life

Hypothesis

The Hypothesis that the Genetic Code Originated in Coupled Synthesis of Proteins and the Evolutionary Predecessors of Nucleic Acids in Primitive Cells

Brian R. Francis

Department of Molecular Biology, University of Wyoming, Laramie, WY 82071, USA;
E-Mail: brianrf@uwyo.edu; Tel.: +1-307-460-0408

Abstract: Although analysis of the genetic code has allowed explanations for its evolution to be proposed, little evidence exists in biochemistry and molecular biology to offer an explanation for the origin of the genetic code. In particular, two features of biology make the origin of the genetic code difficult to understand. First, nucleic acids are highly complicated polymers requiring numerous enzymes for biosynthesis. Secondly, proteins have a simple backbone with a set of 20 different amino acid side chains synthesized by a highly complicated ribosomal process in which mRNA sequences are read in triplets. Apparently, both nucleic acid and protein syntheses have extensive evolutionary histories. Supporting these processes is a complex metabolism and at the hub of metabolism are the carboxylic acid cycles. This paper advances the hypothesis that the earliest predecessor of the nucleic acids was a β-linked polyester made from malic acid, a highly conserved metabolite in the carboxylic acid cycles. In the β-linked polyester, the side chains are carboxylic acid groups capable of forming interstrand double hydrogen bonds. Evolution of the nucleic acids involved changes to the backbone and side chain of poly(β-D-malic acid). Conversion of the side chain carboxylic acid into a carboxamide or a longer side chain bearing a carboxamide group, allowed information polymers to form amide pairs between polyester chains. Aminoacylation of the hydroxyl groups of malic acid and its derivatives with simple amino acids such as glycine and alanine allowed coupling of polyester synthesis and protein synthesis. Use of polypeptides containing glycine and L-alanine for activation of two different monomers with either glycine or L-alanine allowed simple coded autocatalytic synthesis of polyesters and polypeptides and established the first genetic code. A primitive cell capable of supporting electron transport, thioester synthesis, reduction reactions, and

synthesis of polyesters and polypeptides is proposed. The cell consists of an iron-sulfide particle enclosed by tholin, a heterogeneous organic material that is produced by Miller-Urey type experiments that simulate conditions on the early Earth. As the synthesis of nucleic acids evolved from β-linked polyesters, the singlet coding system for replication evolved into a four nucleotide/four amino acid process (AMP = aspartic acid, GMP = glycine, UMP = valine, CMP = alanine) and then into the triplet ribosomal process that permitted multiple copies of protein to be synthesized independent of replication. This hypothesis reconciles the "genetics first" and "metabolism first" approaches to the origin of life and explains why there are four bases in the genetic alphabet.

1. Introduction

Despite considerable research devoted to studying how life started on Earth, no widely accepted theory of the origin of life has emerged. Such a theory would have to explain how an initial simple system was established and how the simple system evolved into the complex biological system. At the present time, there are many hypotheses that have varying degrees of support for different aspects of the problem. The most widely accepted hypothesis among biologists, the RNA world hypothesis, still has strong supporters [1] but difficulties with the hypothesis are recognized, especially the problem of synthesizing RNA in the absence of enzymes [2–7]. The RNA world hypothesis is central to the "genetics first" approach to the origin of life. An alternative approach involves a system in which cellular components coevolved [5,7,8]. Because nucleic acids are complex polymers and ribosomal protein synthesis is a highly complex process, a particularly challenging question is how did proteins and nucleic acids coevolve?

Many years ago, Lipmann [9] wrote, "In the domain of the transfer of genetic information I felt naïvely that, before translation into amino acid sequences could become the target of the genetic code, the extraordinary functional pliability of proteins would have to be established". Given what is now known about ribosomal protein synthesis, it does not seem to be such a naïve observation. Previously, I proposed that ribosomal protein synthesis evolved from coupled synthesis of nucleic acids and proteins. This proposal allowed coded protein synthesis to precede ribosomal protein synthesis [8]. Subsequently, I extended this hypothesis to describe how the coupled process evolved with continuity into ribosomal protein synthesis and, in so doing, provided an alternative to the RNA world hypothesis [5]. In order to develop a coevolutionary theory, the need for coded protein synthesis prior to ribosomal protein synthesis has assumed greater importance. In this paper, I present the argument for coupled protein and nucleic acid synthesis in an updated and different way from that which was published in 2000.

2. Autocatalysis in the Origin of Life

What seems to be a point of agreement in a contentious field is that the origin of life involved some type of simple autocatalytic system in which the products of reactions catalyzed their own synthesis.

Proteins often acting together with coenzymes/cofactors perform almost all of the biological catalysis. The complex autocatalytic system of biology involves symmetry breaking in which amino acids of only one configuration, L-amino acids, are incorporated into proteins to provide reproducible structures that catalyze reactions in the cell including synthesis of the L-amino acids themselves as well as numerous other chiral molecules including D-ribose in RNA and D-deoxyribose in DNA. All chirality in biochemistry begins with proteins. The enzymes involved are the aminoacyl-tRNA synthetases that preferentially bind L-amino acids. A reasonable assumption is that the simple autocatalytic system for the origin of life used proteins in symmetry breaking and catalysis. I will return to this issue later.

3. Energetics, Genetics, Catalysis and Membranes

Four major activities of cells are energy transduction, replication of genetic polymers, catalysis of chemical reactions, and formation of the membranes that define the three dimensional cellular space and the interaction of the cell with the environment. Each of these activities has to be addressed in a theory of the origin of life. It may be that replication of genetic polymers was not an initial requirement—perhaps growth without replication was possible at the beginning—but soon after the initiation of the simple system, genetic polymers would need to evolve so that selection could occur. How could a simple system incorporate these four activities? Using chemistry and biochemistry as a guide, I will suggest what these activities might have been leading to a model for a primitive cell.

4. Thioesters Were Used for Synthesis in Primitive Cells

Maps of cellular metabolic pathways show a central hub, the carboxylic acid cycles, a major spoke, gluconeogenesis/glycolysis, and a large number of other pathways leading from the hub and spoke to synthesis and degradation of nucleic acids, amino acids, lipids, carbohydrates, coenzymes and cofactors. Proponents of the "metabolism first" hypothesis for the origin of life suggest that reactions such as those found in the reverse citric acid cycle, the acetogenic pathway, or the methanogenic pathway may have occurred prior to the evolution of genetic polymers and enzymes [10–17]. The carboxylic acid cycles include molecules that form thioester bonds between carboxylic acids and coenzyme A (CoA), such as acetyl-CoA, succinyl-CoA and malyl-CoA. After the discovery of CoA and the role of phosphopantetheine in fatty acid and non-ribosomal peptide synthesis. Lipmann [9] proposed that use of thioester bonds provided a simpler biosynthetic mechanism for peptide bond synthesis than ribosomal synthesis. Subsequently, the idea of an early metabolic system powered by thioesters has been championed by other researchers [18–21]. CoA has an unusual and complex structure. Simple thiols such as CH_3SH formed from the reaction of CO_2, H_2S, and FeS [22] may have been used initially for thioester synthesis. The alternative view is that phosphate is so extensively used in metabolism that it had to be involved from the beginning [23–25]. If phosphate was not initially required, the model for a primitive cellular system is greatly simplified because there is no need for phosphate uptake and utilization.

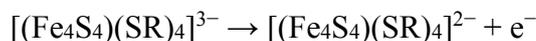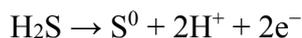## 5. Thioester Bonds Were Used to Form Ester and Amide Bonds

Given the choice, acyl groups move from thiol to hydroxyl or amino groups [26]. Therefore, generation of thioester bonds could have driven formation of ester and amide bonds. For example,

aminoacyl-thioesters form oligopeptides in the absence of enzymes [27,28], and α-thioglutamic acid oligomerizes [29,30]. A glyceroyl thioester oligomerizes chemically to form the polyester, oligoglyceric acid [31], and in biology ester bonds in poly(β-L-hydroxybutyric acid) (PHBA) are formed from β-L-hydroxybutyryl-CoA [32]. Therefore, an energy flux in primitive cells producing thioester bonds could have led to formation of ester and amide linked polymers in a simple system.
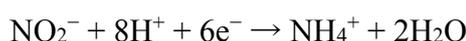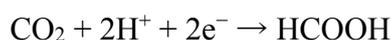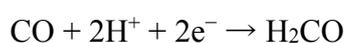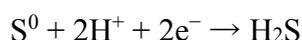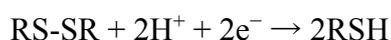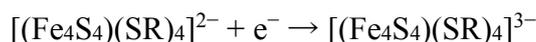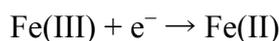
6. Energy Transduction in Biology

What was the earliest mechanism by which cells were energized? The following analysis of existing mechanisms suggests it was a redox gradient across a membrane. Energy transduction by respiration is an ancient process [33]. In "modern" cells, it is extraordinarily complicated, requiring formation of ion gradients across cell membranes and use of these gradients to drive ATP synthesis. The ion gradients are generated using the energy derived from redox processes in which electrons from oxidation of molecules such as NADH are transported to a receptor such as oxygen. Specialized lipid molecules, including ubiquinone and methanophenazine, and complexes of membrane proteins and cofactors, particularly heme groups and FeS clusters, are involved in electron transport and pumping of H^+ or Na^+ across a membrane to establish the gradient. A variety of electron donors and receptors are used in biochemistry. Conversion of the energy in a H^+ or Na^+ gradient into ATP synthesis uses a complex of proteins acting as a molecular motor, ATP synthase. The ion gradient appears to be secondary to the redox gradient suggesting that the redox gradient is the ancient means by which cells were energized. An alternative view suggested by geochemistry is that an acidic carbon dioxide-rich cool ocean and an alkaline hydrogen-rich warm interior of a metal sulfide deposit on the ocean floor near a hydrothermal system provided naturally occurring gradients across inorganic membranes [17,34].

How was a redox gradient established in a simple cell? An important feature of respiration is that the sites of the oxidation and reduction reactions are separated. In the case of NADH as electron donor and oxygen as electron acceptor, electrons can enter and leave the electron transport system on the inside of the membrane. The slow rate of reaction of NADH with oxygen allows this to occur. Generally, electrons can enter or leave electron transport systems from the inside or the outside of a membrane. In principle, a simple redox system would involve separation of the oxidation and reduction events by a membrane with transfer of electrons across the membrane into a primitive cell. The inside of the cell would therefore act as a cathode. Reduced products could accumulate inside primitive cells. What were the oxidation and reduction reactions involved in energy transduction in the simplest cells? The early Earth was anaerobic. Some of the possible electron donors on the primitive earth suggested by biochemistry are hydrogen, hydrogen sulfide, Fe(II), and sulfite. The Fe(II) is often part of an FeS cluster. The electron donor used previously [8] was sulfite produced by dissolving volcanic sulfur dioxide in water. I will use it again in this study.

$$H_2S \rightarrow S^0 + 2H^+ + 2e^-$$

$$H_2 \rightarrow 2H^+ + 2e^-$$

$$SO_3^{2-} + H_2O \rightarrow SO_4^{2-} + 2H^+ + 2e^-$$

$$Fe(II) \rightarrow Fe(III) + e^-$$

$$[(Fe_4S_4)(SR)_4]^{3-} \rightarrow [(Fe_4S_4)(SR)_4]^{2-} + e^-$$

What was the earliest electron acceptor? Some of the possibilities are shown below.

$$Fe(III) + e^- \rightarrow Fe(II)$$

$$[(Fe_4S_4)(SR)_4]^{2-} + e^- \rightarrow [(Fe_4S_4)(SR)_4]^{3-}$$

$$RS\text{-}SR + 2H^+ + 2e^- \rightarrow 2RSH$$

$$S^0 + 2H^+ + 2e^- \rightarrow H_2S$$

$$CO + 2H^+ + 2e^- \rightarrow H_2CO$$

$$CO_2 + 2H^+ + 2e^- \rightarrow HCOOH$$

$$NO_2^- + 8H^+ + 6e^- \rightarrow NH_4^+ + 2H_2O$$

Two of the simplest electron acceptors in biochemistry are a disulfide bond and $Fe(III)$ with $Fe(III)$ often part of an FeS cluster, such as $[(Fe_4S_4)(SR)_4]^{2-}$ (R = organic group). Sulfur-based respiration is commonly found in prokaryotes, especially hyperthermophilic archaea [35]. Three other possible electron receptors are CO, CO_2, and NO/NO_2^-. The atmosphere of the early Earth may have contained CO as well as CO_2 [36], allowing the possibility that CO was reduced to organic compounds inside cells. CO with FeS and thiols at elevated temperatures and pressures produces organic molecules including pyruvic acid via intermediary formation of Fe-carbonyl compounds [37]. FeS clusters, CO, and H_2 in the gas phase form formaldehyde and methanol [38]. Biochemically, reducing ("fixing") CO_2 requires complex enzyme catalyzed reactions [39]. CO_2 is a very stable molecule, and reduction to formate by H_2 is thermodynamically unfavorable, although the reaction of H_2 with bicarbonate is favorable [40]. In the ocean hydrothermal vent scenario, it is thought that the reactions of the acetogenic or methanogenic pathways for CO_2 fixation mimic ancient non-enzymatically catalyzed reactions [16,34]. H_2S can reduce CO_2 when FeS is oxidized to FeS_2 [10]. If the inside of a primitive cell acted as a cathode, CO_2 fixation could have been accomplished similar to the electrolytic process known as electrocarboxylation [40], as in electrolytic reduction of CO_2 to formic acid catalyzed by $[Fe_4S_4(SR)_4]^{2-}$ [41]. Simple redox reactions that occur chemically include oxidation of sulfite to sulphate and thiols to disulfides by $Fe(III)$ [42,43]. In this proposal I will use FeS as the primary electron acceptor. $NAD(P)^+/NAD(P)H$ and $FAD/FADH_2$ are commonly used redox reagents in biochemistry but their complex syntheses make it unlikely that they were present in an early redox system.

Previously, I assumed that early cells obtained reduced nitrogen for amino and amide chemistry by diffusion across primitive membranes of cyanides, and the thioamides and amides resulting from the reaction of cyanides with H_2S and water [8]. Diffusion of molecules such as formamide and acetamide remains an attractive possibility. An alternative explanation is that ammonium ions were produced by reduction of NO or nitrite inside cells. NO has been proposed as the first deep electron sink

for respiration [44]. NO produced in the atmosphere could have undergone reactions leading to accumulation of nitrite and nitrate in water [45]. If primitive cells were leaky to ions, nitrite and nitrate could diffuse across membranes and be reduced to ammonia using Fe(II) or FeS [46–48]. Ethanethiol and other simple thiols react with NO or nitrite to produce S-nitroso compounds [49]. Glutathione SH groups undergo the same reaction with NO in biochemistry [50]. If primitive cells were restricted towards ion transport, nitrite could have diffused into cells as the uncharged S-nitroso derivatives of thiols. Nitrite, released from the S-nitroso compound by hydrolysis inside the cells with formation of a metal thiolate, could be reduced to ammonia.

7. Formation of the Earliest Cell Membranes

If primitive cells were energized by electron transport across a cell membrane, what material constituted the membrane? Two very different views of the origin of cell membranes have been proposed. Both approaches seek to prevent dilution of metabolites in water by compartmentalization. One view is that the original cell membranes were formed from organic materials. Synthesis of membrane phospholipids for biological membranes is a complex biochemical process whether it involves the linkage of 2-carbon units from acetyl-CoA that occurs in bacteria and eukarya or the linkage of 5-carbon isoprene units that occurs in archaea. Lipids of this type were probably not available on the primitive Earth. Organic molecules for cell membranes could have been delivered to the early Earth by asteroids, meteors, comets, and interstellar dust particles (IDPs). Among a diverse range of organic molecules in meteorites are single chain amphiphiles such as fatty acids that are considered to be good candidates for prebiotic vesicles because they have higher permeability to ions than the double chain amphiphiles found in modern cell membranes [51–55]. Polyaromatic hydrocarbons (PAH) that are abundant in meteorites and IDPs [56] may have been included in vesicle membranes. Organic molecules could have formed on the early Earth from the interaction of energy sources such as UV light and lightning with volcanic gases, as demonstrated by Miller-Urey type experiments ([57], references therein), and accumulated in ponds near volcanoes. The other view is that the original membranes were formed from semi-porous inorganic materials, particularly metal sulfides, in deposits at hydrothermal vents on the ocean floor. Lipid-like molecules are not exuded from modern day vents. Lipid membrane synthesis is proposed to have occurred within metal sulfide bubbles eventually enabling cells to exist free of the sulfide deposits [17]. Another proposal is that the semi-porous metal sulfide deposits were at the Earth's surface associated with volcanic fields [58,59].

Membrane biogenesis involves insertion of newly synthesized membrane components, primarily lipids and proteins, into existing membranes. Accordingly, the easiest mechanism to consider for membrane biogenesis in primitive cells involves insertion of newly synthesized membrane components into pre-existing organic membranes. This seems more reasonable than creation of a *de novo* organic membrane within an inorganic membrane. Although most attention has been afforded to primitive lipid vesicles based on biological lipids, another possibility is that the organic material was formed from a different type of organic material. The most recognized products of Miller-Urey type experiments are carboxylic acids including amino and hydroxy acids [60], but there is also a water insoluble heterogeneous mixture of organic molecules called tholin [61]. Tholin is also produced by UV irradiation of gas mixtures and is found on the surfaces of Jupiter's moons, Ganymede and Callipso [62], and in the

haze of Saturn's moon, Titan [63–65]. Tholin can form laminated structures that may have formed the earliest cells [66]. Extractable components of tholin have a $C_iH_jN_k$ core terminally linked to amino and carboxylic acids [67]. Miller-Urey type experiments have been performed with H_2S in the mixture [60,68–70]. Tholin obtained from these experiments contains thiol groups [71]. Observations of the tholin in the haze around Titan show a complex mixture of hydrocarbons, including PAH, nitriles [64], and polyynes [65]. Polymers made from hydrogen cyanide, a major intermediate of amino acid synthesis in Miller-Urey type experiments, have an unsaturated structure in which the dominating polymers contain polyimine chain-like structures [72,73]. Therefore, tholin is a material with a high level of unsaturation including conjugated bonds and heterocyclic and polyaromatic ring components.

8. Energization of Primitive Cells Using Tholin as an Electron Transporting Organic Membrane

Accepting that the earliest electron transport system involved transport from the outside to the inside of a membrane, how could electrons have been transported across an organic membrane in the absence of complicated proteins/cofactors and specialized lipids that are found in biology (reviewed in [24,74])? One possibility comes from the observation that PAHs form quinone type molecules upon photoirradiation in the presence of water. These types of molecules might have accumulated in organic material on the early Earth and functioned in membranes like ubiquinone [75]. Another idea is that electron transport could have used a relatively simple type of cofactor such as an FeS cluster. Electron transport by FeS clusters bound to proteins generally occurs in protein domains associated with but not in membranes. Electrons can tunnel between FeS clusters [76]. It is possible that electrons were transported by FeS clusters embedded in a thiol bearing organic material. A third mechanism might involve the unsaturated components of tholin material. "Electrical" organic polymers contain conjugated unsaturated bonds [77]. Main chain peptide bonds can transport electrons [78]. Aromatic side chains in proteins can facilitate electron transport within and between proteins. An example is found in ribonucleotide reductase class I where the side chains of Tyr and Trp (using three letter abbreviations for amino acids) facilitate electron transfer from Tyr to Cys [79]. The nature of tholin suggests that it was an organic material that could transport electrons. The evolutionary pathway for membrane biogenesis would involve insertion into the tholin-containing membranes of specialized proteins, cofactors and lipids for electron transport, together with proteins performing other functions and lipids that did not transport electrons, eventually eliminating the need for tholin.

9. Iron Sulfide in Origin of Life Scenarios

If tholin formed primitive electron-transporting cell membranes, what was the electron acceptor and how was the energy from electron transport harnessed for synthesis within the cells? Because thioesters played a central role in primitive cell synthesis, the question becomes: How was electron transport linked to thioester synthesis? Before enzymes were available, this would probably have required an alternate type of catalyst. Building on an earlier proposal by Koch and Schmidt [80], I proposed that the first cells were made when tholin enclosed FeS using its thiol groups for binding to FeS. Within these cells, FeS particles or FeS clusters provided sites for catalysis of chemical reactions including sites for thioester synthesis [8]. Next, I will mention alternative ideas regarding metal sulfides and thioester synthesis in the origin of life, and in the following two sections, provide background on

the chemistry and biochemistry of FeS clusters and thioesters that will lead to the proposal of a simple FeS-catalyzed mechanism for thioester synthesis.

Many studies of the origin of life have focused on mineral surfaces as binding and catalytic sites for small molecules [14]. Because cofactors made from metal sulfides are often the sites of redox and catalytic activities in proteins and FeS, NiS, and other sulfides are likely to have been present on the primitive Earth in association with volcanic or hydrothermal activity, reactions involving transition metal sulfides, particularly FeS, have been extensively studied in simulations of reactions on the primitive Earth. Wächtershäuser [21,81] proposes a chemoautotrophic origin of life in an iron-sulfur world. The favored site for the origin of life is a hydrothermal vent where metal sulfides are deposited. Experimental reactions under hydrothermal conditions using H_2S as electron source, CO as carbon source, and NiS/FeS as catalyst produce a thioester, CH_3COSCH_3 [82]. In these studies, FeS is not a catalyst because it is converted to FeS_2. In a related proposal, FeS forms primitive cell membranes allowing concentration of reaction products in hydrothermal deposits [17]. In these cells, it is proposed that reactions analogous to the biological Wood-Ljungdahl acetogenic pathway for thioester synthesis could have occurred in the absence of enzymes using H_2 and CO_2 energized by redox, proton, and temperature gradients [34]. Enzymatically, this is a complicated pathway. Cofactors involved are folic acid, a corrin ring, and an FeNiS complex in a process that involves reduction of CO_2 separately to CO and a methyl group, which then react with CoA to form acetyl-CoA [83].

10. FeS Cluster Biochemistry and Chemistry

FeS clusters bound to proteins, exemplified by $[Fe_4S_4]^{2+}$, have both redox and non-redox functions [84]. The ability of FeS clusters to carry different charges allows FeS proteins to be involved in redox reactions, often in electron transport chains. An extensively studied class of FeS proteins is ferredoxin, a mobile carrier of electrons, which has Fe_4S_4 or Fe_2S_2 clusters bound to cysteine side chains as thiolate ligands. Ferredoxin is thought to be an ancient FeS protein in which the clusters cycle between reduced and oxidized forms through one-electron transfers [85,86]. Use of ferredoxin as a reducing agent is demonstrated by pyruvate synthase (pyruvate ferredoxin oxidoreductase)-catalyzed reductive carboxylation of acetyl-CoA to form pyruvate in the presence of thiamine pyrophosphate (TPP) [87,88].

Iron in FeS proteins is usually coordinated by Cys thiolate anions but oxygen can provide ligands to Fe in FeS clusters during enzymatic catalysis [89]. In the active site of aconitase, one of the Fe in a Fe_4S_4 cluster binds to water, a carboxylate and a hydroxyl group during interconversion of citrate and isocitrate in a catalytic mechanism that does not involve a redox component. Fumerase uses FeS clusters to catalyze interconversion of malic and fumaric acids [90]. Aconitase and fumarase are examples of a wide range of dehydratases with FeS clusters that catalyze dehydration reactions producing carbon-carbon double bonds [89].

FeS clusters like $[Fe_4S_4(SR)_4]^{n-}$ and $[Fe_2S_2(SR)_2]^{n-}$ that are similar to those found in FeS proteins have been synthesized chemically [91–93]. In the absence of proteins, they tend to be unstable and insoluble in water [94], but are stabilized by the presence of thiols [95]. Their reactions have mostly been studied in organic solvents. As in biochemistry, $Fe_4S_4(SR)_4$ can exist in different ionic states with more negatively charged clusters acting as reducing agents [92]. Reactions that do not involve electron transfer can produce clusters in which Fe is bound to oxygen. For example, the reaction of $Fe_4S_4(SR)_4$

with acetic acid produces an acetate-bound cluster and RSH. Reaction with the more electrophilic carbonyl of acetyl chloride produces a thioester, CH_3COSR [91].

11. Biochemical Reduction of Disulfides and Thioester Formation

The ability of the thiol group to carry an acyl group as a thioester and be oxidized to a disulfide provides the most obvious simple link in early cells between electron transport and chemical energy for biosynthesis. The dithiol/disulfide redox reaction is reversible under mild conditions. Two enzymes illustrate a connection between FeS cluster chemistry and disulfide reduction. Ferredoxin-thioredoxin reductase (FTR) catalyzes transfer of electrons from reduced ferredoxin to thioredoxin [96]. In the mechanism of the reaction, two electrons, one from a reduced ferredoxin and one from the FeS cluster in FTR, are used to cleave an adjacent cystine. The thiolate produced attacks the disulfide of thioredoxin. A second one-electron transfer from reduced ferredoxin completes the reduction of the thioredoxin disulfide. A key feature of the reaction is the ability of the Fe_4S_4 cluster of FTR to vary its charge and allow one Fe to adopt five-fold coordination during the reaction. Heterodisulfide reductase (HDR) catalyzes two one-electron transfers to the disulfide formed between coenzyme B (CoB) and coenzyme M (CoM). CoB-S-S-CoA is the terminal electron receptor in methanogenesis [97]. A mechanism similar to FTR is involved but the reaction is not mediated by a cystine. CoMS-SCoB interacts directly with a Fe_4S_4 cluster and the sulfur of CoM binds to the cluster [98,99].

The equilibrium for thioester synthesis from a thiol and carboxylic acid is highly unfavorable for formation of a thioester and water. Consequently, thioester synthesis requires coupling to an energy releasing reaction. Mechanisms for thioester biosynthesis have redox components. In synthesis of acetyl-CoA from pyruvate catalyzed by pyruvate dehydrogenase, a critical reaction is between a lipoamide molecule bearing a disulfide bond and an acetaldehyde formed by decarboxylation of pyruvate that is bound to TPP as a hydroxyethyl group. Formation of the acetyl thioester derivative of lipoamide involves oxidation of the acetaldehyde to an acetyl group and reduction of the lipoamide disulfide. Subsequently, the acetyl group is transferred to CoA. This reaction provides a biochemical link between disulfide reduction and thioester synthesis. Formation of a thioester from an aldehyde is also involved in the mechanism of the reaction catalyzed by glyceraldehyde-3-phosphate dehydrogenase. In most organisms, the thioester is produced by oxidation of a hemithioacetal formed between a Cys-SH and glyceraldehyde-3-phosphate using NAD^+ as the oxidizing agent, but in some cases ferredoxin is used as the electron acceptor instead of NAD^+ [100,101].

One can conclude that thioester synthesis in primitive cells was likely to be associated with FeS redox chemistry and a simple mechanism for thioester synthesis could have involved coupling to reduction of a disulfide bond.

12. A Proposal for Primitive Cellular Synthesis of Thioesters

From the characteristics of reactions involving thiols, disulfides, carboxylic acids, and FeS compounds, and the analysis of chemical and biochemical thioester syntheses, it is plausible that FeS chemistry played a role in prebiotic thioester synthesis. My proposal is that electrons generated from oxidation reactions outside primitive cells were transported and used inside the cells for formation of thioesters by the following reaction mediated by FeS clusters or the surface of an FeS particle. Intermediate formation

of an aldehyde by reduction of the carboxylic acid prior to reaction with the disulfide may not have been necessary.

$$RS\text{-}SR + 2e^- + 2H^+ + R'COOH \rightarrow RSH + R'COSR + H_2O$$

Using sulfite oxidation as the electron source, the overall reaction is

$$RS\text{-}SR + SO_3^{2-} + R'COOH \rightarrow RSH + R'COSR + SO_4^{2-}$$
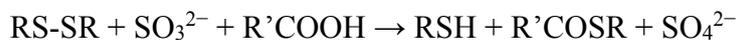
A proposal for acetyl-thioester synthesis based on the HDR/FTR mechanism using the surface of an FeS particle is given in Figure 1. In the presence of thiols and carboxylic acids, FeS could have provided a simple catalytic surface for thioester synthesis without formation of FeS$_2$.
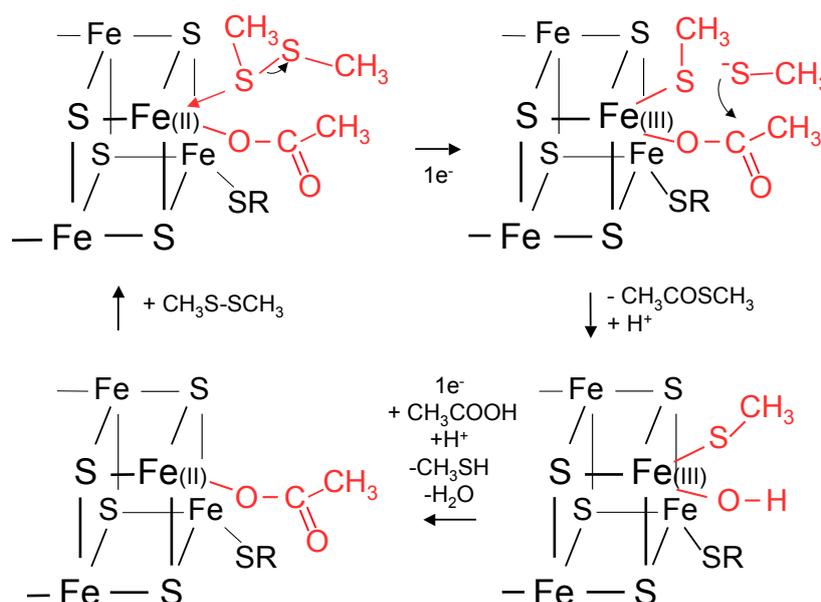
Figure 1. Mechanism proposed for catalysis of thioester synthesis in primitive cells by an FeS particle. A thiolate anion from reduction of a disulfide, CH$_3$SSCH$_3$, reacts with acetate bound to Fe of an FeS particle to form a thioester, a thiol, and hydroxylated Fe. Using the mechanisms of FTR and HDR as models, one of the two electrons for reaction with the disulfide is provided by electron transport from an oxidation reaction and the other electron by conversion of Fe(II) to Fe(III). Transfer of another electron regenerates the FeS binding site and loss of water permits a new round of thioester synthesis.

13. Diffusion of Small Polar and Nonpolar Molecules across Primitive Cell Membranes

Simple molecules such as water, carbon dioxide, oxygen, nitrogen, hydrogen, hydrogen sulfide, and nitric oxide diffuse across biological membranes but transport of ions and organic molecules involves integral membrane protein transporters or channels. If primitive cells were leaky to ions, as they would be if they were composed of fatty acids [52], and many different molecules and ions were present in the environment, the number of possible cellular components would have been high. Miller-Urey type experiments produce a large number of molecules including carboxylic acids, amino acids, and hydroxyacids although their relative amounts vary widely. It is thought that this was highly beneficial for the origin of life because these molecules could have been assembled into macromolecules by the

action of condensing agents like carbonyl sulfide [102] or through wet dry cycles using salts [103] to produce a wide variety of possible catalytic activities. In this line of reasoning, cellular membrane evolution involved increasing the barrier provided to ion permeability [25,104]. On the other hand, leaky membranes would have meant that the inside and outside of a cell would have been equilibrated with regard to small molecules, which would have made it difficult to synthesize macromolecules. Those who advocate an inorganic membrane for early cells refer to its semiporous nature, which would allow diffusion of CO_2 from the environment, and have the inorganic compartment accessible to molecules from within the metal sulfide deposits [17,58].

I favor early organic membranes that presented a barrier to the passage of ions including amino acids, but were permeable to simple non-polar and polar molecules including H_2S, CH_3SH, CO_2, CO, $HCONH_2$ and CH_3SNO. It is difficult to stop protons from leaking across lipid bilayer membranes [105], and presumably the earliest organic membranes did not have the ability to prevent proton transfer. At first sight, it seems odd that amino acids might be available in the environment but not used inside cells, but cells that were not leaky to ions would not be equilibrated with the external environment, energy from redox reactions could be used to maintain cells away from thermodynamic equilibrium, and the buildup of macromolecules could proceed. Monomers, including amino acids, required for assembly into macromolecules would have been synthesized inside cells from diffusible materials.

14. Reduction and Carboxylation Reactions inside Primitive Cells

In addition to thioester synthesis, other reactions might have occurred in these primitive cells either at the FeS surface or by reduction reactions with thiols in solution that would be required to continue thioester synthesis by recycling of thiols to disulfides. Possible reactions, listed in Table 1, are described below.

Reductions of CO, CO_2, and NO involving FeS have been mentioned previously. Another potential reduction reaction would be conversion of a carboxyl group to an aldehyde. Biochemistry often overcomes the energy barrier for this reaction by intermediate formation of an acyl phosphate before reduction with NAD(P)H, but fatty acid reduction to an aldehyde uses a thioester with CoA as substrate [106]. De Duve [107] speculates that thioester reduction to an aldehyde, perhaps using FeS as a catalyst, was the earliest mechanism. Direct evidence for aldehyde formation from thioesters catalyzed chemically by FeS or FeS clusters has not yet been reported. The reaction would require transfer of hydride for C–H bond formation. However, lactic acid synthesis from pyruvic acid catalyzed by $FeS/H_2/H_2S$ produces a C–H bond [108]. The mechanism for this reaction indirectly involves H_2. Reduction of α-keto acids is also found using FeS/H_2S [109]. The Fukuyama reduction reaction that is used chemically for aldehyde synthesis from thioesters involves formation of an acyl-Pd-H intermediate [110]. An acyl-Fe-H intermediate may be involved in FeS catalyzed lactic acid synthesis, but a carboxylate-Fe-H intermediate is also possible.

Table 1. Possible reactions in a primitive cell formed by enclosure of an FeS particle in tholin.

Activity	Reaction
1. Thioester synthesis	$CH_3COOH + 2e^- + 2H^+ + RSSR \rightarrow CH_3COSR + RSH + H_2O$
2. CO reduction	$CO + 2e^- + 2H^+ \rightarrow H_2CO$
3. CO_2 reduction	$CO_2 + 2H^+ + 2e^- \rightarrow HCOOH$
4. Nitrite reduction	$NO_2^- + 8H^+ + 6e^- \rightarrow NH_4^+ + 2H_2O$
5. Reductive carboxylation	$CH_3COSR + 2e^- + CO_2 + 2H^+ \rightarrow CH_3COCOOH + RSH$
6. Reductive amination	$CH_3COCOOH + 2RSH + NH_3 \rightarrow CH_3CH(NH_2)COOH + RSSR + H_2O$
	$CH_3COCOOH + 2e^- + 2H^+ + NH_3 \rightarrow CH_3CH(NH_2)COOH + H_2O$
7. Glycine synthesis	$H_2CO + NH_3 + CO_2 + 2RSH \rightarrow HOOCCH_2NH_2 + H_2O + RSSR$
8. Claisen condensation	$CH_3COSR + HCOCOOH \rightarrow HOOCCH(OH)CH_2COSR$
9. Carboxylation	$CH_3COCOOH + CO_2 \rightarrow HOOCCH_2COCOOH$
10. Reduction of ketone	$HOOCCOCH_2COOH + 2e^- + 2H^+ \rightarrow HOOCCH(OH)CH_2COOH$
11. Reductive thiolation	$HCOCOOH + 2e^- + H_2S + 2H+ \rightarrow HSCH_2COOH + H_2O$
12. Aldol condensation	$2CH_3COCOOH \rightarrow HOOCC(CH_3)OHCH_2COCOOH$
13. Acyl group transfer	$R'COSR + R''OH \rightarrow R'COOR'' + HSR$
	$R'COSR + R''NH_2 \rightarrow R'CONHR'' + HSR$
14. β-Polyester synthesis	$-(COCH_2CH(COOH)O)_n- + RSCOCH_2CH(COOH)OH \rightarrow$
	$-(COCH_2CH(COOH)O)_{n+1}- + RSH$
15. Polypeptide synthesis	$-(NHCHR'CO)_n- + H_2NCHR'COSR \rightarrow -(NHCHR'CO)_{n+1}- + HSR$

Early expansion of metabolism required the synthesis of three- and four-carbon molecules. Pyruvic acid, a pivotal three-carbon molecule in biochemistry, is made by carboxylation of acetaldehyde bound to TPP using reduced ferredoxin for two one-electron transfers catalyzed by pyruvate ferredoxin oxidoreductase [111,112]. A reductive carboxylation reaction of this type using dithionite as electron donor is chemically catalyzed by an FeS complex but the reaction is performed in a partially aqueous solution [113]. An FeMoS cluster complex catalyzes electrolytic production of pyruvic acid from an acetyl thioester and CO_2 in non-aqueous solution. A thioester is required because other acetylated compounds do not yield pyruvic acid. Lactic acid is not reported to be a product [114]. Pyruvic acid is also produced under simulated hydrothermal vent conditions from alkyl thiols and CO in the presence of transition metal sulfides at 250 °C [37]. Hence, TPP-independent synthesis of pyruvic acid from an acetyl-thioester and CO_2 catalyzed by FeS is possible.

Carbonyl groups of aldehydes and ketones react with thiols to form thioacetals and thioketals and the reaction does not proceed further to form hydroxyl groups and disulfides. Indeed, this reaction is employed for protection of aldehydes and ketones in organic syntheses [115]. In contrast, imino groups add thiols and the reaction can proceed to produce amino groups and disulfides [116,117]. Carbonyl groups in glyoxylate and pyruvate could have been converted into imines with NH_4^+ and then into Gly and Ala with thiols, regenerating disulfides. Synthesis of amino acids by reductive amination of α-ketoacids occurs in the presence of FeS and NH_4^+ or amines, but not in the presence of Fe(II) or other transition metal sulfides [118], and Ala is produced in mixtures containing pyruvic acid/FeS/H_2S/H_2/NH_4Cl [108].

Glycine synthesis could also have involved reactions similar to those catalyzed by glycine synthase [119] in which NH_4^+, formaldehyde, and CO_2 are converted into Gly with oxidation of dihydrolipoamide to the disulfide form. Formation of aminomethyllipoamide is non-enzymatic [119], pyruvic acid may have been involved instead of pyridoxal pyrophosphate (PLP) [120], and reductive

carboxylation of imines to amino acids has been demonstrated electrolytically [121,122]. Non-enzymatic synthesis of Gly by transamination of glyoxylic acid using Ala in the absence of PLP is also known [123].

Among the earliest four-carbon molecules synthesized in primitive cells would have been oxaloacetic, malic and aspartic acids. Malic acid is synthesized in biology by condensation of acetyl-CoA with glyoxylic acid catalyzed by malate synthase [124], reductive carboxylation of pyruvic acid by malic enzyme [125,126], hydration of fumaric acid by fumarase [90], and reduction of oxaloacetic acid by malate dehydrogenase [127]. A non-enzymatic malate synthase-like route to malic acid would require base catalyzed formation of the enolate anion of an acetyl-thioester. Another path to malate may have involved intermediate formation of oxaloacetate, which is synthesized biochemically using pyruvate carboxylase, a biotin dependent carboxylase using bicarbonate [128]. In this reaction, the enolate anion of pyruvic acid is stabilized prior to interaction with a carboxyl group. The enol form of oxaloacetate is stabilized by an FeS center [129]. Non-enzymatic synthesis of malic acid from oxaloacetic acid, perhaps by reduction of FeS-bound oxaloacetic acid, would require formation of a C–H bond in the same way as reduction of pyruvic acid to lactic acid. FeS/H_2S reduces oxaloacetic and pyruvic acids to malate and lactate [109]. Reductive amination of oxaloacetic acid would yield Asp.

Simple carboxylic acids containing thiol groups could have been available in primitive cells. Reductive thiolation of a Schiff base formed between an amino acid and glyoxylic acid could have yielded thioglycolic acid. This method is used to introduce thiol groups into coenzymes M and B [130,131]. Thiolactic acid is one of the products of the reaction of pyruvic acid and FeS/H_2/H_2S/CO_2 [108].

The feasibility of reducing carboxylic acids to aldehydes for aldol condensation reactions is not well established in FeS cluster chemistry, but a reduced aldol condensation product is obtained from pyruvic acid/FeS/H_2/H_2S mixtures [108]. This result implies an ability of the FeS surface to stabilize the enol or enolate form of pyruvic acid. Stabilizing enols may be a property of FeS clusters. Fumarase A catalyzes the isomerization of enol and keto isomers of oxaloacetic acid [129].

In conclusion, the available evidence suggests that reduction and other reactions that are catalyzed by FeS or involve oxidation of thiols to disulfides could have led to the synthesis of molecules inside a primitive cell, including Gly and Ala for α-polypeptide synthesis, and malic acid for β-polyester synthesis. Activation of amino and hydroxy carboxylic acids as thioesters would facilitate acyl group transfer and polymerization.

15. Base Catalysis in Early Metabolism

Some of the earliest anabolic reactions would have involved synthesis of three- and four-carbon molecules. It is possible that early reaction mechanisms differed considerably from enzyme reaction mechanisms. For example, greater use may have been made of CO instead of CO_2 in carbon fixation. If modern mechanisms do reflect the earliest mechanisms, base catalyzed removal of protons would have been advantageous for many reactions. Acid/base catalysis is one of the most important functions provided by enzymes. The imidazole side chain of His, and the carboxylic acid side chains of Asp and Glu are the most widely used acid/base catalytic groups in proteins. An extractable fraction of tholin contains terminal carboxyl and aminoacyl groups [67], and some of these groups would be expected to locate on the inside of the tholin membrane facing the FeS particle. Hence, tholin might also have supplied an early acid/base catalytic function through –COOH/COO–. For example, bringing a thioester

and a catalytic base into proximity would be important for Claisen condensation reactions involving intermediate formation of enolates of acetyl-thioesters. Malate synthase uses this mechanism with glyoxylic acid as substrate and Asp as the catalytic base for formation of the enolate anion of acetyl-CoA [124].

16. A Model for the Simple System within a Primitive Cell

The preceding analysis supports the proposal that FeS may have acted as a mediator of reduction reactions and thioester synthesis as proposed by Wächtershäuser [21], but suggests that these reactions occurred using FeS as a catalyst within an organic membrane formed from tholin. Electrons released by the oxidation of sulfite could have been transported through tholin and delivered to FeS bound to tholin through thiol-groups. Freshly prepared particles of FeS have a highly rugged surface and contain a "library of different FeS clusters" [132]. Cavities between the FeS particle and tholin could have contained water. A model is proposed for the reactions within such a primitive cell that led to the synthesis of polypeptides and polyesters (Figure 2). This would be a simple system from which biology could evolve. FeS would be a true catalyst because it would not be converted into FeS_2. The flux of reducing power into these cells allowed thioesters rather than ATP to be the main source of chemical energy for polyester and polypeptide synthesis and thiols rather than NADPH and reduced ferredoxin to be the mobile source of reducing power for synthesis in primitive metabolism. The interior of primitive cells would be reducing, as it is in all extant cells.
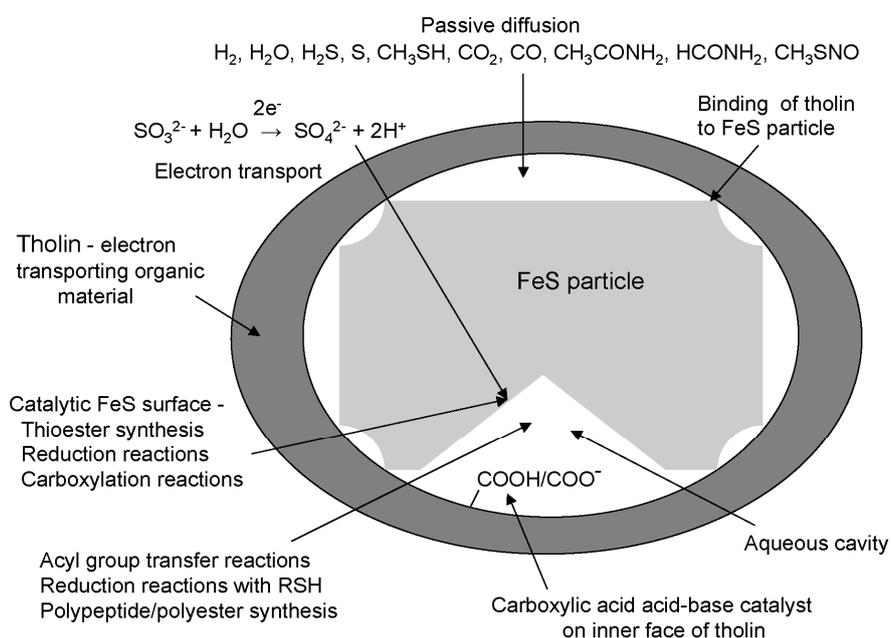
Figure 2. Model of a primitive cell. An FeS particle is enclosed in an organic membrane composed of tholin. Electron transport across the membrane to the FeS particle catalyzes thioester formation, carboxylation, and reduction reactions at the particle surface. Disulfides are reduced to thiols during thioester synthesis and regenerated through reduction reactions for further thioester synthesis. Transfer of acyl groups from thiol to hydroxyl and amino groups drives formation of polyesters and polypeptides. Carboxylic acid groups on the inside surface of tholin provide acid/base catalysis.

17. Genetic Polymers Were Initially Produced from Aliphatic Monomers

So far, I have not mentioned genetic polymers. Previous experimental work on the evolution of nucleic acids has focused primarily on simplifying the nucleic acid backbone, leaving the purines and pyrimidines for base pairing (reviewed in [133]). At the hub of metabolism are the carboxylic acid cycles, suggesting that the metabolites of early cells would have been carboxylic acids and their amino, hydroxy, keto, and thioester derivatives. All of these molecules are aliphatic. Aromatic molecules and molecules with aromatic components have long biosynthetic pathways starting from aliphatic molecules. Therefore, purines and pyrimidines, the aromatic amino acids, the quinone components of ubiquinone and menaquinone, nicotinamide of NADH, isoalloxazine of $FADH_2$, folic acid, pterin ring systems, and other aromatic compounds are unlikely to have been involved in cellular activity until pathways for their syntheses had evolved.

Hydrogen bonding between the purines and pyrimidines of nucleic acids is the foundation of genetics. How did genetics start if purines and pyrimidines were not available from the environment or from synthesis inside early cells? Consideration must be given to the possibility that genetic polymers started with aliphatic side chains rather than purines and pyrimidines. Returning to the structure of PHBA, which is a simple linear β-polyester formed from β-L-hydroxbutyryl-CoA monomers, it is interesting that the methyl groups occur at repeating positions along the polyester chain and the homochiral nature of the polymer ensures that the methyl side chains all point in the same direction relative to the polyester backbone (Figure 3A).
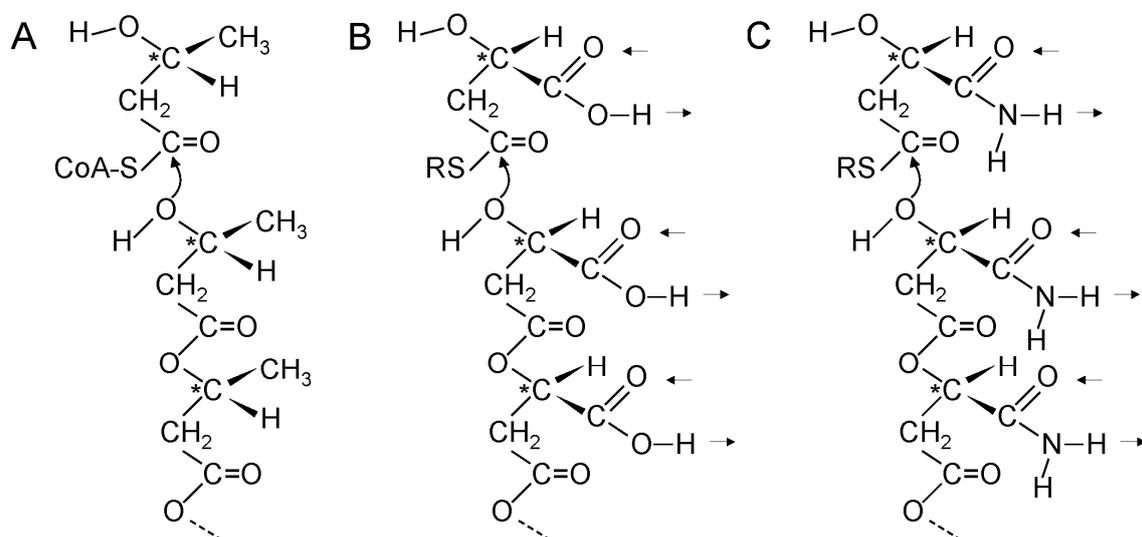
Figure 3. β-Polyester synthesis. (**A**) Enzyme catalyzed synthesis of poly(β-L-hydroxybutyric acid). The CH_3 side chains in poly(β-L-hydroxybutyric acid) occur at repeating distances along the polyester chain and point in the same direction relative to the backbone; (**B,C**) Proposed non-enzymatic synthesis of poly(β-D-malic acid) (B) and poly(β-D-malamide) (C) in primitive cells. The double hydrogen bonds of the side chains occur at repeating distances and are in the same direction relative to the backbone. Arrows indicate directions of possible hydrogen bonds. * indicates chiral carbon atom.

Could a simple β-polyester with a different side chain have formed interstrand hydrogen bonds? The simplest functional groups in organic chemistry that can form at least two roughly parallel hydrogen bonds are carboxylic acids and carboxamides. Malic acid is a β-hydroxyacid and the simplest β-polyester with carboxylic acid side chains is water soluble poly(β-malic acid) [134–136] (Figure 3B). Most notably, malic acid is a central component of the carboxylic acid cycles. Consequently, it is likely to be been available early in the origin of life. β-polyesters formed from a single optical isomer are more hydrolytically stable than those formed from a racemic mixture and are more stable than α-polyesters [137]. Hence, chiral polymers would accumulate preferentially. Homochiral poly-β-malic acid would have carboxylic acid side chains all pointing in the same direction relative to the backbone. If D-malic acid were used for β-polyester synthesis it would be the predecessor of D-ribose that is found in the backbone of RNA. L-malic acid is used primarily in the carboxylic acid cycles.

A carboxyl group can accept hydrogen bonds via its oxygen atoms and donate a hydrogen bond through its hydroxyl group. Two hydrogen bonds are used when a carboxylic acid forms an intermolecular hydrogen bonded cyclic dimer. Dimerization is possible when the carboxyl group is unionized, which might have been the case under the slightly acid conditions on the early Earth resulting from a high pressure of CO_2 [138]. Does poly(β-D-malic acid) form cyclic double hydrogen bonds between the carboxylic acid side chains producing a double stranded polyester in water? At present, there is no direct evidence to answer this question. Studies of poly(β-L-malic acid), which presumably has properties similar to poly(β-D-malic acid), have not distinguished the types of hydrogen bonding [139]. Formation of interstrand hydrogen bonds between carboxyl groups would be expected to involve a balance between hydrogen bonding to another carboxyl group and hydrogen bonding to water, similar to interstrand hydrogen bonding in DNA. Double stranded DNA is converted into single stranded DNA by increasing the temperature above the melting temperature. Bases in single stranded DNA form hydrogen bonds with water. The melting temperature for a double-stranded poly(β-D-malic acid) would likely be considerably lower than that of double stranded DNA which has triple hydrogen bonded GC base pairs and has been optimized for base pairing during the evolution of DNA.

Some carboxylic acids, including malic acid [140], crystallize with cyclic double hydrogen bonds between carboxyl groups and others, including acetic acid [141], crystallize with a catemer motif in which each carboxyl hydrogen bonds to two others in a chain. However, in water acetic acid forms mainly cyclic dimers [142]. Lactic acid dissolved in water is also present as the cyclic dimer [143]. Indeed, the cyclic dimer seems to be the lowest energy state for a carboxylic acid in water Consequently, hydrogen bonding between carboxyl groups is preferred over hydrogen bonding to water at ambient temperatures. If the carboxylic acid side chains of poly(β-D-malic acid) behave in a similar manner, interstrand double hydrogen bonding would be preferred over single stranded hydrogen bonding with water. They might even form cyclic double bonds cooperatively. However, the enthalpy gained by hydrogen bonding would have to offset the loss of entropy involved in creating a more ordered double stranded structure. In summary, it is possible that poly(β-D-malic acid) forms interstrand double hydrogen bonds especially at low temperatures and under slightly acidic conditions.

The hydroxyl group of malic acid evolved into the 3'-OH of ribose and the β-carboxylic acid into the 5'-CH_2OH. An early evolutionary development for the side chain may have involved use of malamide thioester monomers instead of malic acid thioester monomers (Figure 3C). Proposed steps in the evolution of the β-polyester backbone into the phosphoribose backbone of RNA are shown in Figure 4.

Incorporation of phosphate into the backbone required evolution of a mechanism for phosphate uptake and utilization, possibly by reaction with an acetyl-thioester to generate acetyl-phosphate. The ribose-phosphate backbone was selected because it allowed the fully aromatic bases in RNA to stack at the distance and positions required for maximum stability.
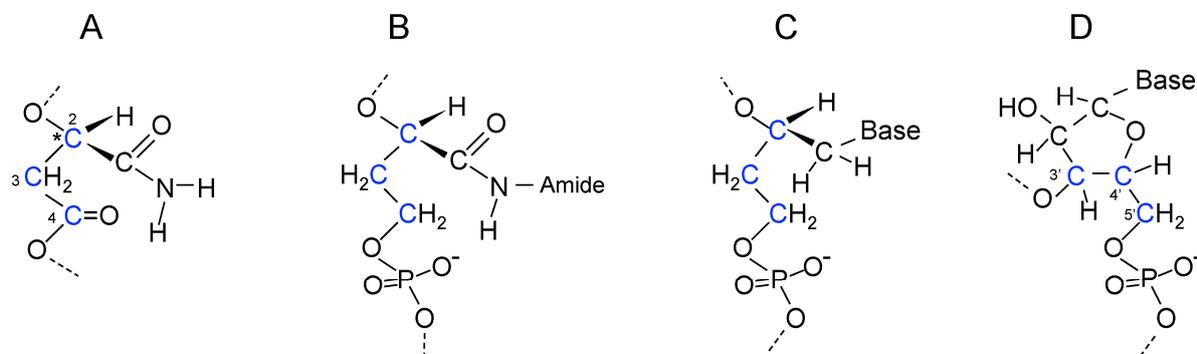
Figure 4. Evolutionary pathway proposed for the RNA backbone. (**A**) A malamide monomer in poly(β-D-malamide). C2, C3, and C4 carbon atoms are highlighted in blue; (**B**) Reduction of the 4-carboxyl group of malamide to a hydroxymethyl group forms D-2,4-dihydroxybutyramide allowing phosphodiester linking of monomers; (**C**) Reduction of the carboxamido carbonyl group of D-2,4-dihydroxybutyramide to a methylene group produces a D-3-deoxyerythritol backbone that forms phosphodiester links and binds to amines or heterocyclic bases through C1; (**D**) Replacement of the D-3-deoxyerythritol backbone by 3'-5' phosphodiester linked D-ribose. Conservation of the C2, C3, and C4 atoms of malamide as the 3', 4', and 5' carbon atoms of ribose is shown in blue. Evolution of the polymer side chains into pyrimidines and purines is discussed in Section 18 in relation to Figures 5A,B and 6. * indicates chiral carbon atom.

18. Evolution of the Purines and Pyrimidines of Nucleic Acids from Aliphatic Side Chains

An information polymer has to have at least two different side chains. Poly(β-D-malic acid) is not an information polymer. The carboxylic acid side chains of poly(β-D-malic acid) are obviously a long evolutionary distance from purines and pyrimidines and the pathway to them would involve multiple steps, which will be described in the following paragraphs. As mentioned previously, the first step could have been the use of D-malamide instead of D-malic acid, which would have provided a unionizable side chain (Figures 3C and 4A). Transamidases catalyze conversion of carboxylic acids into amides with Gln often acting as the ammonia donor [144]. In the primitive setting, amides such as acetamide and formamide may have provided the amide NH_2. In the absence of enzymes, transamidation reactions occur in the presence of Fe(III) and water using different NH_2-containing reagents [145]. Modification of this amide side chain with a β-carboxamide such as Asn or β-alaninamide formed by α-decarboxylation of Asn could have produced a longer side chain that could form a double hydrogen bond in the same direction as the unmodified side chain (Figure 5A,B). At this point, the β-polyester formed from monomers with two different side chains became an information polymer. Hydrogen bonding between polyester strands using different amide side chains allowed "amide pairing".
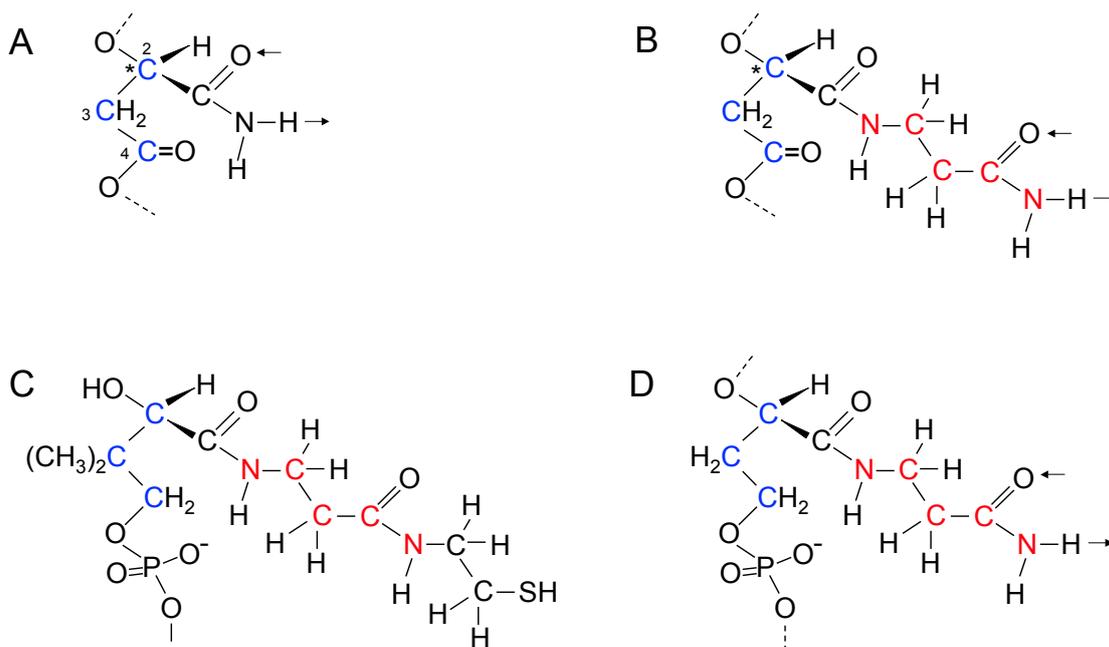
Figure 5. Poly(β-ᴅ-malamide) as an information polymer. (**A**) Malamide in a β-polyester strand with backbone carbon atoms shown in blue and amide side chain double hydrogen bonding shown by arrows. The 2', 3', and 4' carbon atoms of malamide are labeled and highlighted in blue; (**B**) Proposed extended side chain of malamide with β-alaninamide in a polyester strand. Carbon and nitrogen atoms in the extended side chain are highlighted in red. Hydrogen bonding is in the same direction relative to the backbone as that shown in A; (**C**) Structure of the phosphopantetheine component of CoA; (**D**) Structure of a proposed intermediate in nucleic acid evolution with a phosphodiester-linked backbone and extended malamide side chain, showing the similarities to phosphopantetheine. Two methyl groups in phosphopantetheine are not present in the proposed nucleic acid predecessor. * labels the chiral carbon atom.

De novo synthesis of pyrimidines and purines presents a conundrum. Many of the steps are dependent upon energy provided by ATP or GTP hydrolysis. These enzymes may have replaced earlier enzymes that used a simpler energy source such as pyrophosphate, acetyl-phosphate, phosphoenol pyruvate, or a thioester. The reactions in purine synthesis that formylate the amino group of 5-aminoimidazole-4-carboxamide and then cyclize the product to form inosine with loss of water do not involve ATP hydrolysis and might have occurred non-enzymatically or with simple polypeptides as catalysts, perhaps using a formyl-thioester as substrate. Two enzymes in pyrimidine synthesis, carbamoyl-phosphate synthase and CTP synthase, use ATP as a substrate, and two other steps involve complicated enzymatic mechanisms. Overall, evidence from biochemistry does not provide many clues to explain the evolution of pyrimidine and purine synthesis. Nevertheless, I will propose pathways starting with the extended side chain shown in Figure 5B.

Examination of the structures of purines and pyrimidines suggests the way in which the modified aliphatic side chains evolved into purines and pyrimidines. In Figures 5 and 6, carbon and nitrogen atoms in the extended aliphatic side chain are highlighted in red. Rotations allow the side chain to adopt conformations that are *cis*- and *trans*-like with respect to the two methylene groups of β-alaninamide and
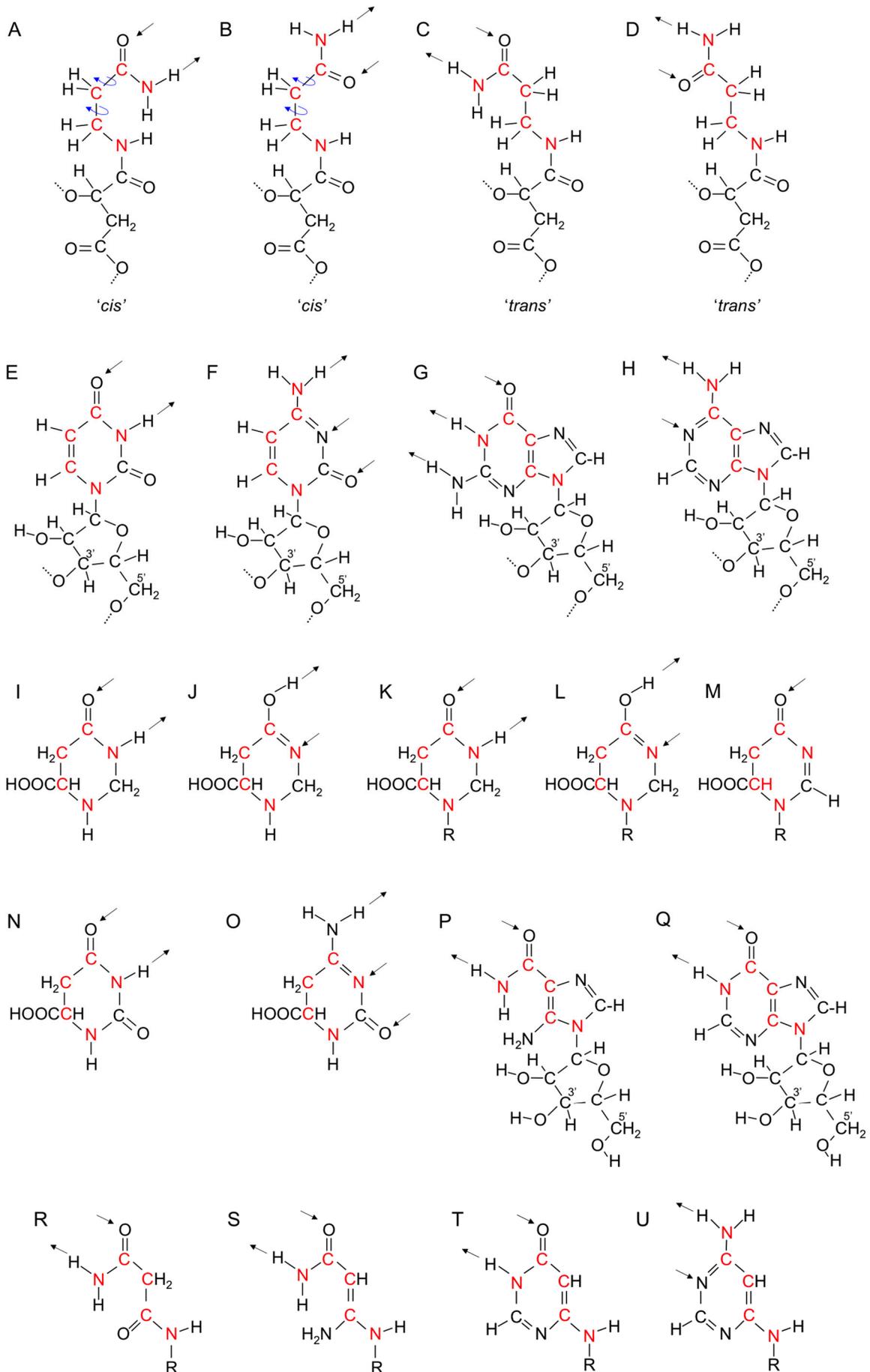
form double hydrogen bonds in two conformations (Figure 6A–D). The evolutionary pathways to the bases involve cyclization of the *cis*-like side chain to form pyrimidines and cyclization on each side of the *trans*-like side chain to form purines (Figure 6E–H).

The general approach to ring formation in pyrimidine and purine synthesis is to form carbon-nitrogen bonds by removal of water. Asn or β-alaninamide might be expected to provide the hydrogen bonding atoms at the 3 and 4 positions of pyrimidines. Indeed, one of the simplest ways of making a pyrimidine ring is the reaction of Asn with formaldehyde. Cyclization by crosslinking the amino and amido groups yields a tetrahydropyrimidine ring, tetrahydroorotate, which exists as two epimers in equilibrium [146,147] (Figure 6I,J). The amido group is trapped in the "*cis*" form in a heterocyclic, but not aromatic, intermediate in the evolution of pyrimidine synthesis. Cyclization of Asn or β-alaninamide side chains could produce tetrahydroorotate rings bound to the backbone of a nucleic acid predecessor (Figure 6K,L). The two epimers form double hydrogen bonds similar to the two amide conformations of the extended "*cis*" amide (Figure 6A,B). Hence, at this point cyclization did not distinguish between the two amide conformations. The strategy for ring closure employed in purine synthesis involves amino group formylation followed by ring closure with an amide NH_2 and loss of water. When this strategy is applied to cyclization of Asn or β-alaninamide, N–H bonds at both the 1 and 3 positions of the ring are not produced and double hydrogen bonding is not possible for a backbone bound pyrimidine ring (Figure 6M).

Biochemical pyrimidine synthesis generating an aromatic ring starts with Asp rather than Asn. Use of carbamoyl phosphate, a rather unstable reagent produced using ATP, allows cyclization with loss of water to leave an amide-like N–H at position 3. The additional C=O in the ring makes a dihydrouracil derivative with a more unsaturated and nearly planar ring that would be expected to exist preferentially as one epimer (Figure 6N). This step would have needed uptake and use of phosphate, perhaps as acetyl phosphate [34]. Further reactions to produce the aromatic pyrimidine ring require complicated enzymatic reactions. Flavin mononucleotide is required for conversion of –CH_2–CH(COOH)– to –CH=C(COOH)–, and a decarboxylation reaction involves stabilization of a carbene [148,149]. The last steps in the evolution of uracil synthesis may have occurred after steps were taken in purine synthesis.

A pyrimidine base that is specific for pairing with inosine or its predecessor required synthesis of an exocyclic amino group similar to that found in cytosine. Biochemically, this step follows formation of uracil. One possibility is that the unsaturated ring shown in Figure 6N was converted into a ring containing an exocyclic amino group (Figure 6O), before the complicated reactions evolved to produce the aromatic pyrimidine rings of uracil and cytosine. This dihydropyrimidine ring would pair with inosine or its predecessor.

Trapping of the "*trans*" amide form of β-alaninamide or Asn into a ring by crosslinking of the nitrogens with formaldehyde or formic acid is not possible. Purines are fusions of five and six-membered rings. The six-membered ring, which happens (confusingly) to be a pyrimidine ring, contains the hydrogen bonding function for base pairing. The "*trans*" carboxamide side chains (Figure 6C,D) are cyclized in the six-membered rings of purines (Figure 6G,H). However, purine biosynthesis assembles the five-membered imidazole ring first and builds the six-membered ring on the imidazole ring. The carboxamide group that provides the hydrogen bonding atoms of inosine is inserted into the imidazole ring. A carbonyl group in the imidazole ring, originating from Gly, plays a role in making the six-membered ring because prior to insertion of the carboxamide group it is converted into an amino group that is formylated for cyclization with the carboxamide NH_2 to produce inosine (Figure 6P,Q).

Figure 6. Evolution of purines and pyrimidines. Arrows show directions of possible hydrogen bonds. (**A–D**) Structures of the malamide side chain extended with β-alaninamide in "*cis*"- and "*trans*"-like conformations relative to the two CH2 groups. Blue arrows indicate rotations producing different conformations. Nitrogen and carbon atoms in the extended side chain are highlighted in red; (**E–H**) Structures of ribonucleic acid bases in RNA, (uracil, cytosine, guanine, and adenine) showing their relationships to the "*cis*"- and "*trans*"-like structures shown in A–D; (**I,J**) Epimers of tetrahydroorotate, a possible intermediate in pyrimidine evolution formed by reaction of Asn with formaldehyde; (**K,L**) Tetrahydroorotate epimers linked to the backbone of nucleic acid predecessors; (**M**) A pyrimidine derivative, formed by crosslinking of Asn using formylation, lacking the N–H bond at position 3 needed for double hydrogen bonding; (**N**) Dihydroorotate intermediate in uracil synthesis; (**O**) 4-Aminodihydroorotate, a proposed evolutionary predecessor of cytosine; (**P**) 4-carboxamido-5-aminoimidazole ribonucleotide intermediate in purine synthesis; (**Q**) Inosine, the first purine ring synthesized in the biochemical pathway to adenine and guanine; (**R,S**) Proposed intermediates in the evolution of the six-membered ring of purines lacking the imidazole ring. The carbonyl group of malonamide (R) is converted to an amino group (S) using the same reactions used in inosine biosynthesis; (**T**) Proposed predecessor of inosine. Six-membered pyrimidine ring is linked through a 6 amino group to the nucleic acid backbone. The amino and amido nitrogen atoms of S are linked via formylation; (**U**) Proposed predecessor of adenine. T is converted into an amino derivative similar to the conversion of inosine into adenine.

The sequence of events in purine biosynthesis—imidazole ring first, pyrimidine ring second—may mirror the sequence of events in the evolution of purine synthesis. If this were the case, the synthetic steps in imidazole ring synthesis evolved prior to the introduction of the carboxamide group, presumably for a function other than purine synthesis. Another possibility is that the imidazole ring was incorporated into the reaction sequence for purine synthesis after introduction of the carboxamide group allowed the imidazole ring to be a substrate for the reactions leading to the six-membered ring. Therefore, an alternative, shorter evolutionary pathway for synthesis of the six-membered ring could have preceded assimilation with the imidazole ring. Crosslinking of the trans-amide conformation into a six-membered ring would require similar introduction of an amino group in the side chain corresponding to the amino group of the imidazole ring. A malonamide side chain in place of Asn or β-alaninamide could have supplied the carbonyl group that was converted into an amino group for cyclization (Figure 6R,S). This evolutionary development would have required the ability to synthesize malonic acid, perhaps in the absence of biotin, by carboxylation of the enol form of an acetyl-thioester.

A substituted pyrimidine ring with double hydrogen bonding capability linked to the backbone via a six-amino group could then be formed via formylation, which introduces unsaturation into the ring and makes it aromatic (Figure 6T). The formylation method could not be used for cyclization of a predecessor of uracil because protonation of N1 was required for linking of the ring to the backbone leaving N3 unprotonated and unable to be involved in hydrogen bonding (Figure 6M), but it can be used for cyclization of ring shown in Figure 6T because N1 is not used for linking to the backbone. The evolutionary predecessor of inosine shown in Figure 6T probably existed as one primary epimer and

therefore double hydrogen bonding was similar to one of the two "*trans*" amide conformations, (compare Figure 6T with Figure 6C). It could have been a substrate for the reactions that convert inosine into adenine that could have introduced an exocyclic amino group in place of oxygen (Figure 6U), allowing double hydrogen bonding similar to the other "*trans*" amide conformation (compare Figure 6U with Figure 6D). Oligomers containing 2,4-disubstituted pyrimidines linked through a five-amino rather than a 6-amino group base pair weaker than biological bases, leading to the observation that "canonical nucleobases represent a functional optimum with respect to informational base pairing in aqueous solution" [150]. Subsequent inclusion of the imidazole ring would have improved base pairing. Cytosine and its proposed predecessor had the capacity to form a triple hydrogen bond, and guanine synthesis evolved from inosine to provide a triple hydrogen bonding partner, thereby increasing the strength of double stranded RNA interactions.

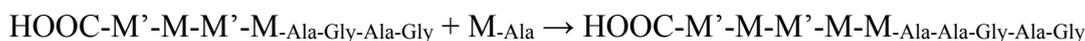## 19. Coenzyme A as a Relic of Early Nucleic Acid Evolution

Synthesis of CoA, one of the most ancient molecules in biology, uses a pyruvoyl-dependent α-decarboxylation reaction for synthesis of its β-alanine component from Asp [151]. The phosphopantetheine component of CoA is shown in Figure 5C. Remarkably, its structure has similarities to the extended side chain and an intermediate in RNA backbone evolution in which the 4-carboxylic acid of malic acid has been reduced to a hydroxymethyl group that allows formation of phosphodiester bonds (Figures 4B and 5D). The main difference is the presence of the two CH_3 groups in the phophopantetheine. Perhaps, these groups prevented pantetheine from being recognized by the ancient proteins involved in activation of monomers for polymerization. The structure of CoA seems to contain a relic of the early evolution of nucleic acids.

20. Coupling of β-Polyester Synthesis and Protein Synthesis

The mechanism for synthesis of PHBA suggests that activation of malic acid and malamide monomers for polymerization could initially have involved formation of β-thioesters. Polymerization would proceed with the terminal OH group attacking the thioester bond of the next monomer (Figure 3B,C). Because the OH group is the predecessor of the 3'-OH of ribose, the direction of synthesis is equivalent to the 5' to 3' direction in RNA polymerization. Aminoacyl thioesters form polypeptides in the absence of enzymes [27,28]. Could these two processes be combined to allow coupled synthesis of proteins and polyesters? As noted earlier, acyl groups move from thiol to hydroxyl groups [26] that would allow formation of aminoacyl esters from aminoacylthioesters. Unlike a simple ester bond such as that in ethylacetate, aminoacyl ester bonds have energies similar to thioester or phosphoanhydride bonds [152], and it is the energy in aminoacyl-tRNAs that drives peptide bond formation in ribosomal protein synthesis [153,154]. Consequently, aminoacyl group transfer from an aminoacyl thioester to the OH of malate or malamide thioesters would create aminoacylated monomers that could also be used for polymerization of β-polyesters. Aminoacylation of the OH group of malic acid or malamide would be the evolutionary predecessor of aminoacylation of the 3'(2')-OH of ribose by the aminoacyl-tRNA synthetases.

At a time when the coding system of ribosomal protein synthesis was beginning to be revealed, Rich [155] proposed that life started with coupling of nucleic acid and amino acid polymerization using

a singlet coding system that could have initially involved only two nucleotides, each of which was ester linked to a specific amino acid. In the context of polymerization of β-polyesters and proteins, this idea can be adopted for coupling of ester bond formation for extension of the β-polyester chain with peptide bond formation for extension of the protein chain. If two different monomers, with extended (=M') and unextended (=M) side chains preferentially aminoacylated with Gly and Ala, respectively, a simple singlet coding system would be created for coupled synthesis.

$$\text{HOOC-M'-M-M'-M}_{\text{-Ala-Gly-Ala-Gly}} + \text{M}_{\text{-Ala}} \rightarrow \text{HOOC-M'-M-M'-M-M}_{\text{-Ala-Ala-Gly-Ala-Gly}}$$

This would have been the first mechanism for coded protein synthesis. Peptide bond synthesis in this process would resemble that of diglycine directed by polyU from 2'(3')-glycyl esters of adenosine derivatives [156]. Figure 7 shows addition of an alanyl-malamide thioester monomer to a nascent β-polyester strand using another polyester strand as template for amide pairing. The polypeptide at the end of the nascent strand is transferred to alanine when the new ester bond is formed.

Although insertion of proteins into membranes appears to be an unlikely primitive process, insertion of segments synthesized from a small number of simple amino acids such as Gly and Ala would have been a simpler method for membrane biogenesis than synthesizing lipids. Gly-Ala polypeptides inserted into tholin were the first membrane components to be synthesized. Another, far less studied, membrane component, PHBA, forms pores in membranes [157] and also has a shorter synthetic pathway than lipids.
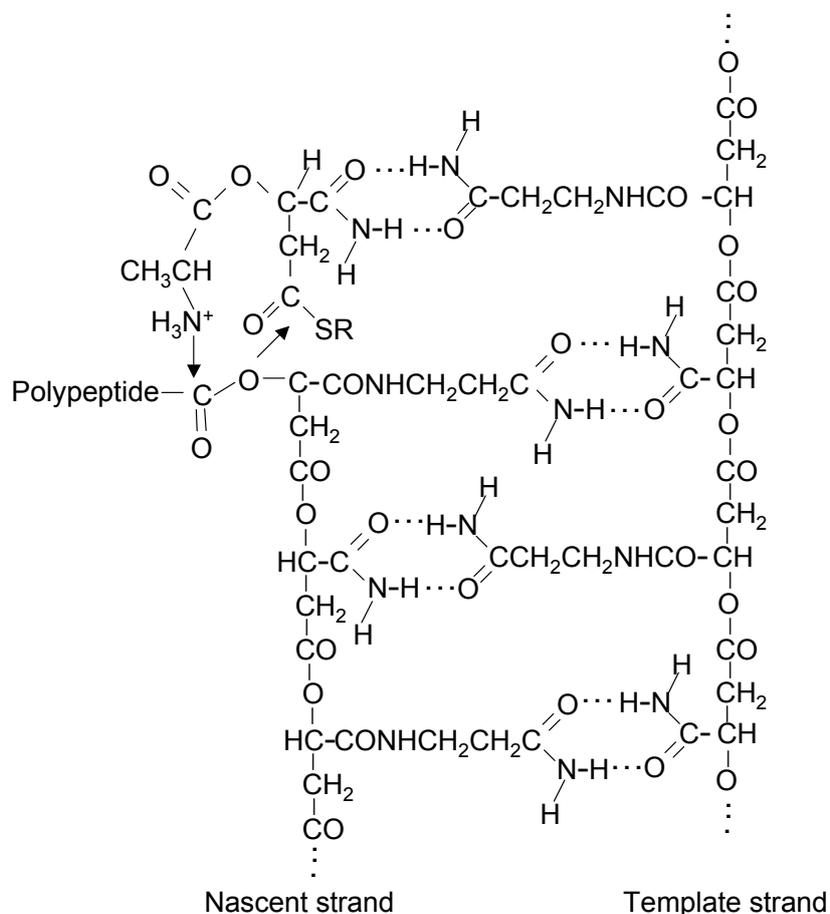
Figure 7. Coupled synthesis of poly(β-malamide) and polypeptide. Extension of a nascent poly(β-malamide) strand using a poly(β-malamide) strand as template and an alanyl-malamide

thioester monomer. The monomer amide pairs with an extended malamide side chain on the template strand. Formation of the ester bond is accompanied by transfer of the polypeptide at the end of the nascent strand to alanine producing a new peptide bond. The leaving group is a thiolate anion. Not shown are loss of a proton from the alanyl NH_3^+ group and gain of a proton by the thiolate anion to produce a thiol.

21. A Thiol at the N-Terminal End of Early Proteins

In the preceding analysis, the thiol group occupies central functions in formation of thioesters and disulfides, and as a ligand to FeS compounds. The thiol group is also a strong nucleophile. Hence, it may have been advantageous, and perhaps essential, to have a thiol group in the earliest protein sequences. However, only the simplest two amino acids, Gly and Ala, were coded in the first singlet system. Cys appears to have been a late addition to the triplet coding system [158,159]. Eck and Dayhoff [85] in a study of ferredoxin evolution suggested that a primitive sequence containing Ala, Asp, Ser, and Gly evolved by duplication. Later, Cys was added to the sequence allowing binding to FeS and then as more Cys were added binding to Fe_4S_4. "On the principle that evolution proceeds one step at a time we assume that the cell was already using iron sulfide as a catalyst probably attached to cysteine alone or to some peptide less suitable than protoferredoxin". What was this peptide? The question can be restated: How could a thiol have been incorporated into the singlet coding system? Met/fMet is the initiating amino acid in protein synthesis. Why has such an unusual amino acid assumed such a prominent function? Met is synthesized from homocysteine. Therefore, homocysteine may have been the initiating amino acid in the triplet coding system prior to incorporation of Cys into proteins, and may also have been the evolutionary successor to a thiol at the N-terminal end of proteins in the early singlet process. By being at the N-terminal end the simple initiating thiol did not necessarily have to be an amino acid. A simple acid such as thioglycolic acid or thiolactic acid may have been sufficient. A thiol at the N-terminal end of all proteins allowed them to bind to the FeS particle and be involved in catalyzing transfer of acyl groups and reduction reactions at the FeS surface. Later, assembly of FeS clusters enabled cells to become independent of FeS particles while retaining critical FeS-mediated functions.

The catalytic function of carboxylates of tholin was taken over by the C-terminal ends of the polypeptides inserted into the membrane material. Early Gly-Ala polypeptides could have brought together a thiol group at the N-terminal end and a carboxylic acid at the C-terminal end for acid/base catalysis. Gly-rich loops supplied an anion binding function [159].

22. Autocatalysis Revisited

As mentioned previously, autocatalysis with symmetry breaking in biology depends upon the use of proteins composed of L-amino acids to catalyze production of proteins composed of L-amino acids. It is now possible to propose how symmetry breaking in biochemistry began. Activation of malamide monomers by aminoacylation with simple amino acids such as Gly and L-Ala was catalyzed by simple proteins made from Gly and L-Ala. The amino acid binding site would preferentially bind L-Ala over D-Ala. When this activation process evolved to allow preferential binding of L-Ala to the malamide monomer with the unextended side chain and binding of Gly to the malamide monomer with the

extended side chain, a simple coding system became possible that allowed conservation of amino acid sequence and configuration, and symmetry breaking was achieved.

23. Replication of β-Polyesters

Because different malamide side chains would be capable of forming double hydrogen bonds, a simple replication process would involve using one strand of β-polyester to act as a template for synthesis of a complementary strand (Figure 7). Duplicated sequences, like gene duplications, would allow selection for improved or changed protein functions. However, in a singlet coding system mutations in base sequence lead directly to changes in amino acid sequence. There are no silent mutations. Multiple copies of genetic sequences would be needed to evolve new protein sequences but also to mitigate the deleterious effects of mutations that altered amino acid sequences.

24. Evolution of the Singlet Coding System into the Triplet Coding System

As the polyester backbone evolved into the phosphoribose backbone of RNA and the purine and pyrimidine bases evolved, the singlet coding system allowed four bases to code for four amino acids (Gly, Ala, Asp, Val). The triplet coding system of ribosomal protein synthesis evolved from the singlet system permitting RNA synthesis and protein synthesis to be uncoupled and multiple copies of protein to be synthesized using a single strand of RNA as template. If the four codons to be used in ribosomal protein synthesis were GNC (N = any base), the second base would determine the amino acid (U = Val, C = Ala, A = Asp, G = Gly) allowing continuity with the singlet coding system [5]. Evolution of the standard genetic code from an early code that included only these four amino acids is widely supported [5,8,159–163].

25. Overview of the Evolution of Nucleic Acid Synthesis, and the Origin and Evolution of the Genetic Code

Synthesis of the four bases in RNA did not evolve at the same time. As far as the aromatic bases are concerned, uracil biosynthesis precedes cytosine and inosine precedes both adenine and guanine. The first products are those in which the "amide NH_2", not the "amide C=O", is incorporated into the rings. Evolution of base pairing must take this into account. Moreover, some steps in biosynthesis of the bases appear to be more difficult than others. Uracil synthesis has some difficult steps, as do the syntheses of cytosine from uracil and guanine from inosine. This hypothesis adopts the position that the four bases were not obtained from the environment but were the products of evolutionary processes within cells. Prior to synthesis of the aromatic bases, intermediates in their syntheses were used in pairing. These intermediates would also not have become available at the same time.

By combining the analysis in this study with those given previously [5,153], an overview of nucleic acid base evolution, amide/base pairing, and the origin and evolution of the genetic code can be presented, starting with the singlet coding system and ending with the triplet coding system:

(1) carboxamide side chain (Figure 5A), coding for Ala, pairs with extended carboxamide side chain (Figure 5B), coding for Gly.

(2) tetrahydropyrimidine derivative (Figure 6K or 6L), coding for Ala, pairs with extended carboxamide side chain (Figure 5B), coding for Gly,

(3) tetrahydropyrimidine derivative (Figure 6L), coding for Ala, pairs with inosine predecessor (Figure 6T), coding for Gly,

(4) tetrahydropyrimidine (Figure 6L), coding for Ala, pairs with inosine predecessor (Figure 6T); tetrahydropyrimidine (Figure 6K), coding for Ala, pairs with adenine predecessor (Figure 6U), coding for Asp.

(5) uracil predecessor (Figure 6N) coding for Val, pairs with adenine predecessor (Figure 6U), coding for Asp; cytosine predecessor (Figure 6O), coding for Ala, pairs with inosine predecessor (Figure 6T), coding for Gly,

(6) uracil (Figure 6E), coding for Val, pairs with adenine (Figure 6H), coding for Asp; cytosine (Figure 6F), coding for Ala, pairs with inosine, coding for Gly (Figure 6Q)

(7) uracil (Figure 6E), coding for Val, pairs with adenine (Figure 6H), coding for Asp; cytosine (Figure 6F), coding for Ala, and pairs with guanine (Figure 6G), coding for Gly.

(8) Codons: G<u>U</u>C codes for Val, G<u>C</u>C codes for Ala, G<u>A</u>C codes for Asp, G<u>G</u>C codes for Gly. The second base in the codon, underlined, is the same as the base in the singlet coding system.

(9) Expansion of genetic code to include 20 amino acids.

A biochemical example of a carboxamide double hydrogen bonding with uracil is found in the structure of photolase where uracil binds to the carboxamide of Asn (Figure 4 in [164]).

26. Discussion

26.1. An Evolutionary Pathway from Primitive Cells to Biology

The events that enabled a simple cellular system to assemble on the primitive Earth and evolve into the complex system of biology are difficult to understand. The four major activities of modern cells—membrane synthesis, energy transduction, catalysis, and replication of genetic polymers—have to be considered together. The "metabolism first" approach to the origin of life leaves out the genetic component and the "genetics first" approach leaves out the metabolism that was needed to support synthesis of genetic polymers. In this paper and others in this series [5,8,159], I have used chemistry and biochemistry to propose an evolutionary pathway from the first cells, using molecules found at the center of metabolism for synthesis of simple linear polyesters and polypeptides, to the complex metabolism of biology with nucleic acid synthesis and ribosomal protein synthesis.

26.2. Comparisons with Other Proposals on the Origins of Cells

The origin of cells described here and earlier [8] bears similarities and differences with other proposals for primitive cells. It is similar to proposals that the first cells had membranes composed of organic molecules [165] but, in agreement with Folsome [66], uses tholin rather than single chain lipids as the organic material. The proposal resembles the "lipid world" hypothesis [166] in that the membrane material, tholin, played an important early function by binding to FeS through thiol groups and supplying carboxylic acid groups for acid/base catalysis. It is similar to proposals that FeS surfaces provided reactive sites for reductive synthesis [21,167] but differs in that FeS (a) was not converted into FeS_2;

(b) was enclosed in thiol-containing tholin that facilitated electron transport and diffusion of small non-polar and polar molecules while only allowing slow diffusion of ions; and (c) catalyzed formation of thioesters and other molecules using energy from oxidation reactions such as oxidation sulfite to sulphate. The idea of encapsulating a mineral, including FeS, in an organic enclosure has precedents in other studies [168–170].

26.3. Why Is the Simple Method of Thioester Synthesis Not Used in Biology?

If FeS-mediated thioester synthesis occurred by a simple mechanism such as that shown in Figure 1, why is there no evolutionary footprint of this mechanism in modern cells? The proposed mechanism is non-specific. While such a mechanism may have been useful initially it may have become superseded by protein-catalyzed reactions that provided higher levels of specificity. An FeS catalyzed reaction may have supplied a mechanism for thioester synthesis but a side reaction reducing efficiency may have been reduction of a disulfide to thiols similar to an electrolytic reduction of a disulfide [171]. This level of efficiency may have been inadequate as evolution proceeded. The model presented involves a simple thiol such as CH_3SH and other simple thiols, thioglycolic or thiolactic acids at the N-terminal end of proteins. The different functions of thiols in redox reactions, metal binding, and carrying acyl groups and NO, have largely become specialized in the course of evolution. This may also have contributed to the loss of the primitive mechanism for thioester synthesis.

26.4. Coded Proteins Enabled Evolution of Nucleic Acid and Ribosomal Protein Synthesis

A widely held view is that prior to evolution of ribosomal protein synthesis catalysis of chemical reactions was performed by RNA (reviewed in [133]). Catalysis by RNA is strongly dependent upon the properties of RNA including the presence of the 2'-OH group, triple hydrogen bonding by GC base pairing, and A-minor interactions. RNA is an extremely complex polymer. If simpler evolutionary predecessors of RNA did not have a ribose-phosphate backbone with its 2'-OH or GC base pairing, the catalytic properties of those polymers would have been different from those of RNA. Yet, it would have been these polymers that would have been responsible for catalyzing RNA synthesis [172]. A more plausible explanation is that coded protein synthesis was accomplished early in the origin of life beginning with Gly and Ala and then Gly, Ala, Val and Asp in a singlet coding system. As a result, proteins would have been available to catalyze reactions needed for evolution of RNA synthesis and ribosomal protein synthesis, and to facilitate energy transduction, membrane biosynthesis, and membrane transport.

26.5. The Need for Simplicity of Metabolism, Catalysis, and Replication in the Primitive Cell

In critiques of "metabolism first" ideas about the origin of life, Ross [173] and Orgel [174] note two particular difficulties with metabolic cycles as precursors to genetics. One is that chemical reactions generally have low specificity and increased side reactions compared to enzyme catalyzed reactions. Reactions of a particular type, conversion of a carboxylic acid into a carboxamide for example, might apply to all carboxylic acids. Lactate formation from pyruvate would be a side reaction that would impede carboxylic acid cycles in metabolism first scenarios. The other difficulty is that most syntheses

under presumed prebiotic conditions yield a complex mixture of products. Side reactions reduce the efficiency of chemical reactions. Orgel [174] suggests that, "Simplification of product mixtures through the self-organization of organic reaction sequences, whether cyclic or not, would help enormously, as would the discovery of very simple replicating polymers" and "… a theory in which metal sulfide-catalyzed reactions provided some or all of the organic molecules needed for the formation of a primitive genetic system would have many attractive features". In addition, it has been concluded that catalysts in the origin of life should allow accumulation of large concentrations of a few components rather than low concentrations of many components [175]. In the proposed primitive cells, the number of substrates is limited by restricting diffusion across tholin membranes of ions and large polar molecules, and the types of chemical reactions catalyzed by FeS is limited. The end products of electron flow into the cell are polyesters and polypeptides. Poly(β-D-malic acid) is a simple predecessor of the nucleic acids. One functional group, a thiol, acts in redox reactions, carries activated carboxylic acids, and binds to Fe. Therefore, the needs for a simple primitive metabolic system and a simple replicating polymer are met by this hypothesis. If nucleic acid synthesis began with malic acid and its amide derivatives, the "genetics first" and "metabolism first" approaches to the origin of life are reconciled because malic acid features prominently in carboxylic acid cycles. However, the reverse citric acid cycle was probably too complex to have been established at the beginning.

26.6. Reversible Reactions in Primitive Cells

Some enzyme-catalyzed reactions in cells operate under near equilibrium conditions and others are essentially irreversible. Interconversion of lactic acid and pyruvic acid catalyzed by lactate dehydrogenase is a near equilibrium reaction and can operate in each direction dependent upon the conditions, especially the concentrations of lactic and pyruvic acids. Was the ability of reactions to be reversible also a feature of primitive cells where the catalysts for reactions were FeS and the inner surface of the tholin membrane? The study by Wang *et al.* [109] shows that some reduction reactions may have been reversible. An FeS/H_2S redox system reduces glyoxylic, pyruvic and oxaloacetic acids into the corresponding hydroxy acids, glycolic, lactic and malic acids. If this were irreversible, pyruvate would be "trapped" in lactic acid and could not be available for synthesis of other molecules including malic acid. But when sulfur is added to the mix with FeS and H_2S, the hydroxy acids are oxidized to glyoxylic, pyruvic and oxaloactic acids, showing that the reactions are reversible. This would allow some synthetic pathways to be favored over others. In this proposal, the two most favored pathways are those leading to β-polyester and α-polypeptide synthesis.

26.7. The Four Bases in Nucleic Acids Are Derived from the Four Orientations of Carboxamide Side Chains

The four bases used for pairing in nucleic acids are not the only ones that could have been used [176]. Why use a genetic alphabet with just four? Szathmáry [177] suggests that the bases are a relic of the RNA world. The explanation that comes out of the proposed evolution of the pyrimidine and purine bases is that the four bases evolved from the two possible orientations of the carboxamide group in the *cis*- and *trans*-like carboxamide evolutionary intermediates (Figure 6A–D).

26.8. Did "Chemical Evolution" Happen before Biological Evolution?

There is a long history to the idea that biological evolution with a genetic selection or Darwinian capability was preceded by a period of "chemical evolution" [178,179]. In the course of chemical evolution, reactions occurred in aqueous solution (the prebiotic soup) or through wet-dry cycles, driven by the energy intrinsic to certain chemicals, such as cyanides and aldehydes, by photolysis, or by surface catalysis, to produce pathways leading to increasingly complex metabolic cycles, polypeptides and oligonucleotides. It is then assumed that these chemicals became encapsulated in lipid vesicles initiating biological evolution [180]. Leaky vesicles enabled diffusion of nutrients and replication [181]. The hypothesis presented here supports the view that there was not a period of chemical evolution. Life started with energized cellular structures and the molecules necessary for biological evolution were generated within these structures from simple molecules that diffused into the cells. Evolution began with the formation of the first cells.

Conflicts of Interest

The author declares no conflict of interest.

References

1. Neveu, M.; Kim, H.-J.; Benner, S.A. The "strong" RNA world hypothesis: Fifty years old. *Astrobiology* **2013**, *13*, 391–403.
2. Joyce, G.F. The antiquity of RNA-based evolution. *Nature* **2002**, *418*, 214–221.
3. Kurland, C.G. The RNA dreamtime. *Bioessays* **2010**, *32*, 866–871.
4. Bernhardt, H.S. The RNA world hypothesis: The worst theory of the early evolution of life (except for all the others). *Biol. Direct* **2012**, *7*, doi:10.1186/1745-6150-7-23.
5. Francis, B.R. An alternative to the RNA world hypothesis. *Trends Evol. Biol.* **2011**, *3*, doi:10.4081/eb.2011.e2.
6. Li, L.; Francklyn, C.; Carter, C.W. Aminoacylating urzymes challenge the RNA world hypothesis. *J. Biol. Chem.* **2013**, *288*, 26856–26863.
7. Caetano-Anollés, G.; Seufferheld, M.J. The coevolutionary roots of biochemistry and cellular organization challenge the RNA world paradigm. *J. Mol. Microbiol. Biotechnol.* **2013**, *23*, 152–177.
8. Francis, B.R. A hypothesis that ribosomal protein synthesis evolved from coupled protein and nucleic acid synthesis. *Chemtracts Biochem. Mol. Biol.* **2000**, *13*, 153–191.
9. Lipmann, F. Attempts to map a process evolution of peptide biosynthesis. *Science* **1971**, *173*, 875–885.
10. Wächtershäuser, G. Evolution of the first metabolic cycles. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 200–204.
11. Morowitz, H.J.; Kostelnik, J.D.; Yang, J.; Cody, G.D. The origin of intermediary metabolism. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 7704–7798.
12. Smith, E.; Morowitz, H.J. Universality in intermediary metabolism. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 13168–13173.

13. Russell, M.J.; Martin, W. The rocky roots of the acetyl-CoA pathway. *Trends Biochem. Sci.* **2004**, *29*, 358–363.

14. Hazen, R.; Sverjensky, D.A. Mineral surfaces, geochemical complexities, and the origins of life. *Cold Spring Harb. Perspect. Biol.* **2010**, *2*, doi:10.1101/cshperspect.a002162.

15. Trefil, J.; Morowitz, H.; Smith, E. The Origin of Life. *Am. Sci.* **2009**, *97*, doi:10.1511/2009.78.206.

16. Nitschke, W.; Russell, M.J. Beating the acetyl coenzyme A-pathway to the origin of life. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2013**, *368*, doi:10.1098/rstb.2012.0258.

17. Sousa, F.L.; Thiergart, T.; Landan, G.; Nelson-Sathi, S.; Pereira, I.A.; Allen, J.F.; Lane, N.; Martin, W.F. Early bioenergetic evolution. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2013**, *368*, doi:10.1098/rstb.2013.0088.

18. Weber, A.L. Nonenzymatic formation of "energy rich" lactoyl and glyceroyl thioesters from glyceraldehyde and a thiol. *J. Mol. Evol.* **1984**, *20*, 157–166.

19. Weber, A.L. Prebiotic amino acid thioester synthesis: Thiol-dependent amino acid synthesis from formose substrates (formaldehyde and glycoaldehyde) and ammonia. *Orig. Life Evol. Biosph.* **1998**, *28*, 259–270.

20. De Duve, C. *Blueprint for a Cell: The Nature and Origin of Life*; Neil Patterson Publishers: Burlington, NC, USA, 1991.

21. Wächtershäuser, G. Groundworks for an evolutionary biology: The iron-sulphur world. *Prog. Biophys. Mol. Biol.* **1992**, *58*, 85–201.

22. Heinen, W.; Lauwers, A.-M. Organic sulfur compounds resulting from the interaction of iron sulfide, hydrogen sulfide and carbon dioxide in an anaerobic environment. *Orig. Life Evol. Biosph.* **1996**, *26*, 131–150.

23. Baltscheffsky, H.; Schultz, A.; Baltscheffsky, M. Energy for the origin of life. In *Exobiology: Matter, Energy, and Information in the Origin and Evolution of Life in the Universe*; Chela-Flores, J., Raulin, F., Eds.; Kluwer Academic Publishers: South Holland, The Netherlands, 1998; pp. 95–102.

24. Deamer, D.W.; Weber, A.L. Bioenergetics and life's origins. *Cold Spring Harb. Perspect. Biol.* **2010**, *2*, doi:10.1101/cshperspect.a004929.

25. Koonin, E.V. The origins of cellular life. *Antonie Van Leeuwenhoek* **2014**, *106*, 27–41.

26. Janssen, M.J. Thiolo, thiono, and dithio acids and esters. In *The Chemistry of Carboxylic Acids and Esters*; Patai, S., Ed.; Interscience: London, UK, 1969; pp. 705–764.

27. Wieland, T.; Schäfer, W. Synthese von oligopeptiden unter zellmoglichen bedingungen. *Angew. Chem.* **1951**, *63*, 146–147.

28. Wieland, T.; Pfleiderer, G. Activation of amino acids. *Adv. Enzymol. Relat. Subj. Biochem.* **1957**, *19*, 235–266.

29. Maurel, M.C.; Orgel, L.E. Oligomerization of α-thioglutamic acid. *Orig. Life Evol. Biosph.* **2000**, *30*, 423–430.

30. Zepik, H.H.; Rajamani, S.; Maurel, M.C.; Deamer, D. Oligomerization of thioglutamic acid: Encapsulated reactions and lipid catalysis. *Orig. Life Evol. Biosph.* **2007**, *37*, 495–505.

31. Weber, A.L. Oligoglyceric acid synthesis by autocondensation of glyceroyl thioester. *J. Mol. Evol.* **1987**, *25*, 191–196.

32. Dawes, E.A.; Senior, P.J. The role and regulation of energy reserve polymers in micro-organisms. *Adv. Microbiol. Physiol.* **1973**, *10*, 135–266.

33. Castresana, J.; Saraste, M. Evolution of energetic metabolism: The respiration-early hypothesis. *Trends Biochem. Sci.* **1995**, *20*, 443–448.

34. Martin, W.; Russell, M.J. On the origin of biochemistry at an alkaline hydrothermal vent. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2007**, *362*, 1887–1925.

35. Liu, Y.; Beer, L.L.; Whitman, W.B. Sulfur metabolism in archaea reveals novel processes. *Environ. Microbiol.* **2012**, *14*, 2632–2644.

36. Khare, B.N.; Thompson, W.R.; Chyba, C.F.; Arakawa, E.T.; Sagan, C. Organic solids produced from simple C/H/O/N ices by charged particles: Applications to the outer solar system. *Adv. Space Res.* **1989**, *9*, 41–53.

37. Cody, G.D.; Boctor, N.Z.; Filley, T.R.; Hazen, R.M.; Scott, J.H.; Yoder, H.S., Jr. Primordial carbonylated iron-sulfur compounds and the synthesis of pyruvate. *Science* **2000**, *289*, 1337–1340.

38. Yin, S.; Wang, Z.; Bernstein, E.R. Formaldehyde and methanol formation from reaction of carbon monoxide and hydrogen on neutral Fe_2S_2 clusters in the gas phase. *Phys. Chem. Chem. Phys.* **2013**, *15*, 4699–4706.

39. Fuchs, G. Alternative pathways of carbon dioxide fixation: Insights into the early evolution of life? *Ann. Rev. Microbiol.* **2011**, *65*, 631–658.

40. O'Connell, C.; Hommeltoft, S.I.; Eisenberg, R. Electrochemical approaches to the reduction of carbon dioxide. In *Carbon Dioxide as a Source of Carbon*; Aresta, M., Forti, G., Eds.; D. Reidel Publishing Co.: Dordrecht, The Netherlands, 1987; pp. 33–54.

41. Tezuka, M.; Yajima, T.; Tsuchiya, A.; Matsumoto, Y.; Uchida, Y.; Hidai, M. Electroreduction of carbon dioxide catalyzed by iron-sulfur cluster compounds $[Fe_4S_4(SR)_4]^{2-}$. *J. Am. Chem. Soc.* **1982**, *104*, 6834–6836.

42. Lucas, B.H.; Ritcey, G.M. Reduction of ferric iron by SO_2 with heat or SO_2 with activated carbon. In *Department of Energy, Mines and Resources, Technical Bulletin TB 107*; Queen's Printer and Controller of Stationary: Ottawa, ON, Canada, 1969.

43. Cremlyn, R.J. *An Introduction to Organosulfur Chemistry*; John Wiley & Sons: Chichester, UK, 1996.

44. Ducluzeau, A.L.; van Lis, R.; Duval, S.; Schoepp-Cothenet, B.; Russell, M.J.; Nitschke, W. Was nitric oxide the first deep electron sink? *Trends Biochem. Sci.* **2009**, *34*, 9–15.

45. Mancinelli, R.L.; McKay, C.P. The evolution of nitrogen cycling. *Orig. Life Evol. Biosph.* **1988**, *18*, 311–325.

46. Summers, D.P.; Chang, S. Prebiotic ammonia from reduction of nitrite by iron(II) on the early Earth. *Nature* **1993**, *365*, 630–633.

47. Summers, D.P. Ammonia formation by the reduction of nitrite/nitrate by FeS: Ammonia formation under acidic conditions. *Orig. Life Evol. Biosph.* **2005**, *35*, 299–312.

48. Summers, D.P.; Basa, R.C.; Khare, B.; Rodoni, D. Abiotic nitrogen fixation on terrestrial planets: Reduction of NO to ammonia by FeS. *Astrobiology* **2012**, *12*, 107–114.

49. Saville, B. A scheme for colorimetric determination of microgram amounts of thiol. *Analyst* **1958**, *83*, 670–672.

50. Broniowska, K.A.; Diers, A.R.; Hogg, N. S-nitrosoglutathione. *Biochim. Biophys. Acta.* **2013**, *1830*, 3173–3181.

51. Chyba, C.E.; Sagan, C. Endogenous production, exogenous delivery, and impact-shock synthesis of organic molecules: An inventory for the origin of life. *Nature* **1992**, *355*, 125–132.

52. Deamer, D.W. Origins of life: How leaky were primitive cells. *Nature* **2008**, *454*, 37–38.

53. Mansy, S.S. Model protocells from single-chain lipids. *Int. J. Mol. Sci.* **2009**, *10*, 835–843.

54. Chen, I.A.; Walde, P. From self-assembled vesicles to protocells. *Cold Spring Harb. Perspect. Biol.* **2010**, *2*, doi:10.1101/cshperspect.a002170.

55. Pizzarello, S.; Shock, E. The organic composition of carbonaceous meteorites: The evolutionary story ahead of biochemistry. *Cold Spring Harb. Perspect. Biol.* **2010**, *2*, doi:10.1101/cshperspect.a002105.

56. Allamandola, L.J.; Sandford, S.A.; Wpoenka, B. Interstellar polycyclic aromatic hydrocarbons and carbon in interplanetary dust particles and meteorites. *Science* **1987**, *237*, 56–59.

57. Parker, E.T.; Zhou, M.; Burton, A.S.; Glavin, D.P.; Dworkin, J.P.; Krishnamurthy, R.; Fernandez, F.M.; Bada, J.L. A plausible simultaneous synthesis of amino acids and simple peptides on the primordial Earth. *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 8132–8136.

58. Mulkidjanian, A.Y.; Bychkov, A.Y.; Dibrova, D.V.; Galperin, M.Y.; Koonin, E.V. Origin of first cells at terrestrial, anoxic geothermal fields. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E821–E830.

59. Mulkidjanian, A.Y.; Bychkov, A.Y.; Dibrova, D.V.; Galperin, M.Y.; Koonin, E.V. Open questions on the origin of life at anoxic geothermal fields. *Orig. Life Evol. Biosph.* **2012**, *42*, 507–516.

60. Parker, E.T.; Cleaves, H.J.; Dworkin, J.P.; Glavin, D.P.; Callahan, M.; Aubrey, A.; Lazcano, A.; Bada, J.L. Primordial synthesis of amines and amino acids in a 1958 Miller H_2S-rich spark discharge experiment. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 5526–5531.

61. Sagan, C.; Khare, B.N. Tholins: Organic chemistry of interstellar grains and gas. *Nature* **1979**, *277*, 102–107.

62. McCord, T.B.; Carlson, R.W.; Smythe, W.D.; Hansen, G.B.; Clark, R.N.; Hibbitts, C.A.; Fanale, F.P.; Granahan, J.C.; Segura, M.; Matson, D.L.; *et al.* Organics and other molecules in the surfaces of Callisto and Ganymede. *Science* **1997**, *278*, 271–275.

63. Khare, B.N.; Bakes, E.L.; Cruikshank, D.; McKay, C.P. Solid organic matter in the atmosphere and on the surface of outer Solar System bodies. *Adv. Space Res.* **2001**, *27*, 299–307.

64. Waite, J.H., Jr.; Young, D.T.; Cravens, T.E.; Coates, A.J.; Crary, F.J.; Magee, B.; Westlake, J. The process of tholin formation in Titan's upper atmosphere. *Science* **2007**, *316*, 870–875.

65. Gu, X.; Kim, Y.S.; Kaiser, R.I.; Mebel, A.M.; Liang, M.C.; Yung, Y.L. Chemical dynamics of triacetylene formation and implications to the synthesis of polyynes in Titan's atmosphere. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 16078–16083.

66. Folsom, C.E. *The Origin of Life*; W.H. Freeman & Co.: San Francisco, CA, USA, 1979.

67. McDonald, G.D.; Khare, B.N.; Thompson, W.R.; Sagan, C. $CH_4/NH_3/H_2O$ spark tholin: Chemical analysis and interaction with Jovian aqueous clouds. *Icarus* **1991**, *94*, 354–367.

68. Khare, B.N.; Sagan, C. Synthesis of cystine in simulated primitive conditions. *Nature* **1971**, *232*, 577–579.

69. Van Trump, J.E.; Miller, S.L. Prebiotic synthesis of methionine. *Science* **1972**, *178*, 859–860.

70. Raulin, F.; Toupance, G. Effect of H_2S on the formation of organic compounds from C, N, H, S model atmospheres submitted to a glow discharge. *Orig. Life* **1975**, *6*, 507–512.

71. Khare, B.N.; Sagan, C.; Bandurski, E.L.; Nagy, B. Ultraviolet-photoproduced organic solids synthesized under simulated jovian conditions: Molecular analysis. *Science* **1978**, *199*, 1199–1201.

72. Minard, R.D.; Hatcher, P.G.; Gourley, R.C.; Matthews, C.N. Structural investigations of hydrogen cyanide polymers: New insights using TMAH thermochemolysis/GC-MS. *Orig. Life Evol. Biosph.* **1998**, *28*, 461–473.

73. He, C.; Lin, G.; Upton, K.T.; Imanaka, H.; Smith, M.A. Structural investigation of Titan tholins by solution-state ^1H, ^{13}C, and ^{15}N NMR: One-dimensional and decoupling experiments. *J. Phys. Chem. A* **2012**, *116*, 4760–4767.

74. Martin, W.F.; Sousa, F.L.; Lane, N. Evolution. Energy at life's origin. *Science* **2014**, *344*, 1092–1093.

75. Bernstein, M.P.; Sandford, S.A.; Allamandola, L.J.; Gillette, J.S.; Clemett, S.J.; Zare, R.N. UV irradiation of polycyclic aromatic hydrocarbons in ices: Production of alcohols, quinones, and ethers. *Science* **1999**, *283*, 1135–1138.

76. Hayashi, T.; Stuchebrukhov, A.A. Electron tunneling in respiratory complex I. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 19157–19162.

77. Schenning, A.P.; Meijer, F.W. Supramolecular electronics; nanowires from self-assembled π-conjugated systems. *Chem. Commun. (Camb.)* **2005**, doi:10.1039/B501804H.

78. Gray, H.B.; Winkler, J.R. Electron transfer in proteins. *Annu. Rev. Biochem.* **1996**, *65*, 537–561.

79. Seyedsayamdost, M.R.; Yee, C.S.; Reece, S.Y.; Nocera, D.G.; Stubbe, J. pH Rate profiles of F_nY_{356}-R2s ($n = 2, 3, 4$) in *Escherichia coli* ribonucleotide reductase: Evidence that Y356 is a redox-active amino acid along the radical propagation pathway. *J. Am. Chem. Soc.* **2006**, *128*, 1562–1568.

80. Koch, A.L.; Schmidt, T.M. The first cellular bioenergetic process: Primitive generation of a proton-motive force. *J. Mol. Evol.* **1991**, *33*, 297–304.

81. Wächtershäuser, G. Before enzymes and templates: Theory of surface metabolism. *Microbiol. Rev.* **1988**, *52*, 452–484.

82. Huber, C.; Wächtershäuser, G. Activated acetic acid by carbon fixation on (Fe,Ni)S under primordial conditions. *Science* **1997**, *276*, 245–247.

83. Ragsdale, S.W.; Pierce, E. Acetogenesis and the Wood-Ljungdahl pathway of CO_2 fixation. *Biochim. Biophys. Acta* **2008**, *1784*, 1873–1898.

84. Johnson, D.C.; Dean, D.R.; Smith, A.D.; Johnson, M.K. Structure, function, and formation of biological iron sulfur clusters. *Annu. Rev. Biochem.* **2005**, *74*, 247–281.

85. Eck, R.V.; Dayhoff, M.O. Evolution of the structure of ferredoxin based on living relics of primitive amino acid sequences. *Science* **1966**, *152*, 363–366.

86. Davis, B.K. Evolution before the origin of species. *Prog. Biophys. Mol. Biol.* **2002**, *79*, 77–133.

87. Evans, M.C.; Buchanan, B.B. Photoreduction of ferredoxin and its use in carbon dioxide fixation by a subcellular system from a photosynthetic bacterium. *Proc. Natl. Acad. Sci. USA* **1965**, *53*, 1420–1425.

88. Reed, G.H.; Ragsdale, S.W.; Mansoorabadi, S.O. Radical reactions of thiamin pyrophosphate in 2-oxoacid oxidoreductases. *Biochim. Biophys. Acta* **2012**, *1824*, 1291–1298.

89. Flint, D.H.; Allen, R.M. Iron-sulfur proteins with nonredox functions *Chem. Rev.* **1996**, *96*, 2315–2334.

90. Van Vugt-Lussenburg, B.M.; van der Weel, L.; Hagen, W.R.; Hagedoorn, P.L. Biochemical similarities and differences between the catalytic [4Fe-4S] cluster containing fumarases FumA and FumB from *Escherichia coli. PLoS One* **2013**, *8*, e55549.

91. Holm, R.H. Synthetic approaches to the active sites of iron-sulfur proteins. *Acc. Chem. Res.* **1977**, *10*, 427–434.

92. Holm, R.H. Electron transfer: Iron-sulfur clusters. In *Comprehensive Coordination Chemistry II*; McCleverty, J.A., Meyer, T.J., Eds.; Elsevier: New York, NY, USA, 2003; Volume 8, pp. 61–90.

93. Bonomi, F.; Werth, M.T.; Kurtz, D.M. Assembly of $Fe_nS_n(SR)^{2-}$ ($n = 2, 4$) in aqueous media from iron salts, thiols, and sulfur, sulfide, or thiosulfate plus rhodonase. *Inorg. Chem.* **1985**, *24*, 4331–4335.

94. Lo, W.; Scott, T.A.; Zhang, P.; Ling, C.C.; Holm, R.H. Stabilities of cubane type $[Fe_4S_4(SR)_4]^{2-}$ clusters in partially aqueous media. *J. Inorg. Biochem.* **2011**, *105*, 497–508.

95. Job, R.C.; Bruice, T.C. Iron-sulfur Clusters II: Kinetics of ligand exchange studied on a water-soluble $Fe_4S_4(SR)_4^{n-}$ cluster. *Proc. Natl. Acad. Sci. USA* **1975**, *72*, 2478–2482.

96. Dai, S.; Friemann, R.; Glauser, D.A.; Bourquin, F.; Manieri, W.; Schürmann, P.; Eklund, H. Structural snapshots along the reaction pathway of ferredoxin-thioredoxin reductase. *Nature* **2007**, *448*, 92–96.

97. Welte, C.; Deppenmeier, U. Bioenergetics and anaerobic respiratory chains of aceticlastic methanogens. *Biochim. Biophys. Acta* **2014**, *1837*, 1130–1147.

98. Duin, E.C.; Madadi-Kahkesh, S.; Hedderich, R.; Clay, M.D.; Johnson, M.K. Heterodisulfide reductase from *Methanothermobacter marburgensis* contains an active-site [4Fe-4S] cluster that is directly involved in mediating heterodisulfide reduction. *FEBS Lett.* **2002**, *512*, 263–268.

99. Shokes, J.E.; Duin, E.C.; Bauer, C.; Jaun, B.; Hedderich, R.; Koch, J.; Scott, R.A. Direct interaction of coenzyme M with the active-site Fe-S cluster of heterodisulfide reductase. *FEBS Lett.* **2005**, *579*, 1741–1744.

100. Sakuraba, H.; Oshima, T. Novel energy metabolism in anaerobic hyperthermophilic archaea: A modified Embden-Meyerhof pathway. *J. Biosci. Bioeng.* **2002**, *93*, 441–448.

101. Reher, M.; Gebhard, S.; Schonheit, P. Glyceraldehyde-3-phosphate ferredoxin oxidoreductase (GAPOR) and nonphosphorylating glyceraldehyde-3-phosphate dehydrogenase (GAPN), key enzymes of the respective modified Embden-Meyerhof pathways in the hyperthermophilic crenarchaeota *Pyrobaculum aerophilum* and *Aeropyrum pernix. FEMS Microbiol. Lett.* **2007**, *273*, 196–205.

102. Leman, L.; Orgel, L.; Ghadiri, M.R. Carbonyl-sulfide-mediated prebiotic formation of peptides. *Science* **2004**, *306*, 283–286.

103. Rode, B.M. Peptides and the origin of life. *Peptides* **1999**, *20*, 773–786.

104. Mulkidjanian, A.Y.; Galperin, M.Y.; Koonin, E.V. Co-evolution of primordial membranes and membrane proteins. *Trends Biochem. Sci.* **2009**, *34*, 206–215.

105. Deamer, D.W. Proton permeation of lipid bilayers. *J. Bioenerg. Biomembr.* **1987**, *19*, 457–479.

106. Hofvander, P.; Doan, T.T.; Hamberg, M. A prokaryotic acyl-CoA reductase performing reduction of fatty acyl-CoA to fatty alcohol. *FEBS Lett.* **2011**, *585*, 3538–3543.

107. De Duve, C. Clues from present-day biology: The thioester world. In *The Molecular Origins of Life*; Brach, A., Ed.; Cambridge University Press: Cambridge, UK, 1998; pp. 219–236.

108. Nivikov, Y.; Copley, S.D. Reactivity landscape of pyruvate under simulated hydrothermal vent conditions. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 13283–13288.

109. Wang, W.; Yang, B.; Qu, Y.; Liu, X.; Su, W. FeS/S/FeS$_2$ redox system and its oxidoreductase-like chemistry in the iron-sulfur world. *Astrobiology* **2011**, *11*, 471–476.

110. Fukuyama, T.; Lin, S.C.; Li, L. Facile reduction of ethyl thiol esters to aldehydes: Application to a total synthesis of a (+)-neothromycin A methyl ester. *J. Am. Chem. Soc.* **1990**, *112*, 7050–7051.

111. Ragsdale, S.W. Pyruvate ferredoxin oxidoreductase and its radical intermediate. *Chem. Rev.* **2003**, *103*, 2333–2346.

112. Chabrière, E.; Vernède, X.; Guigliarelli, B.; Charon, M.H.; Hatchikian, E.C.; Fontecilla-Camps, J.C. Crystal structure of the free radical intermediate of pyruvate:ferredoxin oxidoreductase. *Science* **2001**, *294*, 2559–2563.

113. Nakajima, T.; Yabushita, Y.; Tabushi, I. Amino acid synthesis through biogenetic-type CO$_2$ fixation. *Nature* **1975**, *256*, 60–61.

114. Tanaka, K.; Matsui, T.; Tanaka, T. Catalytic formation of α-keto acids by artifical CO$_2$ fixation. *J. Am. Chem. Soc.* **1989**, *111*, 3765–3767.

115. Olsen, R.K.; Currie, J.O. Synthetic uses of thiols. In *The Chemistry of the Thiol Group, Part 2*; Patai, S., Ed.; John Wiley & Sons: London, UK, 1974; pp. 519–588.

116. Stacy, G.W.; Day, R.I.; Morath, R.J. Schiff bases and related substances II. Reactions of thiols with N-benzylidene aniline and N-benzylidene anthranilic acid. *J. Am. Chem. Soc.* **1955**, *77*, 3869–3873.

117. Oakes, T.R.; Stacy, G.W. Reactions of thiols with Schiff bases in nonaqueous solvents. Addition equilibria, cleavage, and reduction. *J. Am. Chem. Soc.* **1972**, *94*, 1594–1600.

118. Huber, C.; Wächtershäuser, G. Primordial reductive amination revisited. *Tetrahedron Lett.* **2003** *44*, 1695–1697.

119. Kochi, H.; Kikuchi, G. Reactions of glycine synthesis and glycine cleavage catalyzed by extracts of *Arthrobacter globiformis* grown on glycine. *Arch. Biochem. Biophys.* **1969**, *132*, 359–369.

120. Van Poelje, P.D.; Snell, E.E. Pyruvoyl-dependent enzymes. *Annu. Rev. Biochem.* **1990**, *59*, 29–59.

121. Hess, U.; Theile, R. Electosynthesis of N-substituted DL-arylglycineesters and 1,2-diarylamino-1,2-diarylethanes by cathodic reduction of azomethines in the presence of carbon dioxide. *J. Prakt. Chem.* **1982**, *324*, 385–399.

122. Silvestri, G.; Gambino, S.; Filardo, G. Electrochemical syntheses involving carbon dioxide. In *Enzymatic and Model Carboxylation and Reduction Reactions for Carbon Dioxide Utilization*; Arresta, M., Schloss, M.V., Eds.; Kluwer Academic Publishers: South Holland, The Netherlands, 1990; pp. 101–127.

123. Nakada, H.I.; Weinhouse, S. Non-enzymatic transamination with glyoxylic acid and various amino acids. *J. Biol. Chem.* **1953**, *204*, 831–836.

124. Howard, B.R.; Endrizzi, J.A.; Remington, S.J. Crystal structure of *Escherichia coli* malate synthase G complexed with magnesium and glyoxylate at 2.0 Å resolution: Mechanistic implications. *Biochemistry* **2000**, *39*, 3156–3168.

125. Hsu, R.Y.; Mildvan, A.; Chang, G.; Fung, C. Mechanism of malic enzyme from pigeon liver. Magnetic resonance and kinetic studies of the role of Mn^{2+}. *J. Biol. Chem.* **1976**, *251*, 6574–6583.

126. Aktas, D.F.; Cook, P.F. A lysine-tyrosine pair carries out acid-base chemistry in the metal ion-dependent pyridine dinucleotide-linked β-hydroxyacid oxidative decarboxylases. *Biochemistry* **2009**, *48*, 3565–3577.

127. Goward, C.R.; Nicholls, D.J. Malate dehydrogenase: A model for structure, evolution, and catalysis. *Protein Sci.* **1994**, *3*, 1883–1888.

128. Mukhopadhyay, B.; Stoddard, S.F.; Wolfe, R.S. Purification, regulation, and molecular and biochemical characterization of pyruvate carboxylase from *Methanobacterium thermoautotrophicum* strain ΔH. *J. Biol. Chem.* **1998**, *273*, 5155–5166.

129. Flint, D.H. *Escherichia coli* fumarase A catalyzes the isomerization of enol and keto oxalacetic acid. *Biochemistry* **1993**, *32*, 799–805.

130. White, R.H. Intermediates in the biosynthesis of Coenzyme M (2-mercaptoethane sulfonic acid). *Biochemistry* **1986**, *25*, 5304–5308.

131. White, R.H. Biosynthesis of the 7-mercaptoheptanoic acid subunit of component B [(7-mercaptoheptanoyl) threonine phosphate] of methanogenic bacteria. *Biochemistry* **1989**, *28*, 860–865.

132. Dörr, M.; Käßbohrer, J.; Grunert, R.; Kreisel, G.; Brand, W.A.; Werner, R.A.; Geilmann, H.; Apfel, C.; Robl, C.; Weigand, W.; *et al*. A possible prebiotic formation of ammonia from dinitrogen on iron sulfide surfaces. *Angew. Chem. Int. Ed.* **2003**, *42*, 1540–1543.

133. Robertson, M.P.; Joyce, G.F. The origins of the RNA world. *Cold Spring Harb. Perspect. Biol.* **2012**, *4*, doi:10.1101/cshperspect.a003608.

134. Braud, C.; Brunel, C.; Vert, M. Poly(β-malic acid): A new polymeric drug-carrier. *Polym. Bull. (Berl.)* **1985**, *13*, 293–299.

135. Karl, M.; Anderson, R.; Holler, E. Injection of poly(β-L-malate) into the plasmodium of *Physarum polycephalum* shortens the cell cycle and increases the growth rate. *Eur. J. Biochem.* **2004**, *271*, 3805–3811.

136. Liu, S.; Steinbüchel, A. Investigation of poly(β-L-malic acid) production by strains of *Aureobasidium pullulans*. *Appl. Microbiol. Biotechnol*. **1996**, *46*, 273–278.

137. Goodman, I. *Polyesters, Volume I: Saturated Polymers*; Elsevier: New York, NY, USA, 1965.

138. Kasting, J.F. Earth's earliest atmosphere. *Science* **1993**, *259*, 920–926.

139. Portilla-Arias, J.A.; García-Alvarez, M.; Martínez de Ilarduya, A.; Holler, E.; Muñoz-Guerra, S. Thermal decomposition of fungal poly(β-L-malic acid) and poly(β-L-malate)s. *Biomacromolecules* **2006**, *7*, 3283–3290.

140. Van der Sluis, P.; Kroon, J. The structure of (+/−) malic acid, (+/−)-$C_4H_6O_5$. *Acta Crystallogr.* **1985**, *C41*, 956–959.

141. Gavezzotti, A. Hydrogen bond strength and bond geometry in cyclic dimers of crystalline carboxylic acids. *Acta Crystallogr. B* **2008**, *864*, 401–403.

142. Heisler, I.A.; Mazur, K.; Yamaguchi, S.; Tominaga, K.; Meech, S.R. Measuring acetic acid dimer modes by ultra fast time-domain Raman spectroscopy. *Phys. Chem. Chem. Phys*. **2011**, *13*, 15573–15579.

143. Losada, M.; Tran, H.; Xu, Y. Lactic acid in solution: Investigations of lactic acid self-aggregation and hydrogen bonding interactions with water and methanol using vibrational absorption and vibrational circular dichroism spectroscopies. *J. Chem. Phys.* **2008**, *128*, doi:10.1063/1.2806192.

144. Nakamura, A.; Yao, M.; Chimnaronk, S.; Sakai, N.; Tanaka, I. Ammonia channel couples glutaminase with transamidase reactions in GatCAB. *Science* **2006**, *312*, 1954–1958.

145. Beccera-Figueroa, L.; Ojeda-Porras, A.; Gamba-Sánchez, D. Transamidation of carboxamides catalyzed by Fe(III) and water. *J. Org. Chem.* **2014**, *79*, 4544–4552.

146. Levy, M.; Silberman, D.E. The reactions of amino and imino acids with formaldehyde. *J. Biol. Chem.* **1937**, *118*, 723–734.

147. Kitamoto, Y.; Maeda, H. Reevaluation of the reaction of formaldehyde at low concentrations with amino acids. *J. Biochem.* **1980**, *87*, 1519–1530.

148. Argyrou, A.; Washabaugh, M.W.; Pickart, C.M. Dihydroorotate dehydrogenase from *Clostridium oroticum* is a class 1B enzyme and utilizes a concerted mechanism of catalysis. *Biochemistry* **2000**, *39*, 10373–10384.

149. Wise, E.; Yew, W.S.; Babbitt, P.C.; Gerlt, J.A.; Rayment, I. Homologous (β/α)8-barrel enzymes that catalyze unrelated reactions: Orotidine-5'-monophosphate decarboxylase and 3-keto-L-gulonate-6-phosphate decarboxylase. *Biochemistry* **2002**, *41*, 3861–3869.

150. Mittapalli, G.K.; Osornio, Y.M.; Guerrero, M.A.; Reddy, K.R.; Krishnamurthy, R.; Eschenmoser, A. Mapping the landscape of potentially primordial informational oligomers: Oligodipeptides tagged with 2,4-disubstituted 5-aminopyrimidines as recognition elements. *Angew. Chem. Int. Ed.* **2007**, *46*, 2478–2484.

151. Stuecker, T.N.; Hodge, K.M.; Escalante-Semerena, J.C. The missing link in coenzyme A biosynthesis: PanM (formerly YhhK), a yeast GCN5 acetyltransferase homologue triggers aspartate decarboxylase (PanD) maturation in *Salmonella enterica*. *Mol. Microbiol.* **2012**, *84*, 608–619.

152. Chladek, S. The synthesis, reactions and properties of the 2'(3')-O-aminoacyl and peptidyl nucleosides and nucleotides. In *Chemistry of Nucleosides and Nucleotides*; Townsend, L.B., Ed.; Springer Science & Business Media: New York, NY, USA, 1994; Volume 3, pp. 107–143.

153. Gavrilova, L.P.; Kostiashkina, O.E.; Koteliansky, V.E.; Rutkevitch, N.M.; Spirin, A.S. Factor-free ("non-enzymic") and factor-dependent systems of translation of polyuridylic acid by *Escherichia coli* ribosomes. *J. Mol. Biol.* **1976**, *101*, 537–552.

154. Southworth, D.R.; Brunelle, J.L.; Green, R. EFG-independent translocation of the mRNA:tRNA complex is promoted by modification of the ribosome with thiol-specific reagents. *J. Mol. Biol.* **2002**, *324*, 611–623.

155. Rich, A. On the problems of evolution and biochemical information transfer. In *Horizons in Biochemistry*; Kasha, M., Pullman, B., Eds.; Academic Press: New York, NY, USA, 1962; pp. 103–126.

156. Weber, A.L.; Orgel, L.E. Poly(U)-directed peptide bond formation from the 2'(3')-glycyl esters of adenosine derivatives. *J. Mol. Evol.* **1980**, *16*, 1–10.

157. Reusch, R.N. Physiological importance of poly-(R)-3-hydroxybutyrates. *Chem. Biodivers.* **2012**, *9*, 2343–2366.

158. Brooks, D.J.; Fresco, J.R. Increased frequency of cysteine, tyrosine, and phenylalanine residues since the last universal ancestor. *Mol. Cell. Proteomics* **2002**, *1*, 125–131.

159. Francis, B.R. Evolution of the genetic code by incorporation of amino acids that improved or changed protein function. *J. Mol. Evol.* **2013**, *77*, 134–158.

160. Eigen, M.; Winkler-Oswatitsch, R. Transfer-RNA, an early gene? *Naturwissenschaften* **1981**, *68*, 217–228.

161. Ikehara, K.; Omori, Y.; Arai, R.; Hirose, A. A novel theory on the origin of the genetic code: A GNC-SNS hypothesis. *J. Mol. Evol.* **2002**, *54*, 530–538.

162. Higgs, P.G. A four-column theory for the origin of the genetic code: Tracing the evolutionary pathways that gave rise to an optimized code. *Biol. Direct* **2009**, *4*, doi:10.1186/1745-6150-4-16.

163. Van der Gulik, P.; Massar, S.; Gilis, D.; Burhman, H.; Rooman, M. The first peptides: The evolutionary transition between prebiotic amino acids and early proteins. *J. Theor. Biol.* **2009**, *261*, 531–539.

164. Fujihashi, M.; Numoto, N.; Kobayashi, Y.; Mizushima, A.; Tsujimura, M.; Nakamura, A.; Kawarabayasi, Y.; Miki, K. Crystal structure of archaeal photolyase from *Sulfolobus tokodaii* with two FAD molecules: Implication of a novel light-harvesting cofactor. *J. Mol. Biol.* **2007**, *365*, 903–910.

165. Deamer, D.W.; Dworkin, J.P.; Sandford, S.A.; Bernstein, M.P.; Allamandola, L.J. The first cell membranes. *Astrobiology* **2002**, *2*, 371–381.

166. Segré, D.; Ben-Eli, D.; Deamer, D.W.; Lancet, D. The Lipid World. *Orig. Life Evol. Biosph.* **2001**, *31*, 119–145.

167. Russell, M.J.; Hall, A.J. The emergence of life from iron monosulphide bubbles at a submarine hydrothermal redox and pH front. *J. Geol. Soc. Lond.* **1997**, *154*, 377–402.

168. Alpermann, T.; Rüdel, K.; Rüger, R.; Steiniger, F.; Nietzsche, S.; Filiz, V.; Förster, S.; Fahr, A.; Weigand, W. Polymersomes containing iron sulfide (FeS) as primordial cell model: For the investigation of energy providing redox reactions. *Orig. Life Evol. Biosph.* **2010**, *41*, 103–119.

169. Menzel, K.; Menzel, K.; Apfel, U.P.; Wolter, N.; Rüger, R.; Alpermann, T.; Steiniger, F.; Gabel, D.; Förster, S.; Weigand, W.; *et al.* [FeFe]-hydrogenase models assembled into vesicular structures. *J. Liposome Res.* **2014**, *24*, 59–68.

170. Summers, D.P.; Noveron, J.; Basa, R.C. Energy transduction inside of amphiphilic vesicles: Encapsulation of photochemically active semiconducting particles. *Orig. Life Evol. Biosph.* **2009**, *39*, 127–140.

171. Emiliozzi, R.; Pichat, L.; Herbert, M. Preparation of cysteine-S^{35} hydrochloride. *Bull. Soc. Chem. Fr.* **1959**, 1544–1545.

172. Orgel, L.E. Some consequences of the RNA world hypothesis. *Orig. Life Evol. Biosph.* **2003**, *33*, 211–218.

173. Ross, D.S. The viability of a nonenzymatic reductive citric acid cycle—Kinetics and thermochemistry. *Orig. Life Evol. Biosph.* **2007**, *37*, 61–65.

174. Orgel, L.E. The implausibility of metabolic cycles on the prebiotic Earth. *PLoS Biol.* **2008**, *6*, doi:10.1371/journal.pbio.0060018.

175. Copley, S.D.; Smith, E.; Morowitz, H.J. The emergence of sparse metabolic networks. *J. Cosmol.* **2010**, *10*, 3345–3361.

176. Engelhart, A.E.; Hud, N.V. Primitive genetic polymers. *Cold Spring Harb. Perspect. Biol.* **2010**, *2*, doi:10.1101/cshperspect.a002196.

177. Szathmáry, E. Why are there four letters in the genetic alphabet? *Nat. Rev. Genet.* **2003**, *4*, 995–1001.

178. Calvin, M. *Chemical Evolution: Molecular Evolution towards the Origin of Living Systems on the Earth and Elsewhere*; Clarendon Press: London, UK, 1969.

179. Rauchfuss, H. *Chemical Evolution and the Origin of Life*; Springer-Verlag: Berlin/Heidelberg, Germany, 2008.

180. Schrum, J.P.; Zhu, T.F.; Szostak, J.W. The origins of cellular life. *Cold Spring Harb. Perspect. Biol.* **2010**, *2*, doi:10.1101/cshperspect.a002212.

181. Mansy, S.S.; Schrum, J.P.; Krishnamurthy, M.; Tobé, S.; Treco, D.A.; Szostak, J.W. Template-directed synthesis of a genetic polymer in a model protocell. *Nature* **2008**, *454*, 122–125.