for the bases and ribose subunits are shown.

synthesis of pyrimidine nucleosides has been found. In biological systems, the pyrimidines are attached to ribose by displacement of a pyrophosphate at the C1 position but attempts to perform a similar displacement under plausible prebiotic conditions have been unsuccessful⁶⁶. Even the successful synthesis of purine nucleosides is plagued by side reactions, including the production of ribosylamine derivatives based on the C6-amine of adenine or the C2-amine of guanine. In biological systems, the purine nucleosides are synthesized by constructing the purine base on a pre-existing ribose-5-phosphate. A similar synthesis under prebiotic conditions seems far more difficult than using the purines formed by HCN polymerization and joining them directly to ribose.

At this point, the best possibility is a mixture of α , β -D, L-purine nucleosides accompanied by a dizzying array of related compounds. Now comes the task of phosphorylating the nucleosides. Heating a mixture of urea, ammonium chloride and hydroxylapatite $[\text{Ca}_{10}(\text{PO}_4)_6(\text{OH})_2]$ to 100 °C for 24 hours results in a good yield of inorganic polyphosphate⁶⁷. If a nucleoside is added to the mixture, about 20–25% of it is converted to a mixture of nucleoside phosphates, including 2'-, 3'-, 2',3'-cyclic- and 5'-phosphates, and various 2'(3'),5'-diphosphates⁶⁸. Alternative routes using cyanogen or cyanoacetylene rather than high temperature to drive the reaction have been demonstrated⁶⁹, although, as already discussed, these condensing agents were probably not widely available on the primitive Earth.

Finally, the nucleoside phosphates must have been activated to provide a substrate for template-directed polymerization. Adenosine 2', 3'-cyclic phosphate is in a sense already activated; it can be oligomerized by heating it to dryness⁷⁰ or by incubating it at room temperature in the presence of a polyamine⁷¹. Given the proper stereochemical orientation, the 2', 3'-cyclic phosphate undergoes condensation with an adjacent nucleoside 5'-

hydroxyl to produce a 2', 5' or 3', 5' phosphodiester linkage⁷². This is comparable to the RNA-catalysed ligation reaction that has been demonstrated using the satellite RNA of tobacco ringspot virus⁷³ or the genomic RNA of hepatitis delta virus⁷⁴.

A more practical approach might be to activate a nucleoside 5'-phosphate and take advantage of the greater nucleophilicity of the *cis*-(2', 3')-diol compared to the 5'-hydroxyl for the condensation between a 5'-phosphate and 2'(3')-hydroxyl. Adenosine 5'-phosphate can be converted to the diphosphate and subsequently the triphosphate by incubating it with carbamyl phosphate in the presence of a divalent metal cation⁷⁵. Carbamyl phosphate is itself produced by condensation of cyanate and inorganic phosphate in aqueous solution⁷⁶. Nucleoside 5'-polyphosphates have also been obtained by heating a solution of nucleotides and inorganic polyphosphate to dryness⁷⁷.

Could this long and difficult road to activated mononucleotides be travelled without biological catalysts (or an organic chemist) to lead the way? Suppose that somehow it did happen; that given a few hundred million years and an entire globe of microenvironments, there was a special time and place in which a rich soup of activated mononucleotides formed. The only flaw in this hypothetical garden of Eden is that the soup would contain a racemic mixture of D- and L-nucleotides, a necessary consequence of the fact that the prebiotic environment was achiral, which, it turns out, makes RNA replication very difficult. By a strange twist of nucleotide chemistry, D-mononucleotides in the *anti* conformation and L-mononucleotides in the *syn* conformation tend to form very close structural homologues when bound to a complementary template⁷⁸ (Fig. 3). As a result, the template-directed synthesis of one enantiomer is strongly inhibited by the presence of the other⁷⁹. The 'wrong' enantiomer adds to the 3' end of a growing chain, which distorts the 3' terminus and thus prevents further chain elongation.

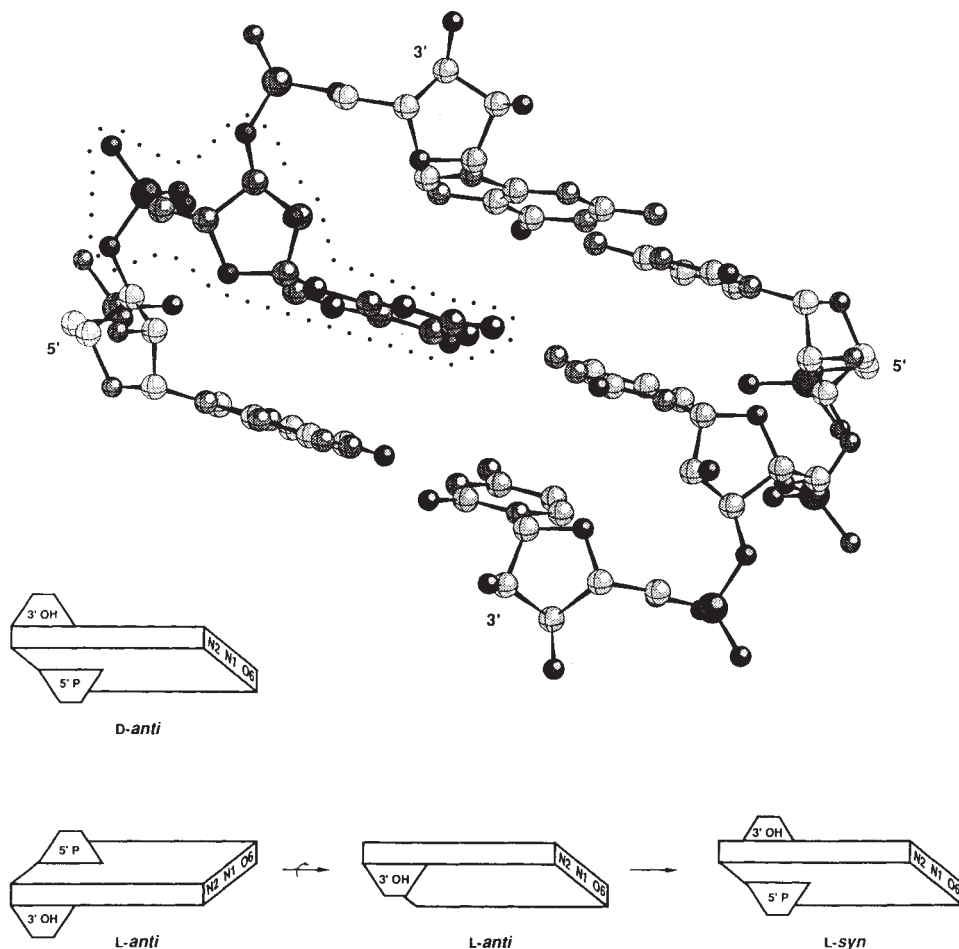


FIG. 3 Top, three successive guanosine · cytosine pairs of the A-form of RNA projected onto a plane perpendicular to the diad axis. The middle guanosine residue is enclosed in a dotted box. Atom coordinates are based on the crystal structure of A-RNA⁹⁶. Graphical representation was performed using *Ball & Stick* software on a Macintosh computer. Bottom, the close structural homology between D-*anti*- and L-*syn*-guanosine mononucleotides when bound to a complementary template. The D-*anti* form corresponds to the middle guanosine residue shown above. Imagine reflecting this in a mirror to produce the L-*anti* form, which is rotated about its long axis to orient the base for complementary pairing and finally rotated about the glycosidic bond to produce the L-*syn* form.

If one admits both α - and β -nucleotides to the hypothetical garden of Eden, another cross-inhibition phenomenon is likely to occur. α -D-mononucleotides in the *syn* conformation and β -D-mononucleotides in the *anti* conformation also form close structural homologues⁸⁰ and any combination of α -D-*syn*-, α -L-*anti*-, β -D-*anti*-, or β -L-*syn*-mononucleotides is likely to be a problem. The addition of arabinose, other pentoses and the hexoses is an open invitation to template-directed chaos. It is difficult to imagine how a genetic message could be transferred from template to complementary product in the face of all this isomeric interference. The most reasonable interpretation is that life did not start with RNA. The RNA world came into existence after many of the problems associated with the prebiotic synthesis and template-directed replication of RNA had been solved. This implies that there was a simpler genetic system, or systems, that preceded RNA and that evolutionary advances made by the ancestral system were somehow carried over to the RNA world.

The origin of self-replication

We know very little about the chemical nature of the first self-replicating molecule. There are four general requirements for the primordial genetic material, regardless of its chemical nature. First, it must have had informational properties, suggesting that it was some type of heteropolymer, composed of at least two monomeric subunits. Second, it must have been capable of directing the ordered assembly of monomeric starting materials to form additional copies of itself. Third, the monomeric starting materials must have been readily available, at least in some locale on the primitive Earth. Fourth, the genetic material must have been sufficiently stable for its rate of reproduction to exceed its rate of decomposition. If the goal is the eventual origin of an RNA world, then a requirement for an evolutionary pathway from the first self-replicating molecule to a genetic system based on RNA must be added.

There are three ways in which the difficulties in arriving at an RNA world may have been reduced. One is that before the origin of self-replication, there was a period of chemical evolution⁸¹, during which non-instructed processes led to progressive alteration of the chemical composition of the environment. The prebiotic environment can be thought of as a flow reactor, driven by the supply of HCN, H₂CO and other high-energy starting materials, and characterized by a steady-state distribution of reactants and products. In this context, the formose reaction for instance, may not have been quite so unruly; rather than the cascade of all possible 3-, 4-, 5-, 6- and 7-carbon sugars obtained from batch polymerization of H₂CO, there might have been more limited distribution of sugars determined by the differential rates of formation and decomposition of the various compounds⁸². Non-instructed chemical ordering may have led to very complex structures, such as colloidal aggregates of polypeptides¹⁸ and even self-propagating membranous vesicles⁸³. These complex structures may have had a profound influence on the local environment and may have set the stage for the emergence of a self-replicating system.

A second possibility is the intervention of a primitive self-replicating system chemically unrelated to RNA, which may have helped to bias the environment in a way that made the appearance of RNA or RNA-like molecules more likely. The most carefully constructed hypothesis along these lines concerns the possibility of self-replicating clay minerals¹³, where the genetic material is considered to be a distribution of ionic charges, embedded organic compounds or structural defects along a mineral surface that serves as a template for the formation of additional mineral layers. At one extreme, these clay 'organisms' may have come and gone, leaving their mark by altering the local environment. At the other extreme, they may have developed the capacity to synthesize RNA or RNA-like molecules that were used initially as catalysts or cofactors in the clay world but eventually themselves became self-replicating and took on the role of genetic material⁸⁴.

The third advance on the way to an RNA world may have been the appearance of a self-replicating chemical system similar to RNA but simpler in the sense that it demanded less from the chemistry of the prebiotic environment^{78,85}. For example, the difficulties associated with the prebiotic synthesis of pyrimidines could be circumvented if the genetic material were based on purines alone. Non-standard pairing combinations such as adenine:inosine, adenine:xanthine and guanine:isoguanine have been suggested^{8,86}, although there is no experimental evidence that they are suitable for template-directed polymerization. The backbone might have been simpler, with ribose replaced by something that was not accompanied by dozens of closely related compounds or subject to enantiomeric cross-inhibition. A number of flexible, acyclic nucleoside analogues that fit this description have been proposed^{78,87} (Fig. 4).

The compound that has been studied in the greatest detail is the glycerol-derived nucleoside analogue (compound *b* in Fig. 4). In principle, it could be synthesized by condensation of H₂CO and glycerol followed by addition of a base. This should produce a mixture of various mono- and di-substituted glycerol derivatives but this would be far simpler than the assortment of ribosyl- and other sugar derivatives that would accompany a nucleoside. This compound, either non-phosphorylated or phosphorylated at both hydroxyl positions, is an achiral molecule. It becomes chiral when incorporated into a heteropolymeric structure. The glycerol-derived analogue of guanosine bisphosphate, when activated at both phosphates, undergoes a poly(C)-directed oligomerization reaction to form pyrophosphate-linked products⁸⁸. A pyrophosphate-linked oligomer formed from deoxycytidine bisphosphate has also been used as a template to direct the oligomerization of deoxyguanosine bisphosphate derivatives (A. W. Schwartz, personal communication). Thus there could have been a genetic molecule with a glycerol-phosphate backbone that replicated by a chemical process and was able to undergo darwinian evolution.

Chemical self-replication, in the glycerol-phosphate system or in any RNA-like system, faces several obstacles. One of the most serious is the need for end-to-end copying of the template before the development of a mechanism to ensure that replica-

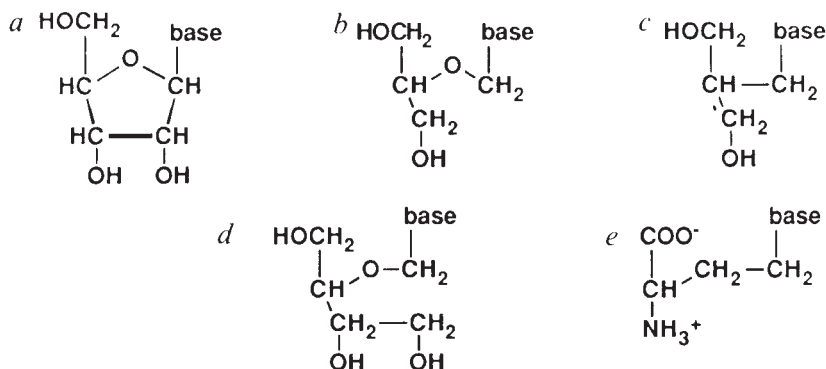


FIG. 4 Comparative structure of a nucleoside and of four simple acyclic nucleoside analogues. *a*, nucleoside; *b*, glycerol-derived acyclonucleoside; *c*, acrolein-derived nucleoside analogue; *d*, erythritol-derived acyclonucleoside; *e*, aspartate-derived nucleoside analogue.

tion begins at a discrete initiation site. In the RNA world this problem may have been solved by a 'genomic-tag' structure that directed a primitive replicase to the appropriate initiation site⁸⁹. In the early stages of self-replication, a chemical tag could have demarcated one end of the template or a chemical ligation mechanism could have joined partial copies together to form complete products.

Another important obstacle to chemical self-replication, which has been demonstrated in experimental systems, is the tendency of molecules with regions of self-complementarity to form intramolecular duplex structures⁹⁰. It is difficult to envisage a solution to this problem in the absence of a factor such as single-strand binding proteins, because the conditions must favour short duplex regions so that template-directed synthesis can occur yet disfavour long duplex regions that would produce stable secondary structures. There may be alternative helical backbones that have non-cooperative properties so that longer duplex structures are increasingly disfavoured⁷⁸. In any event, the development of an unwinding activity would have been a top priority during early evolutionary history.

Finally, there is the problem of accurate copying of the genetic information in the absence of a sophisticated replicase enzyme. The selectivity of complementary base pairing provides the chemical basis for information transfer, with the fidelity depending on the particular bases involved (highest for cytosine → guanine, lowest for adenine → uracil)⁹¹. The efficiency of copying also depends on base sequence, with selection pressure favouring those sequences that can be copied most readily. One of the tasks of a polymerase enzyme is to 'smooth' these differences so that a broader range of sequences can be explored⁹². But given the limits on genome size (as determined by copying fidelity²⁹) and on sequence heterogeneity, would there have been enough potential sequences remaining to evolve a smoothing function that could further expand the range of potential sequences? A crude replicase activity must have arisen so that copying fidelity could be improved, genome size could be increased and ever more sophisticated catalysts could be developed.

Transition to an RNA world

Let us suppose that after a period of chemical evolution, with or without further augmentation by a genetic system completely unrelated to RNA, a genetic system arose based on some simple RNA-like molecule. It is not inconceivable that a genetic system completely unrelated to RNA would have invented RNA for its own purposes. With so many design problems associated with RNA synthesis, it seems more likely that a simple RNA-like molecule was invented instead. Once a genetic system based on an RNA-like molecule existed, the domain of nucleotide chemistry would have been more accessible and the invention of RNA would have been more practical. Why should an increasingly sophisticated RNA-like world come to synthesize mononucleotides rather than optimize the synthesis of its own monomeric units? The answer must be that mononucleotides somehow conferred a selective advantage to the RNA-like world, either by improving its information storage properties or by enhancing its catalytic abilities.

Reviewing the catalytic properties of known RNA enzymes⁹³,

RNA catalysis seems to be limited to sequence-specific cleavage/ligation reactions involving nucleic acid substrates. But the chemical properties of the various functional groups within RNA suggest a broader range of catalytic function (Table 2). Of particular importance to RNA, but not to any of its likely predecessors, is the ribose 2' hydroxyl. A terminal *cis*-(2', 3')-diol is a good nucleophile and an internal 2'-hydroxyl can function either as a nucleophile²² or as a determinant of RNA tertiary structure^{94,95}. Because of the 2'-hydroxyl, RNA may have enhanced the structural and catalytic properties of the RNA-like world and its production could have been a favourable event.

We do not know how the transition from an RNA-like to an RNA world was made. There are several ways in which RNA may have gained entry. One is that nucleotides or nucleotide derivatives were generated to perform some useful function in the RNA-like world and then took root as cofactors in primitive metabolic processes. The RNA cofactors progressed from monomeric to oligomeric forms and eventually became autonomously self-replicating. Another possibility is that mononucleotides were at first inadvertently incorporated into the ancestral genetic material but this proved selectively advantageous. It is not clear how random misincorporations could have been useful or how the incorporation of mononucleotides could have been limited to specific sites. Somehow the genetic material took on more and more of the character of RNA without forsaking its past evolutionary accomplishments. A third possibility is that mononucleotides, originally present as cofactors, became oligomerized in a template-dependent way to give products that contained only RNA components. The ancestral genetic material directed both its own replication and the 'transcription' of RNA. RNA became autonomously self-replicating and either shed its RNA-like predecessor or, through reverse transcription, replaced the predecessor with DNA.

The transition to an RNA world, like the origins of life in general, is fraught with uncertainty and is plagued by a lack of relevant experimental data. Researchers into the origins of life have grown accustomed to the level of frustration in these problems. Perhaps it is time for a new group of scientists to join in the frustration. The discovery of catalytic RNA and the appreciation of the dual role of RNA as both genetic material and catalyst have kindled new interest in the evolution of RNA and the origins of life. It is time to go beyond talking about an RNA world and begin to put the evolution of RNA in the context of the chemistry that came before it and the biology that followed. □

G. F. Joyce is in the Department of Molecular Biology, Research Institute of Scripps Clinic, 10666 North Torrey Pines Road, La Jolla, California 92037, USA.

TABLE 2 Catalytic potential of RNA

Component	Metal-ion coordination	General acid catalysis	General base catalysis	Nucleophile
Ribose	2'OH			2'OH, 2'(3')-diol
Phosphate	O ⁻	OH	O ⁻	
Adenine	N7		N1	
Guanine	N7	N1	N7	
Cytosine	N3		N3	
Uracil	O2	N3		

- Walter, M. R. in *Earth's Earliest Biosphere* (ed Schopf, J. W.) 187–213 (Princeton University Press, 1983).
- Miller, S. L. *Science* **117**, 528–529 (1953).
- Oró, J. *Nature* **191**, 1193–1194 (1961).
- Kruger, K. et al. *Cell* **31**, 147–157 (1982).
- Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. *Cell* **35**, 849–857 (1983).
- Quastler, H. *The Emergence of Biological Organization* Ch. 1 (Yale University Press, New Haven, 1964).
- Woese, C. *The Genetic Code* Ch. 7 (Harper & Row, New York, 1967).
- Crick, F. H. C. *J. molec. Biol.* **38**, 367–379 (1968).
- Orgel, L. E. *J. molec. Biol.* **38**, 381–393 (1968).
- Sharp, P. A. *Cell* **42**, 397–400 (1985).
- Lewin, R. *Science* **231**, 545–546 (1986).
- Gilbert, W. *Nature* **319**, 618 (1986).
- Cairns-Smith, A. G. *J. theor. Biol.* **10**, 53–88 (1966).
- Nelstuen, G. L. *J. molec. Evol.* **15**, 59–72 (1980).
- Weber, A. L. *Origins of Life* **17**, 107–119 (1987).
- Orgel, L. E. *Origins of Life* **17**, 27–34 (1986).
- Oparin, A. I. *Adv. Enzymol.* **27**, 347–380 (1965).
- Fox, S. W. & Dose, K. *Molecular Evolution and the Origin of Life* (Dekker, New York, 1977).
- Dyson, F. J. *J. molec. Evol.* **18**, 344–350 (1982).
- Kauffman, S. A. *J. theor. Biol.* **119**, 1–24 (1986).
- Inoue, T. & Orgel, L. E. *Science* **219**, 859–862 (1983).
- Peebles, C. L., Mecklenberg, K. L., Perlman, P. S., Tabor, J. & Cheng, H. L. *Cell* **44**, 213–223 (1986).
- Uhlenbeck, O. C. *Nature* **328**, 596–600 (1987).
- Sharmeen, L., Kuo, M. Y. P., Dinter-Gottlieb, G. & Taylor, J. J. *Virology* **62**, 2674–2679 (1988).
- Sharp, P. A. *Science* **235**, 766–771 (1987).
- Maniatis, T. & Reed, R. *Nature* **325**, 673–678 (1987).
- Greider, C. W. & Blackburn, E. H. *Cell* **51**, 887–898 (1987).

28. White, H. B. *J. molec. Evol.* **7**, 101-104 (1976).
 29. Eigen, M. *Naturwiss.* **58**, 465-523 (1971).
 30. Shapiro, R. *Origins of Life* **14**, 565-570 (1984).
 31. Papanastassiou, D. A. & Wasserburg, G. J. *Earth planet. Sci. Lett.* **11**, 37-62 (1971).
 32. Ernst, W. G. in *Earth's Earliest Biosphere* (ed Schopf, J. W.) 41-52 (Princeton University Press, 1983).
 33. Maher, K. A. & Stevenson, D. J. *Nature* **331**, 612-614 (1988).
 34. Buick, R., Dunlop, J. S. R. & Groves, D. I. *Alcheringa* **5**, 161-181 (1981).
 35. Schidlowski, M., Appel, P. W. U., Eichmann, R. & Junge, C. E. *Geochim. Cosmochim. Acta* **43**, 189-199 (1979).
 36. Dickerson, R. E. *Sci. Am.* **239**, 70-86 (1978).
 37. Hayes, J. M. in *Earth's Earliest Biosphere* (ed Schopf, J. W.) 291-301 (Princeton University Press, 1983).
 38. Chang, S., DesMarais, D., Mack, R., Miller, S. L. & Strathearn, G. E. in *Earth's Earliest Biosphere* (ed Schopf, J. W.) 53-92 (Princeton University Press, 1983).
 39. Levine, J. S. & Augustsson, T. R. *Origins of Life* **15**, 299-318 (1985).
 40. Deines, P. *Geochim. cosmochim. Acta* **44**, 943-961 (1980).
 41. Oró, J. *Nature* **190**, 389-390 (1961).
 42. Borowska, Z. & Mauzerall, D. *Proc. natn. Acad. Sci. U.S.A.* **85**, 6577-6580 (1988).
 43. Stribling, R. & Miller, S. L. *Origins of Life* **17**, 261-273 (1987).
 44. Miller, S. L. & Urey, H. C. *Science* **130**, 245-251 (1959).
 45. Zahnle, K. J. *J. geophys. Res.* **91**, 2819-2834.
 46. Pinto, J. P., Gladstone, G. R. & Young, Y. L. *Science* **210**, 183-185 (1980).
 47. Bernal, J. D. *The Physical Basis of Life* **34** (Rouledge & Kegan Paul, London, 1951).
 48. Sanchez, R., Ferris, J. & Orgel, L. E. *Science* **153**, 72-73 (1966).
 49. Schwartz, A. W. & Goverde, M. *J. molec. Evol.* **18**, 351-353 (1982).
 50. Sanchez, R. A., Ferris, J. P. & Orgel, L. E. *J. molec. Biol.* **30**, 223-253 (1967).
 51. Ferris, J. P. & Orgel, L. E. *J. Am. Chem. Soc.* **88**, 1074 (1966).
 52. Ferris, J. P. & Hagan, W. J. *Tetrahedron* **40**, 1093-1120 (1984).
 53. Schlesinger, G. & Miller, S. L. *J. molec. Evol.* **19**, 376-382 (1983).
 54. Gabel, N. W. & Ponnampuram, C. *Nature* **216**, 453-455 (1967).
 55. Reid, C. & Orgel, L. E. *Nature* **216**, 455 (1967).
 56. Mizuno, T. & Weiss, A. H. *Adv. Carbohydr. Chem. Biochem.* **29**, 173-227 (1974).
 57. Mopper, K., Dawson, R., Liebezeit, G. & Ittekkot, V. *Mar. Chem.* **10**, 55-66 (1980).
 58. Sanchez, R. A., Ferris, J. P. & Orgel, L. E. *Science* **154**, 784-785 (1966).
 59. Ferris, J. P., Zamek, O. S., Altbuch, A. M. & Freiman, H. *J. molec. Evol.* **3**, 301-309 (1974).
 60. Shapiro, R. & Klein, R. S. *Biochemistry* **5**, 2358-2362 (1966).
 61. Ferris, J. P., Joshi, P. C., Edelson, E. H. & Lawless, J. G. *J. molec. Evol.* **11**, 293-311 (1978).
 62. Ferris, J. P. & Joshi, P. C. *J. Org. Chem.* **44**, 2133-2137 (1979).
 63. Chittenden, G. J. F. & Schwartz, A. W. *Nature* **263**, 350-351 (1976).
 64. Hill, A. R., Nord, L. D., Orgel, L. E. & Robins, R. K. *J. molec. Evol.* **28**, 170-171 (1988).
 65. Fuller, W. D., Sanchez, R. A. & Orgel, L. E. *J. molec. Biol.* **67**, 25-33 (1972).
 66. Orgel, L. E. & Lohrmann, R. *Accounts Chem. Res.* **7**, 368-377 (1974).
 67. Osterberg, R., Orgel, L. E. & Lohrmann, R. *J. molec. Evol.* **2**, 231 (1973).
 68. Lohrmann, R. & Orgel, L. E. *Science* **171**, 490-494 (1971).
 69. Lohrmann, R. & Orgel, L. E. *Science* **161**, 64-66 (1968).
 70. Bishop, M. J., Lohrmann, R. & Orgel, L. E. *Nature* **237**, 162-164 (1972).
 71. Verlander, M. S., Lohrmann, R. & Orgel, L. E. *J. molec. Evol.* **2**, 303-316 (1973).
 72. Usher, D. A. *Nature New Biology* **235**, 207(1972).
 73. Buzayan, J. M., Gerlach, W. L. & Bruening, G. *Nature* **323**, 349-353 (1986).
 74. Wu, H.-N. & Lai, M. M. C. *Science* **243**, 652-654 (1989).
 75. Saygin, O. *Naturwiss.* **67**, 617-619 (1981).
 76. Jones, M. E. & Lipmann, F. *Proc. natn. Acad. Sci. U.S.A.* **46**, 1194-1205 (1960).
 77. Lohrmann, R. *J. molec. Evol.* **6**, 237-252 (1975).
 78. Joyce, G. F., Schwartz, A. W., Miller, S. L. & Orgel, L. E. *Proc. natn. Acad. Sci. U.S.A.* **84**, 4398-4402 (1987).
 79. Joyce, G. F. et al. *Nature* **310**, 602-604 (1984).
 80. Praseuth, D. et al. *J. molec. Biol.* **196**, 939-942 (1987).
 81. Calvin, M. *Chemical Evolution* (Oxford University Press, New York, 1969).
 82. Harsch, G., Harsch, M., Bauer, H. & Voelter, W. *Z. Naturforsch.* **38b**, 1269-1280 (1983).
 83. Morowitz, H. J., Heinz, B. & Deamer, D. *Origins of Life* **18**, 281-287 (1988).
 84. Cairns-Smith, A. G. in *Origin of Life: Proceedings of the Second ISSOL Meeting, Kyoto, Japan* (ed Noda, H.) 399-404 (Center for Academic Publications, Tokyo, 1978).
 85. Spach, G. *Origins of Life* **14**, 433-437 (1984).
 86. Wächtershäuser, G. *Proc. natn. Acad. Sci. U.S.A.* **85**, 1134-1135 (1988).
 87. Orgel, L. E. in *Evolutionary Tinkering in Gene Expression* (ed Grunberg-Manago, M.) (Plenum, London, in the press).
 88. Schwartz, A. W. & Orgel, L. E. *Science* **228**, 585-587 (1985).
 89. Weiner, A. M. & Maizels, N. *Proc. natn. Acad. Sci. U.S.A.* **84**, 7383-7387 (1987).
 90. Joyce, G. F. & Orgel, L. E. *J. molec. Biol.* **188**, 433-441 (1986).
 91. Joyce, G. F. *Cold Spring Harbor Symp. quant. Biol.* **52**, 41-51 (1987).
 92. Joyce, G. F., Inoue, T. & Orgel, L. E. *J. molec. Biol.* **176**, 279-306 (1984).
 93. Cech, T. R. *Science* **236**, 1532-1539 (1987).
 94. Pace, N. R. & Marsh, T. L. *Origins of Life* **16**, 97-116 (1985).
 95. Dock-Bregeon, A. C. et al. *Nature* **335**, 375-378 (1988).
 96. Arnott, S. & Hukins, D. W. L. *Biochem. Biophys. Res. Commun.* **48**, 1392-1399 (1972).

ACKNOWLEDGEMENTS. I thank Leslie Orgel, Stanley Miller and Francis Crick for their helpful comments.

Copies of articles from this publication are now available from the UMI Article Clearinghouse.

For more information about the Clearinghouse, please fill out and mail back the coupon below.

UMI Article Clearinghouse

Yes! I would like to know more about UMI Article Clearinghouse. I am interested in electronic ordering through the following system(s):

- DIALOG/Dialorder ITT Dialcom
 OnTyme OCLC ILL Subsystem
 Other (please specify) _____
 I am interested in sending my order by mail.
 Please send me your current catalog and user instructions for the system(s) I checked above.

Name _____
 Title _____
 Institution/Company _____
 Department _____
 Address _____
 City _____ State _____ Zip _____
 Phone (_____) _____

Mail to: University Microfilms International
 300 North Zeeb Road, Box 91 Ann Arbor, MI 48106