

# Reconstructing the Ancestral Routes to Nucleobases, D-Sugars, RNAs, L- $\alpha$ -Amino Acids, and Lipids

Sylvain L. Smadja

October 29, 2010 L

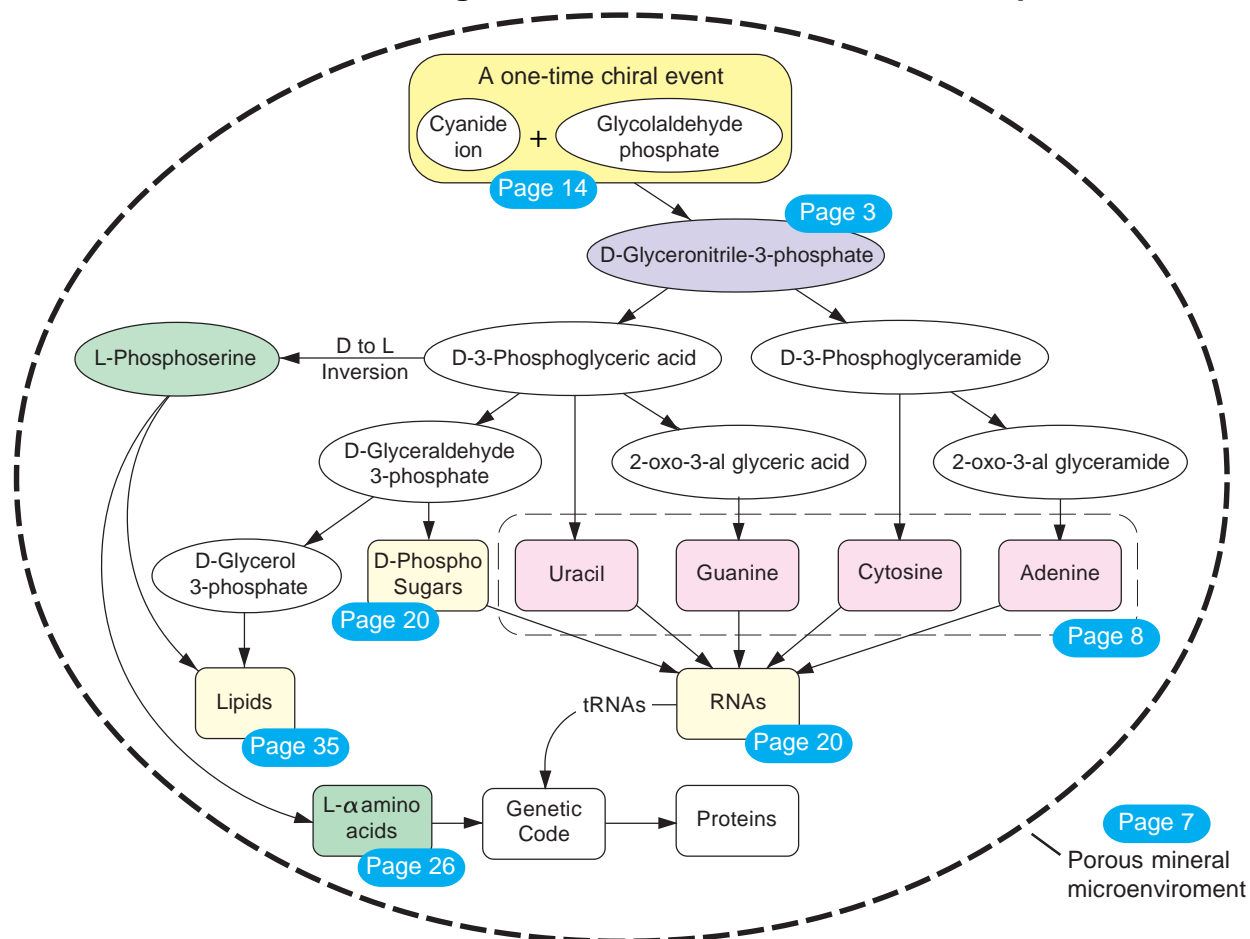
Copyright © 2010 Sylvain Smadja. All rights reserved.

Email: slsmadja@innido.com

## Abstract

After decades of investigation, many challenges remain to understanding Life's emergence from prebiotic chemicals. Adequate models are needed to explain very early pathways that produced the relevant nucleobases, D-sugars, L- $\alpha$ -amino acids, straight-chain fatty acids, and branched-chain isoprenoid compounds. It was the enigma of biological homochirality that first intrigued the author. But the scope of his model grew as his investigations led to insights about the origin of biomolecular building block units. By searching modern biology for molecular fossils, using retrosynthetic analysis, and finding support in experimental data, the author ended up constructing a full-picture model of possible prebiotic pathways to relevant biomolecules. The author proposes that within a local sheltered-microenvironment of early Earth an asymmetric amplification process that propagated from a tiny initial imbalance gave rise to a chiral three-carbon cyanohydrin following a reaction between cyanide ions and glycoaldehyde phosphate entities. Starting with this chiral cyanohydrin acting as the 'Premier Precursor,' the author suggests sequences of cascading chemical steps that yield nucleobases, D-sugars, L- $\alpha$ -amino acids, and lipids. This model holds that biological homochirality was established as a consequence of just a single, fortuitous symmetry-breaking event, and that the proposed early chemical events took place as a 'one-pot' reaction system which yielded all necessary biomolecules to usher in the RNA/Protein/Lipid World.

## Full Picture Model of Ancestral Routes to Nucleobases, D-Sugars, RNAs, L- $\alpha$ -Amino Acids, and Lipids



Proposed 'one-pot' reaction system within a porous mineral microenvironment that yielded all the necessary biomolecules to usher in the RNA/Protein/Lipid World.

Section	Title	Page
1	Proposed Premier Precursor of Biomolecules	3
2	Proposed Mechanisms for Ancestral Routes to Nucleobases	8
3	Induction and Preservation of Chirality	14
4	Proposed Ancestral Routes to D-Ribose and to RNAs	20
5	L-Phosphoserine: Proposed Precursor of Left-handed $\alpha$ -Amino Acids	26
6	Proposed Ancestral Routes to Lipids	35

## Introduction

To understand how Life could have emerged from prebiotic chemicals it is necessary to explain the processes that led to the specific nucleobases, relevant sugars,  $\alpha$ -amino acids, straight-chain fatty acids of even-number carbons, and branched-chain isoprenoid compounds that are found in the living world. Further, is the necessity to explain why all the biologically relevant sugars are only the D-enantiomers, and all the standard amino acids are only the L-enantiomers. To this day, the origin of this biological homochirality is one of the most puzzling problems. In this paper I propose a novel model that provides compelling answers to those challenges. Using a Top-Down approach, retrosynthetic analysis, and molecular archeology, I have endeavored to reconstruct the earliest pathways to the biomolecules. The sequences of events that I propose, takes place during the prebiotic era, not in a primordial soup, but rather in a local, sheltered microenvironment that provided a host structure for entrapped organic and mineral compounds. I show how, within that self-contained environment, the chemical processes that led to Life could have occurred in a logical progression of natural, easy steps. What I offer is a coherent, plausible, full-picture model. My scenario starts with just a few small molecules that, through their interactions within the confined environment, end up generating the complex structures of the biomolecules that yield the RNA world and further, the RNA/Protein/Lipid world from which Life originated.

This paper is divided into six sections: (1) One chiral cyanohydrin, the Premier Precursor; (2) Mechanisms for ancestral routes to nucleobases; (3) Induction and preservation of chirality; (4) Ancestral routes to D-ribose and to RNAs; (5) L-phosphoserine, the precursor of a whole class of L- $\alpha$ -amino acids, and; (6) Ancestral routes to lipids.

## 1. One Chiral Cyanohydrin

### The Proposed Premier Precursor of Biomolecules

In my attempts to reconstruct the ancestral routes to the biomolecules I realized that a single key-reaction in the prebiotic era could have triggered the formation of several three-carbon entities that I propose were the precursors that led to nucleobases, sugars, amino acids, and lipids. I also realized that if homochirality was induced at that earliest step, homochirality would be built-into those resulting three-carbon units, then transferred and maintained from that time forward. (Homochirality is discussed in detail in Section 3.)

That single key-reaction was the attack at the carbonyl group of glycolaldehyde phosphate by the nucleophilic cyanide ion ( $\text{CN}^-$ ). This reaction yielded a chiral cyanohydrin, namely glyceronitrile-3-phosphate that I propose was the premier precursor of the biomolecules (Fig. 1, step 1). This three-carbon nitrile gave rise to seven three-carbon entities as follows: Glyceronitrile-3-phosphate is transformed into 3-phosphoglyceramide by the addition of one water molecule (Fig. 1, step 2), and into 3-phosphoglyceric acid by the addition of two water molecules (Fig. 1, step 3). After dephosphorylation followed by dehydrogenation, 3-phosphoglyceramide is transformed into 2-oxo-3-al glyceramide (Fig. 1, step 4), and 3-phosphoglyceric acid transforms into 2-oxo-3-al glyceric acid (Fig. 1, step 5). 3-phosphoglyceric acid is transformed into glyceraldehyde-3-phosphate by a one-step reduction of the carboxylic group (Fig. 1, step 6), and into phosphoserine after a nucleophilic attack of an ammonia molecule at the C2 position (Fig. 1, step 7). After a reduction of the carbonyl group, glyceraldehyde-3-phosphate is transformed into glycerol-3-phosphate (Fig. 1, step 8).

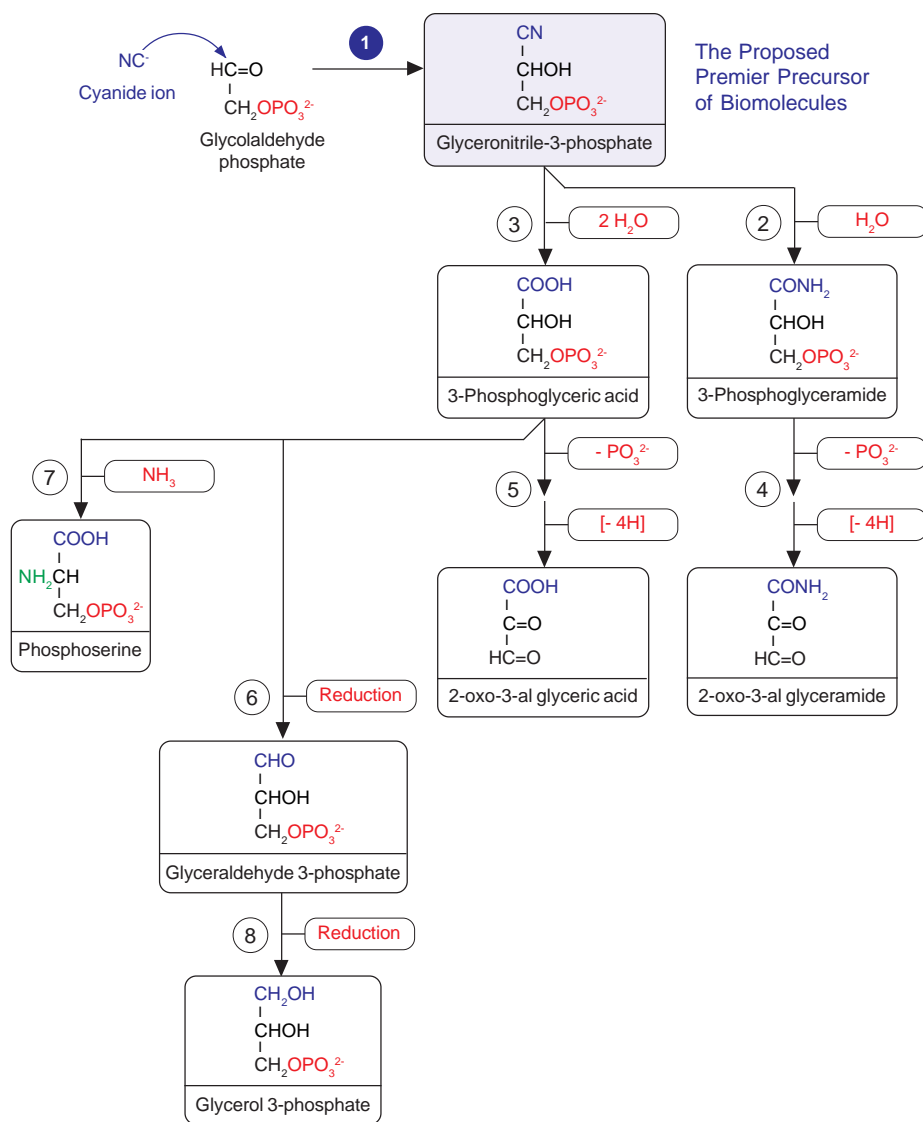


Fig 1. The proposed three-carbon precursors of families of biomolecules

Experimental results from modern-day organic chemistry can help illustrate the chemical mechanisms that would have allowed the synthesis of glyceronitrile-3-phosphate (Fig. 1, step1) to have taken place in the prebiotic era. The route to cyanohydrins was introduced to the world of organic chemistry as early 1887 (1). This classical process was first used to prepare aldonic acids that exhibit one more carbon than the starting aldehyde (through

hydrolysis of the cyano group). Adding one step to this process creates aldoses by the reduction of aldonic acids to their respective aldehydes. Such synthetic sequences, known as Kiliani-Fisher reactions, have been used by several investigators especially in the preparations of aldoses containing radioactive isotopes. For example, authors were able to synthesize carbon-13-enriched triose, tetrose and pentose phosphates following the catalytic reduction of the respective three-, four-, and five-carbon aldonitrile phosphates. Relevant to the reactions illustrated in Figure 1, it is worth mentioning that [1-<sup>13</sup>C] glyceronitrile-3-phosphate was synthesized following the condensation of K<sup>13</sup>CN with glycoaldehyde phosphate (in a pathway that is similar to Fig. 1, step 1). Glyceraldehyde-3-phosphate was then produced after reduction of glyceronitrile-3-phosphate (2).

For the reaction between cyanide (CN<sup>-</sup>) and glycoaldehyde phosphate to take place in the prebiotic era we need to consider the sources and availability of the reactants and the environmental context in which the reactions could have taken place.

Hydrogen cyanide (HCN): This presumably became an abundant component of the early Earth from exogenous and/or endogenous sources. HCN has been detected in the gas phase of the interstellar medium as well as in the comae of a variety of comets (3). It is often assumed that comets delivered compounds such as hydrogen cyanide on the surface of our young planet (4). An endogenous source of HCN could have been from volcanic activity. Cyanoligands of transition metals have been identified in extant volcanic sites (5). For many years scientists have considered HCN to be a plausible source of purines, pyrimidines and amino acids on the primitive Earth (6). Most experiments conducted to illustrate how biomolecules could form from HCN, or cyanide CN<sup>-</sup>, involved aqueous alkaline cyanide.

However, according to some authors, free dissolved cyanide  $\text{CN}^-$  or HCN would not have been plausible as efficient reacting species. Instead, they favored cyanide ions attached as ligands to transition metals. These authors produced a series of  $\alpha$ -hydroxy and  $\alpha$ -amino acids after reactions of  $\text{Fe}^{2+}$  or  $\text{Ni}^{2+}$  cyano complexes with carbon monoxide (CO). Interestingly, among the variety of compounds produced by these experiments, ethylene glycol was identified. This led the authors to suggest a pathway that involved glycolaldehyde (7).

Glycolaldehyde: This has been detected in interstellar clouds and could have been delivered to the early Earth from space (8). There are two main routes to glycolaldehyde. It is produced from two formaldehydes at the early stage of the formose reaction, and from HCN in the following way: Formic acid that is derived from HCN can be reduced to formaldehyde. HCN and formaldehyde can react to produce glycolonitrile that can be transformed to glycolic acid. Reduction then gives rise to glycolaldehyde. A similar route to glycolaldehyde was used by a group of researchers who prepared formaldehyde from a direct reduction of HCN and then produced glycolonitrile from HCN and HCHO. Finally, glycolaldehyde was obtained from the direct reduction of glycolonitrile (9).

Glycolaldehyde phosphate: This could have been created through various pathways as shown by several experiments. Some authors have demonstrated easy phosphorylation of glycolaldehyde with amidotriphosphate (10). Urea-inorganic phosphate has been considered a plausible mixture that can be used in prebiotic phosphorylation (11). It has also been shown that in presence of formamide, a wide range of phosphate minerals act as phosphate donors (12).

One can reasonably assume that hydrogen cyanide, glycolaldehyde, and therefore glycolaldehyde phosphate, would have been readily available in the habitats of early Earth. Now let's consider the environmental settings in which the proposed chemical reactions could have taken place.

My suggested chemical step-pathways take place in a non-aqueous environment. This is in agreement with authors who doubt that early metabolisms could have come about by random collisions in dilute aqueous primordial soups. They suggest that internal surfaces of minerals may have played a crucial role in selecting and concentrating relevant small molecules and ions that led to the formation of complex bio-systems. Comprehensive reviews on the role of minerals in the origin of Life have been published in 1999 (13) and in 2004 (14).

Mineral host-structures could have also provided a sheltered microenvironment, as well as catalytic sites, for the organic and inorganic compounds that they encapsulated. Various categories of minerals have been investigated to determine if they could have met the necessary criteria. Among the most representative groups of minerals are the double-layer hydroxides (DLH) (15), and weathered feldspars (16). DLH can be derived from brucite ( $\text{Mg}(\text{OH})_2$ ) by the replacement of divalent ions with trivalent ions. This creates an excess positive charge that needs to be neutralized. Molecular ions that are important sources of biomolecule structures, such as  $\text{CN}^-$ ,  $\text{Fe}(\text{CN})_6^{2-}$ ,  $\text{CO}_3^{2-}$ ,  $\text{H}(\text{PO}_4)^{2-}$ , etc., are negatively charged and could have been attracted into the interlayer of the DLH structure to compensate the positive charge. Also,  $\text{H}_2\text{O}$  molecules are generally associated with sorbed anions in the interlayer. Further, the internal surface of DLH exhibits sheet structures that are capable of

accommodating molecules of any size, and these groups of minerals have particularly high catalytic activity (15).

Silica-rich surfaces of partly dealuminated feldspars and zeolites may have played important roles during the emergence of life (16). Such weathered alkali feldspar crystals contain tubular etch pits that form an orthogonal honeycomb network. “Sub-micrometer tubes in the honeycomb might have acted as rudimentary cell-walls for proto-organisms, which ultimately evolved a lipid lid...” (16). It is worth noting that feldspars commonly contain a large variety of mineral inclusions such as clays, the phosphate mineral apatite, carbonates, as well as oxides and sulfides of various metals including Fe, Ag, and Ti.

Because the geological record from the prebiotic era is no longer available, one can only speculate on the specific minerals that may have been involved in the creation of the first biomolecules. DLH and weathered feldspars could have been suitable candidates. However, other mineral structures with similar qualities could have provided the compartmentalized habitats in which the processes of early metabolism could have taken place. Such a setting is the local microenvironment of the ‘one-pot’ reaction system that I am proposing.

## **2. Proposed Mechanisms for Ancestral Routes to Nucleobases**

In charting the path to early nucleobases, I suggest that the first four of the three-carbon entities that derived from glyceronitrile-3-phosphate (Fig. 1, steps 2 to 5) were key precursors in the early building of pyrimidine and purine rings. To build the rings, these three-carbon fragments combined with small molecules such as formamide, ammonia and urea. It can be

pointed out that those small molecules can derive from HCN. Formamide is produced from HCN hydrolysis, ammonia results from hydrolysis of formamide, and urea has been identified among the products of HCN oligomerization (17). In what follows, I illustrate the ways the three-carbon fragments could have been pieced together to build early nucleobase rings. Although alternative chronological orders for guanine and adenine can be envisioned, the steps would lead to the same end products. (How nucleobases could have assembled on a sugar-phosphate backbone is discussed in Section 4).

Proposed Early Routes to Pyrimidine Rings:

*Uracil:* 3-phosphoglyceric acid condenses with urea and produces a hydroxylated uracil ring (Fig. 2, step 1a). Then, just after the elimination of water which leads to a double bond formation, uracil is created (Fig. 2, step 2a).

*Cytosine:* 3-phosphoglyceramide combines with urea to give a hydroxylated cytosine ring (Fig. 2, step 1b). Then, just after the elimination of water which leads to a double bond formation, cytosine is created (Fig. 2, step 2b).

It is interesting to compare the proposed early routes with experimental attempts to create cytosine and uracil under presumed prebiotic conditions. Experimentalists have combined three-carbon fragments with urea, but instead of starting with 3-phosphoglyceric acid, or 3-phosphoglyceramide, as I propose, they used cyanoacetylene (or its hydrolysis product, cyanoacetaldehyde) (18). However, there is doubt about cyanoacetylene or cyanoacetaldehyde being part of the prebiotic environment (19). On the other hand, 3-phosphoglyceric acid and 3-phosphoglyceramide appear to be more plausible prebiotic reactants because they are readily formed and appear in the same proposed microenvironment

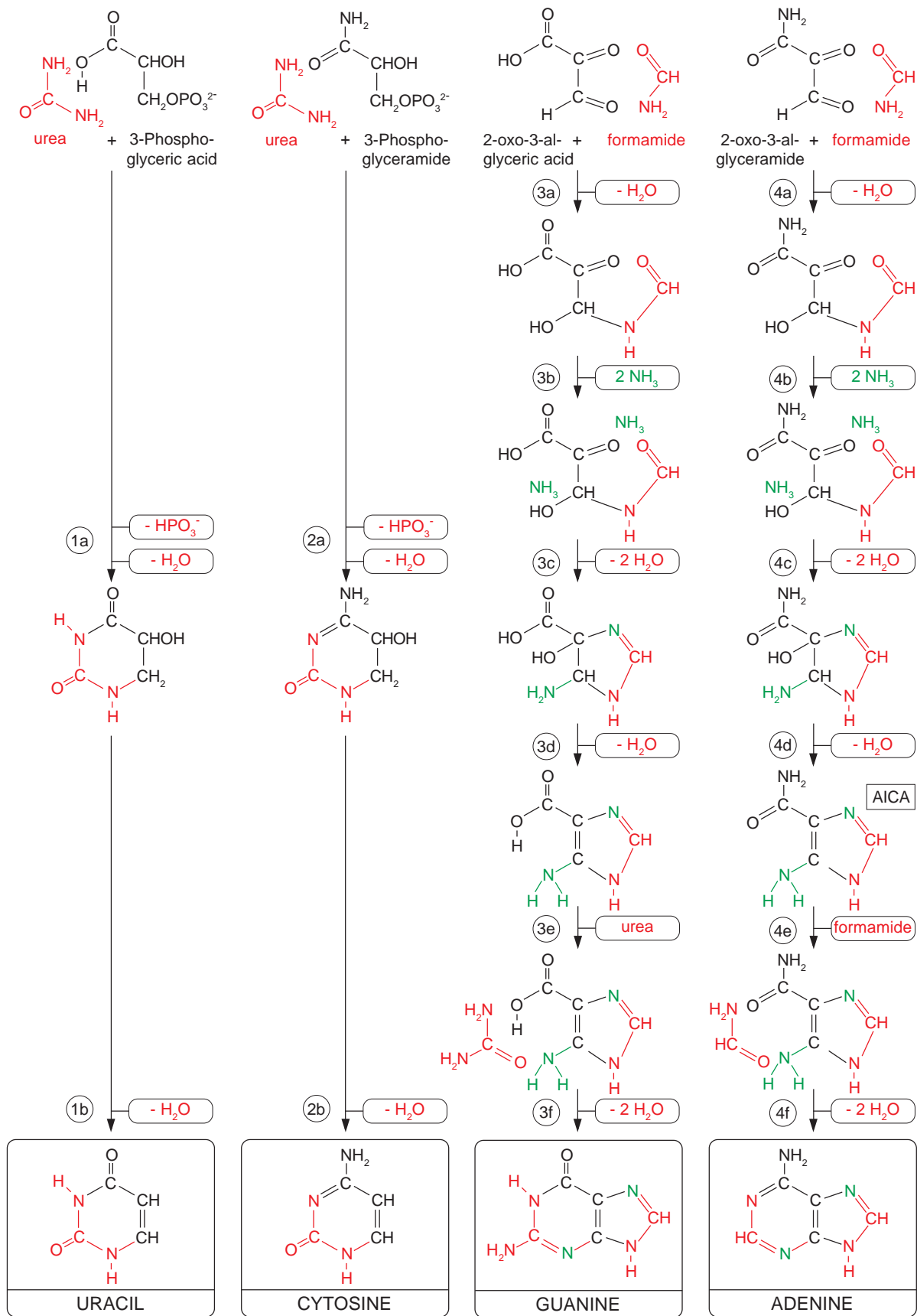


Fig 2. Proposed Step-Pathways for Ancestral Routes to Nucleobases

among the other products that become the precursors of other important biofamilies such as purine rings,  $\alpha$ -amino acids, sugars, and lipids.

#### Proposed Early Routes to Purine Rings:

*Guanine:* The proposed route to guanine involves 2-oxo-3-al glyceric acid, a three-carbon entity that could have resulted from dehydrogenation of 3-phosphoglyceric acid (Fig. 1, step 4) by, for example, the action of  $\text{CO}_2$  which is known to be an electron acceptor.  $\text{CO}_2$  could have been created from formic acid, a derivative of HCN. ( $\text{NO}_3^-$  and  $\text{NO}_2^-$  are also recognized as plausible prebiotic electron acceptors).

In a first step to guanine, 2-oxo-3-al glyceric acid combines with formamide (Fig. 2, step 1c). Two well-positioned ammonia molecules are able to condense and one of them provokes a ring closure (Fig. 2, steps 2c and 3c). A double bond creation after water elimination leads to the formation of an imidazole ring (Fig 2, step 4c). Guanine is created following the condensation of urea with the last intermediate and after elimination of  $2\text{H}_2\text{O}$  (Fig. 2, steps 5c and 6c).

*Adenine:* Creation of adenine involves 2-oxo-3-al-glyceramide, a three-carbon entity resulting from dehydrogenation of 3-phosphoglyceramide (Fig. 1, step 4). The first four steps (Fig. 2, steps 1d to 4d) are similar to the first steps of guanine creation described above (Fig. 2, steps 1c to 4c). Adenine is produced following the condensation of formamide with the last intermediate and after elimination of  $2\text{H}_2\text{O}$  (Fig. 2, steps 5d and 6d).

It is interesting to note that in the suggested route to adenine, the intermediate that forms at step 4d (of Fig. 2), is exactly the same entity, known as 4-amino-imidazole-5-carboxamide (AICA), that is attached to ribose-5' phosphate during one of the last stages

of the biosynthetic route to inosine monophosphate (IMP). In the living-world, IMP is converted to adenine monophosphate and to guanine monophosphate.

Now, let's compare my proposed early pathway to purines with the routes suggested by researchers who attempted to create adenine and guanine under assumed prebiotic conditions. In the 1960s, adenine was produced among numerous (not always well-characterized) compounds by refluxing ammonium cyanide solutions (20). A great deal of effort has been devoted to understanding the process behind the formation of adenine under such experimental conditions. Scientists have suggested complex multi-step mechanisms that require numerous variations in the environmental conditions. Their pathways start with the formation of an HCN tetramer followed by a cyclization process which then requires UV radiation and finally a further attack from HCN. Nevertheless, none of these proposed steps (reactions) has been shown to contribute directly to adenine synthesis under the conditions employed by experimentalists (21).

I submit that in such experimental attempts, adenine is produced in step-pathways that possibly share similarities with the ones that I have just proposed that start, not with an HCN tetramer, but with a three-carbon fragment such as 2-oxo-3-al glyceramide or 2-oxo-3-al glyconitrile. In this regard it is worth mentioning that experiments show that the addition of glycolonitrile (the two-carbon cyanohydrin produced from HCN and formaldehyde) to solutions of HCN in the presence of  $\text{NH}_4\text{OH}$ , increases the yield of adenine by about five fold (22). One possible way to explain the increase is that in a first step a three-carbon entity was created after a reaction of HCN with glycolonitrile. Then, a derivative of this three-carbon fragment, together with ammonia and formamide molecules, led to adenine.

It is interesting to note that different nucleobases have been identified in meteorites such as the Murchison, and in experiments involving HCN's oligomerization. The model that I am proposing (based on the three-carbon fragment precursors to cytosine, uracil, adenine and guanine) may help explain the origin of these nucleobases.

Researchers identified 4-hydroxypyrimidine and 4-hydroxy-2-methyl-pyrimidine in the Murchison meteorite (23). As illustrated in Fig. 3A, the first of those two nucleobases could have been produced from a direct condensation of formamide with glyceramide while the second nucleobase, as shown in Fig. 3B, could have formed after a direct condensation of acetamide with glyceramide.

In experiments involving HCN's oligomerization, authors have identified 5-hydroxyuracil and 4,5-dihydroxypyrimidine as products (6). As illustrated in Fig. 3C, 5-hydroxyuracil could have been created from the condensation of urea with 2-oxo-glyceric acid. This last molecule could have derived from a partial oxidation of glyceric acid. As shown in Fig. 3D,

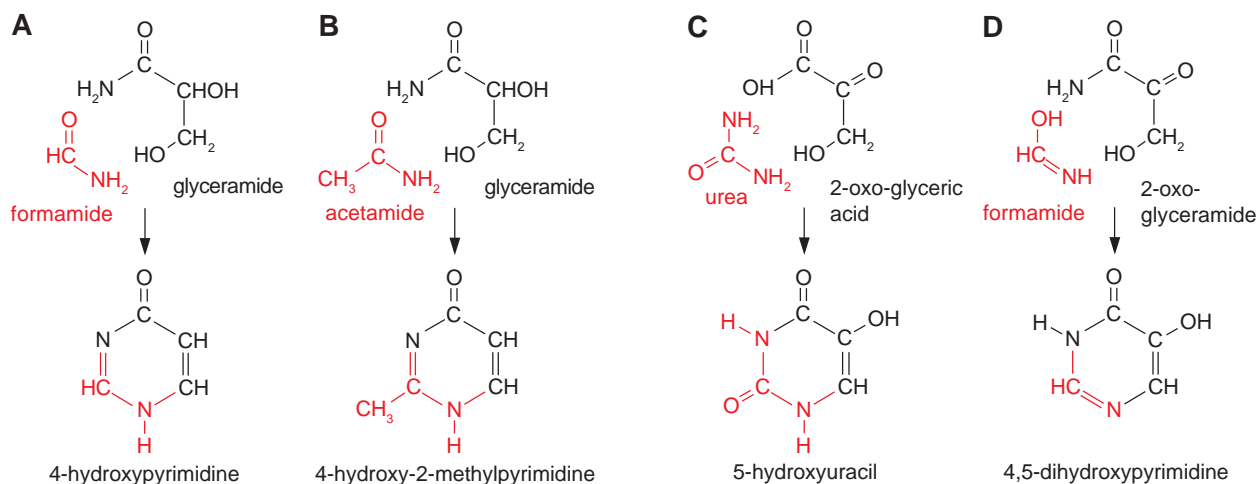


Fig 3. Proposed fragments from which pyrimidines could have been built:  
 A-B: In meteorites (23). C-D: In experiments utilizing cyanide (6).

the base 4,5-dihydropyrimidine could have been the result of the condensation of formamide with 2-oxo-glyceramide, an oxidized derivative of glyceramide.

The just described step-pathways help substantiate my argument that three-carbon fragments, derivatives of glyceronitrile, were the precursors of the nucleobases that became the building block units of our modern RNA. Discussing the origin of RNA, Origin of Life specialists stated: “A structure of this complexity is considered by many unlikely to appear by random molecular evolution. Simple (unidentified) nitriles may possibly have served as precursors of nucleobases in early forms of self-recognizing molecules” (24). I suggest that glyceronitrile-3-phosphate was one of those “unidentified nitriles.”

### **3. Induction and preservation of chirality**

Living entities are found with left-handed amino acids and right handed-sugars. Explaining the emergence of this biological homochirality has been one of the biggest challenges for those working on the Origin of Life. It has been assumed that the left-handed amino acids and right-handed carbohydrates incorporated in the early forms of life were somehow selected from a hypothetical ‘primordial soup’ comprised of racemic mixtures of myriad organics. In this section I offer a novel concept to explain the emergence of homochirality.

I propose that the emergence of the relevant D-sugars and L- $\alpha$ -amino acids took place in a local sheltered microenvironment and they originated in their pure enantiomeric forms because they derived from a series of chiral step-pathways that started with one single

symmetry-breaking event. In my scenario the single symmetry-breaking event was the synthesis of glyceronitrile-3-phosphate in the pure right-handed form (Fig. 1, step 1).

To understand how D-glyceronitrile-3-phosphate could have been abiotically produced on the early Earth, and how its synthesis could have induced the chirality of other derivatives, it is helpful to review relevant experimental results of organic chemistry. Organic chemists use enantiopure cyanohydrins to generate a wide range of chiral compounds. Comprehensive reviews of such syntheses, starting from chiral cyanohydrins, have been published in 1993 (25) and in 2004 (26). Relevant to my model are data showing that chiral cyanohydrins can be converted, with retention of configuration, to chiral  $\alpha$ -hydroxy acids (27) and to chiral  $\alpha$ -hydroxy aldehydes (28). Also of interest is the experimental work of authors who produced chiral  $\alpha$ -amino acids with an inversion of configuration. Such synthesis involved the  $\alpha$ -hydroxyl group of (R)-cyanohydrins. In a first step, the OH function was transformed into a good-leaving group which was further exchanged stereoselectively with a variety of nucleophiles. The mechanism of such nucleophilic substitutions was solely of the  $S_N2$  type. Consequently this provoked a complete inversion of configuration. Through such steps, starting from (R)-cyanohydrins, the authors were able to create (S)- $\alpha$ -amino nitriles that were later converted to (S)- $\alpha$ -amino acids (29).

Now, let's consider how chiral cyanohydrins are produced. Several different general methods are available to access enantiopure cyanohydrins starting from prochiral aldehydes or ketones. These processes require the use of a chiral catalyst such as chiral metals and chiral dipeptides. A good number of chiral cyanohydrins have also been prepared using enzymatic catalysts such as lipases and hydroxy nitrile lyases (25). The main idea behind

these catalytic processes is to create a situation in which one face of the carbonyl group is “protected” thus causing a preferential attack by cyanide from the “unprotected” side of the carbonyl group.

The above experimental data shows that, given the right conditions, chiral cyanohydrins can be converted to chiral  $\alpha$ -hydroxy acids,  $\alpha$ -hydroxy aldehydes and  $\alpha$ -amino acids. Extrapolating these results to prebiotic chemistry, let's return to my proposed scenario in which glyceronitrile-3-phosphate is created as a pure right-handed entity.

One can envision that at some point during the prebiotic era glyceronitrile-3-phosphate emerged in a pure right-handed form following the attack of glycolaldehyde phosphate by  $\text{CN}^-$  ions. But how could such an achiral-to-chiral transition have taken place?

One possibility is that it was initiated by a tiny enantiomeric imbalance that was then amplified by an autocatalytic mechanism. The remarkable work of Soai and his collaborators provide insight of how such a mechanism could have occurred (30). As stated by one author, experimental studies of the Soai reaction tend to indicate “. . . that statistical formation of dimer catalyst species coupled with lower activity of the heterochiral dimer is sufficient to rationalize the evolution of high ee from a tiny initial imbalance.” In her conclusion the author wrote, “. . . this general mechanism could be effective in a world of simple organic molecules such as those likely to have been present in the prebiotic world.” (31)

Another possibility for glyceronitrile-3-phosphate to have been created as a pure right-handed entity is that its enantioselective synthesis took place in a local sheltered microenvironment exhibiting a chiral mineral surface. Chiral mineral surfaces are ubiquitous in Earth's crust. Quartz is known for displaying several different chiral surfaces. A rich

variety of chiral solid surfaces are also provided by rock-forming centric crystals such as found in pyroxene, calcite, gypsum, barite and apatite (calcium phosphate).

Although some authors have argued that quartz and other minerals could not have contributed to the emergence of biological homochirality, their arguments are based on the assumption that the chemical events that led to the Origin of Life took place on a global scale, at several different geographical locations of the early Earth. According to this global-scale view, even if enantiomorphic crystals served as asymmetric inductors, they could not have been the source of biological homochirality because both types of enantiomorphs would randomly occur around the globe (32). But a counter-argument was offered by scientists who stated, “We hold a different view: we assume that the nucleation of self-replicating molecular systems is infrequent, whereas the growth of such systems once nucleated is relatively rapid and efficient. In this scenario, a single successful self-replicating chiral synthesis (for example, on one chiral crystal face on one calcite crystal) purely by chance became the dominant biochemical overprint in spite of initially racemic mixtures both of molecules and surfaces” (33). This last statement is in agreement with suggestions that a chance event could have led to enantiomeric excess and eventually to enantiomeric purity.

It is also possible that a still-unknown process was responsible for right-handed glyceronitile-3-phosphate entities to have emerged very early, ahead of Life. For example this could have involved the participation of a preexisting enantioenriched catalyst that would have driven the reaction toward the production one enantiomeric pure compound.

Given that D-glyceronitile-3-phosphate molecules would have been part of the primitive environment, what could have this led to?

As proposed in Section 1, this chiral three-carbon cyanohydrin would have acted as the ‘Premier Precursor’ producing a set of three-carbon molecules each of which would have played a major role as starting points for various families of relevant biomolecules.

In this section I will focus on the specific compounds whose chirality would have been transferred to their derivatives ultimately giving rise to D-sugars and L- $\alpha$ -amino acids.

As illustrated in Figure 4, hydrolysis of D-glycero-nitrile-3-phosphate could produce D-3-phosphoglyceric acid.

A one-step reduction could create D-glyceraldehyde-3-phosphate. (As discussed in Section 4, this compound could have been involved in the formation of relevant D-sugars).

A further reduction could produce D-glycerol-3-phosphate. In the present-day living world this compound, known as sn-glycerol-1-phosphate, is the backbone of the membrane phospholipids of Archea (See Section 6).

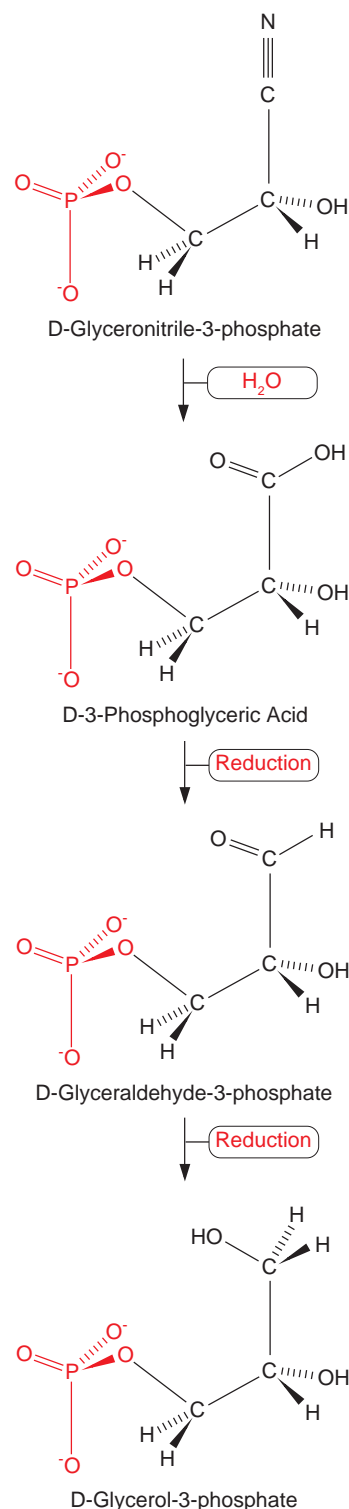


Fig 4. Chemical steps with retention of chirality at the chiral center C2.

Returning to D-3-phosphoglyceric acid, I propose that this was the starting point for the creation of a left-handed  $\alpha$ -amino acid, namely L-phosphoserine. This early chemical event could have taken place as follows.

A surface-bonded D-3-phosphoglyceric acid entity could have been in such a favorable conformation that it would have entrapped, through hydrogen-bonding, an ammonia molecule (Fig. 5). The properly oriented ammonia molecule, being at the vicinity of carbon C2 would have been able to displace the OH group. Such a nucleophilic attack from  $\text{NH}_3$  occurring on the other side of the hydroxyl would have provoked an inversion of configuration and created L-phosphoserine, a left-handed  $\alpha$ -amino acid. It is interesting to note that in the present-day living world, L-phosphoserine is biosynthesized from D-3-phosphoglycerate.

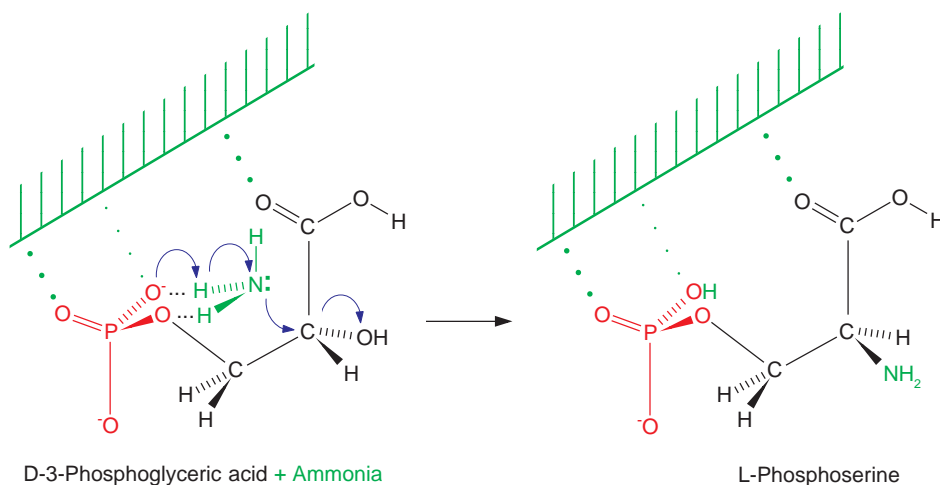


Fig 5. Proposed formation of L-Phosphoserine from D-3-Phosphoglyceric acid. Note the inversion from right-handedness to left-handedness.

According to my proposed model, relevant amino acids derived from L-phosphoserine and thus inherited both the specific carbon skeleton of alpha-amino acid, and left-handedness.

(How L-phosphoserine led to relevant L- $\alpha$ -amino acids is discussed in Section 5).

Before concluding this section I would like to again emphasize that in my proposed scenario, homochirality as exhibited in the living world originated as the result of just one single symmetry-breaking event. This is in stark contrast with previous models in which authors had to account for multitudes of physical and chemical events that would have been required to separate L-amino acids from D-amino acids as well as D-sugars from L-sugars; all those entities presumably coexisting in primordial environments along with myriad of other small and large molecules. I submit that my proposed one-chance symmetry-breaking event is many orders of magnitude more probable.

#### **4. Proposed Ancestral Routes to D-Ribose and to RNAs**

Possible step-pathways that could have led to the early synthesis of the purine and pyrimidine nucleobases were proposed in Section 2. In the present-day living world those nucleobases are linked to a D-ribose-phosphate backbone. How D-ribose (a right-handed pentose) originated, and why the phosphate groups in the polynucleotides occupy the specific 3' and 5' positions, have puzzled scientists for many years. The search for answers has resulted in a great deal of experimental work.

It was initially thought that the early route to ribose could have involved the “classical” formose reaction which is the polymerization of formaldehyde. However this reaction produces a complex mixture that include straight- as well as branched-chained carbohydrates. Further, ribose is obtained in very poor yield, and for such a reaction to occur, high concentrations of formaldehyde are required. For these reasons it was concluded that the “classical” formose reaction is not a plausible route for the prebiotic accumulation of sugars.

More promising results were obtained in the 1990s by scientists who were able to synthesize pentose-2,4-diphosphate from mineral-induced reactions between glycolaldehyde phosphate and racemic glyceraldehyde-2-phosphate (34). Of all the pentoses that this reaction had produced, ribose was a major compound. However, such syntheses of phosphorylated riboses have shortcomings: (i) a racemic mixture of right- and left-handed ribose are produced; (ii) ribose is obtained as a six-membered pyranose rather than a five-membered furanose, and; (iii) the phosphate groups occupy the 2 and 4 positions while, in the living-world, the phosphate groups are attached to positions 3 and 5 of ribose rings.

Continuing with my full-picture scenario, I now suggest an alternative route to phosphorylated D-ribose. Such an ancestral pathway would have involved the cross-aldol condensation between glycolaldehyde phosphate and D-glyceraldehyde-3-phosphate, both entities being constituents of the proposed local microenvironment (Section 1). If D-ribose was preferentially created, it would have required that a stereospecific condensation took place. It is interesting to note that recent studies show stereospecific aldol condensations can be achieved when catalyzed by very short chiral molecules. In one example, glycolaldehyde self-condensation, catalyzed by a chiral dipeptide, produced D-erythrose with an enantiomeric excess greater than 80% (35). In another example, acetaldehyde self-condensation, catalyzed by L-proline, led to a chiral hexenal with an enantiomeric excess of up to 90% (36). With this in mind, I propose a possible pathway that could have led to the asymmetric synthesis of phosphorylated five-carbon entities that started from glycolaldehyde phosphate and D-glyceraldehyde-3-phosphate (D-GAP). These aldol condensations occurring

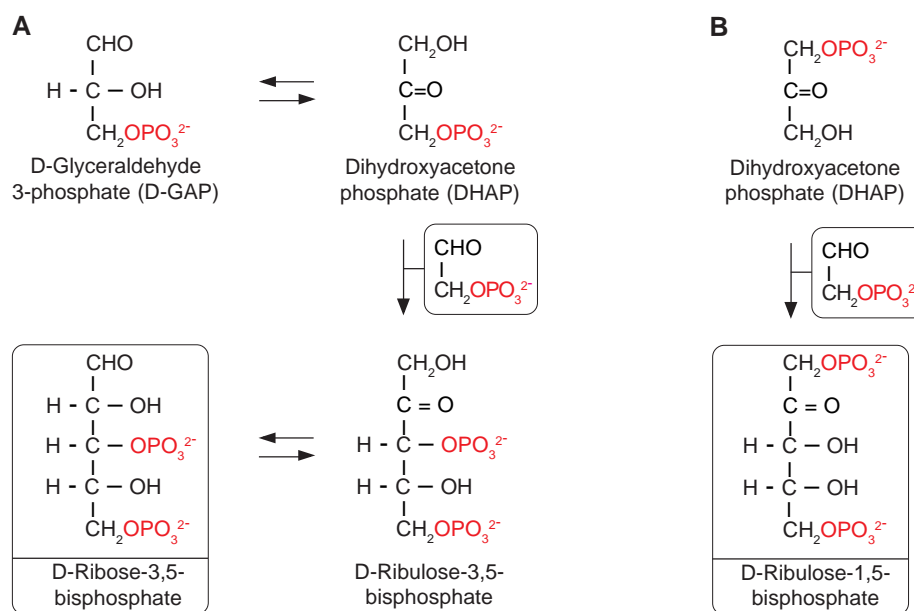


Fig 6. Proposed aldol condensation between glycolaldehyde phosphate and D-glyceraldehyde 3-phosphate (D-GAP) using dihydroxyacetone phosphate (DHAP) as an intermediate.

in the proposed local chiral microenvironment would have been catalyzed by chiral molecules such as the chiral three-carbon entities already in existence in the suggested setting.

As illustrated in Fig.6A and 6B, D-GAP is first isomerized to dihydroxyacetone phosphate (DHAP). Then DHAP reacts with glycolaldehyde phosphate in two different ways:

a) In Fig. 6A, the carbon C3 of DHAP (the one that carries the phosphate group) acts as a nucleophile and attacks the electrophilic carbonyl group of glycolaldehyde phosphate and in the process an asymmetric condensation, catalyzed by preexisting chiral compounds, could have led to ribulose-3,5-biphosphate; then, after tautomerization, to D-ribose-3,5-biphosphate. (Remember, the experiments described above (34) showed that condensation between glyceraldehyde-2-phosphate and glycolaldehyde phosphate produced ribose-2,4-biphosphate as a major compound).

b) In Fig. 6B, the C1 carbon of DHAP (that is not attached to the phosphate group) can also act as the nucleophile that attacks the C=O group of glycolaldehyde phosphate. In this case, an asymmetric condensation could have led to ribulose-1,5-bisphosphate (RuBP).

Before going further, I want to discuss RuBP. In the living-world, RuBP is the key constituent of an enzyme (RuBisCo) responsible for CO<sub>2</sub> fixation. This enzyme, the most abundant and probably the most important on our planet, is found in all three domains of life: Archaea, Eukarya, and Bacteria. RuBisCo, through the Calvin cycle, generates two 3-phosphoglycerates after fixation of CO<sub>2</sub> by RuBP. If RuBP had been produced through an ancestral route similar to the one just proposed, it would imply that a primitive Calvin cycle emerged very early in the prebiotic era.

Let's return to D-ribose-3,5-bisphosphate. If it was synthesized through a route similar to that illustrated in Fig. 6A, and became a constituent of the proposed microenvironment, how could this right-handed pentose have been involved in the creation of nucleotides? Experimentalists first attempted to efficiently bind the nucleobases directly to the ribose rings. However, the results were disappointing. Condensation of purines with ribose rings occurred with low-yield, and there are no known plausible prebiotic condensations of pyrimidines with ribose rings that would produce nucleosides (37).

Alternative routes to pyrimidine nucleosides have also been investigated. Experiments that involved the stepwise assembly of the nitrogenous bases on a sugar, or sugar phosphate, have yielded more promising results. In one experiment, treatment of D-ribose with cyanamide and cyanoacetylene produced  $\alpha$ -ribocytidine (but not  $\beta$ -ribocytidine) (38). In another experiment, the sequential action of cyanamide and cyanoacetylene on arabinose-

3-phosphate produced cytidine-2',3'-cyclophosphate and arabinocytidine-3'-phosphate (39). Even though these experiments didn't lead to the relevant nucleotides, they show that it is possible to build nucleobase rings on sugar phosphate backbones. It is also worth noting that, starting from a phosphorylated ribose, a stepwise assembly takes place in the building of the purine residue in the biosynthetic pathway to inosine monophosphate (IMP, an intermediate in the synthesis of the purine ribonucleotides).

Based on these observations, and utilizing the strategy of reconstruction, I suggest that plausible prebiotic routes to pyrimidine and purine ribonucleotides most likely involved a similar sequential building of the heterocyclic bases on D-ribose-3,5-diphosphate residues. With this in mind, I now propose possible sequences of events that could have led to D- $\beta$ -ribonucleotides.

An ancestral route to pyrimidine ribonucleotides could have started with surface-bonded phosphorylated  $\alpha$ -ribopyranose entities. The bonding with the surface was favorable because of the hydroxyl and phosphate groups that are located at positions 1,2,3,5 on the ribose ring. The surface-bonded ribose ring could have been attached and oriented in such a way that position 1 would have been open to attack by urea. This attack would have led, after inversion of configuration, to  $\beta$ -phosphoribosylurea entities that would have then reacted with 3-phosphoglyceric acid or with 3-phosphoglyceramide to create uracil and cytosine rings as suggested in Section 2, and illustrated in Fig. 2.

The routes leading to the formation of the purine D- $\beta$ -ribonucleotides could have started in a similar way. However, in this case, formamide instead of urea condensed in a first step to form  $\beta$ -ribosylformamide entities. This was followed by the condensation of 2-oxo-3-al-

glyceric acid, or 2-oxo-3-al-glyceramide, and continued with the sequential addition of ammonia, then urea, or ammonia, then formamide, to create guanine and adenine rings as suggested in Section 2, and illustrated in Fig 2.

What makes these proposed routes to the ribonucleotides even more compelling is that they started with the three-carbon precursors described in Section 1, and their assembly would have taken place within the same compartmentalized microenvironment that also favored the creation of D-ribose-3,5-bisphosphate from one of the precursors. These suggested prebiotic processes would have given rise to a pool of  $\beta$ -D-ribonucleotides featuring two phosphate groups at the 3' and 5'-positions of the ribose. One possibility of how those two phosphate groups would have played an important role in favoring 3',5'-phosphodiester linkages between surface-bonded  $\beta$ -D-ribonucleotides is illustrated in Fig. 7. Such attachment events that linked properly aligned nucleotides would have ultimately produced chains of oligoribonucleotides and thereby would have opened the door to the RNA world.

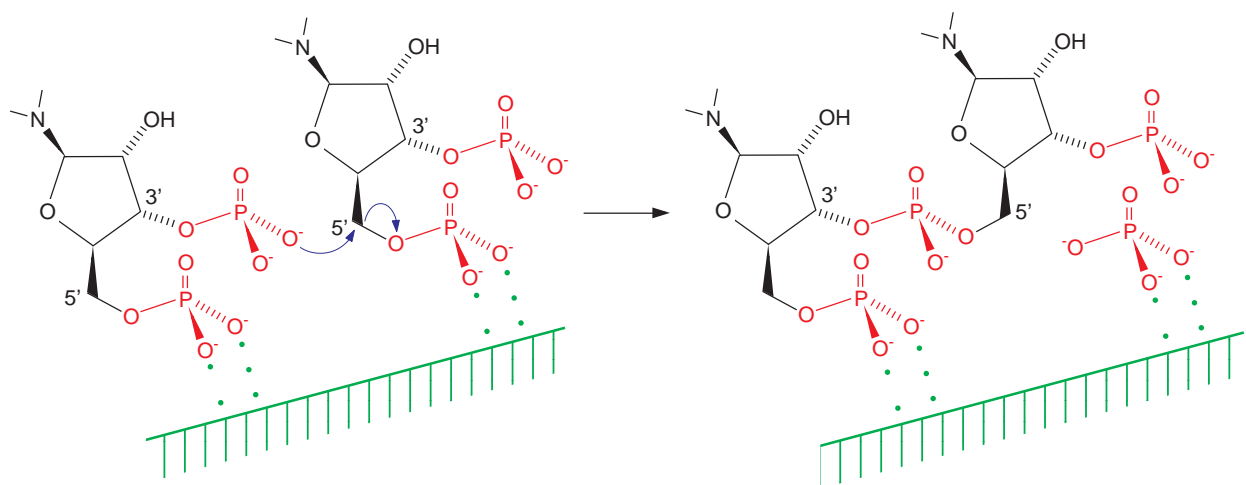


Fig 7. Proposed 3',5' phosphodiester linkage creation between surface-bonded ribonucleotides

## 5. L-Phosphoserine

### The proposed precursor of a whole class of left-handed $\alpha$ -amino acids

The present-day living world incorporates 20 (+2) “standard” L- $\alpha$ -amino acids in proteins. It has been a long-standing challenge to understand why, out of the many possible amino acids, Life uses just 20 (+2), and why these specific 20 (+2). Further, is the necessity to explain why all these  $\alpha$ -amino acids exhibit the same left-handedness.

It is my view that the carbon skeletons of the relevant amino acids were created from a sole precursor, specifically L-phosphoserine. L-phosphoserine itself was created via a  $S_N2$  reaction from D-3-phosphoglyceric acid as discussed in Section 3 and illustrated in Fig. 5.

Before discussing this proposed important role of L-phosphoserine entities, let us examine how the present-day living world synthesizes the “standard” amino acids. The carbon skeletons of the relevant amino acids come from molecular intermediates that are produced during three interrelated biosynthetic pathways: (i) glycolysis; (ii) the citric cycle, and; (iii) the pentose phosphate pathway. The intermediates, presented below in their roles, are the precursors of five families of amino acids:

- 1) The serine-glycine family: Its precursor, 3-phosphoglycerate yields serine, via phosphoserine. Serine gives rise to cysteine and glycine.
- 2) The alanine-valine-leucine family: Its precursor, pyruvate, a three-carbon fragment, yields alanine, valine and leucine.
- 3) The aspartate family: Its precursor, oxaloacetate, a four-carbon fragment, yields aspartate. Aspartate gives rise to asparagine, methionine, lysine and threonine. Threonine leads to isoleucine.

- 4) The glutamate family: Its precursor,  $\alpha$ -ketoglutarate, a five-carbon fragment, yields glutamate. Glutamate gives rise to glutamine, proline and arginine.
- 5) The aromatic amino acids family: Its precursors, phosphoenolpyruvate, a three-carbon fragment, and erythrose-4-phosphate, a four-carbon fragment, yield tyrosine, tryptophan and phenylalanine.

Here are some of the interesting things we learn from these biosynthetic routes: (i) The carbon skeleton of these precursors can be seen as being produced directly or indirectly from the three-carbon skeleton of 3-phosphoglycerate. To be more specific, 3-phosphoglycerate, via the formation of 2-phosphoglycerate and phosphoenolpyruvate, yields pyruvate. Pyruvate, after carboxylation, yields oxaloacetate. Pyruvate after decarboxylation yields acetylCoA which can combine with oxaloacetate to give a five-carbon entity that ultimately leads to  $\alpha$ -ketoglutarate. Finally, a reduced form of 3-phosphoglycerate, namely glyceraldehyde-3-phosphate yields erythrose-4-phosphate. (ii) A few of the amino acids, specifically serine, aspartate, and glutamate, act as precursors for many of the other standard amino acids.

In my proposed model, the carbon skeletons of the early amino acids derived from the three-carbon skeleton of D-3-phosphoglyceric acid (Fig. 1). Note that this is the very same three-carbon precursor that plays a key role in the biosynthetic routes leading to the standard amino acids discussed above. In my model, D-3-phosphoglyceric acid gave rise to L-phosphoserine, the first relevant left-handed  $\alpha$ -amino acid (Fig. 5). Then, through a series of step-pathways, L-phosphoserine entities created a whole class of derivatives that inherited from L-phosphoserine both the specific  $\alpha$ -amino acid group and left-handedness (Figs. 8-9).

To illustrate the plausibility that L-Phosphoserine was the precursor of a whole class of biologically relevant amino acids, I now suggest a sample of possible elongation pathways that could have extended the chain length of L-Phosphoserine entities. Obviously, alternative routes can also be envisioned. However, my proposals are based in part on metabolic pathways I consider to be relics of archaic routes that take place in cyanorganisms. In these suggested elongation processes, cyanations play a major role with HCN, or  $\text{CN}^-$  acting as a source of one-carbon units as well as one-nitrogen units. This essential role of cyanation is illustrated in many experiments involving HCN, or its oligomers, that yield an impressive number of amino acids (40).

In my model, the amino acids that derived from L- phosphoserine can be seen as members of three different families. I call them: (i) The Serine family; (ii) The Cyanoalanine family, and; (iii) The Cyanohomoalanine family. (I want to point out that the step-pathways to the amino acids illustrated in Figures 8 and 9 didn't necessarily happen in the same epoch. Although some members of these families would have formed at the early stages, others could have been transformed from intermediates to their contemporary forms as late-comers).

**The Serine Family:** This proposed family is generated directly from L-phosphoserine in the possible following ways: *L-serine*, *L-cysteine* and *L-selenocysteine* can be created through the displacement of the phosphate group by OH, SH and SeH respectively. *L-Alanine* can be created from L-serine through reduction of the OH group. In modern biology *glycine* is synthesized from L-Serine. However, as will be discussed later, *glycine* can be created from *L-threonine* through a retroaldol condensation.

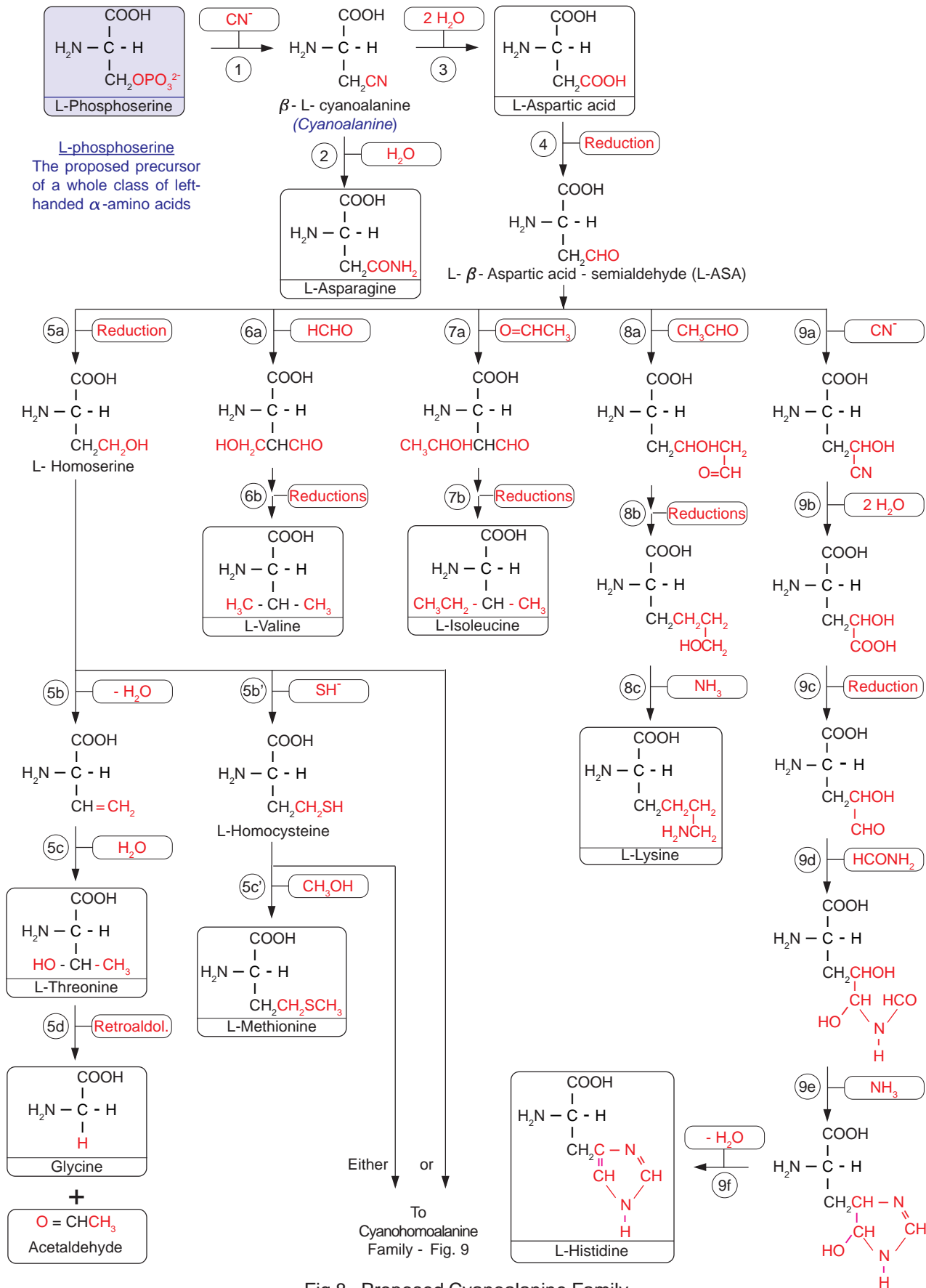


Fig 8. Proposed Cyanoalanine Family

**The Cyanoalanine Family** (Fig. 8): In my scenario, members of this proposed family derive from  $\beta$ -L-cyanoalanine (L-AlaCN). In the living-world,  $\beta$ -L-cyanoalanine (L-AlaCN), a four-carbon nitrile, is found in a wide spectrum of species among bacteria (41), plants (42), and insects (43). It is produced through reactions of HCN with L-cysteine or L-O-acetyl serine, both derivatives of L-serine. In many species that include, for example, higher plants, *L-asparagine* and *L-aspartic acid* are produced by hydrolysis of  $\beta$ -L-cyanoalanine (44).

In my proposed model, L-phosphoserine is converted to  $\beta$ -L-cyanoalanine after a nucleophilic attack by cyanide ions (Fig. 8, step 1). Hydrolysis of this four-carbon nitrile produces *L-asparagine* (Fig. 8, step 2) and *L-aspartic acid* (Fig. 8, step 3). A one-step reduction of *L-aspartic acid* gives rise to L- $\beta$ -aspartic acid-semialdehyde (Fig. 8, step 4). This four-carbon aldehyde will lead to a large group of left-handed  $\alpha$ -amino acids (Fig. 8, steps 5 through 9):

Fig 8 step pathways 5a to 5c': A one-step reduction of the aldehyde function of L- $\beta$ -aspartic acid-semialdehyde (L-ASA) gives L-homoserine (Fig. 8, step 5a). *L-threonine* is created from L-homoserine through dehydration, then rehydration, of L-homoserine (Fig. 8, step 5 and 5c). *Glycine* and acetaldehyde are created from L-threonine through retroaldol condensation (Fig. 8, step 5d). L-homocysteine is created from L-homoserine through an attack by SH<sup>-</sup> (Fig. 8, step 5b'). *L-methionine* is created following the reaction of methanol with L-homocysteine (Fig. 8, step 5c'). (In my model, L-homocysteine, or an activated L-homoserine, are the precursors of a large class of amino acids belonging to what I call the "Cyanohomoalanine family," illustrated in Fig. 9, to be discussed later.)

Fig 8, step pathways 6a and 6b: Aldol condensation of L-ASA acting as a donor with formaldehyde creates an intermediate that after several step-reductions produces *L-valine*.

Fig 8, step pathways 7a and 7b: L-ASA acting as a donor condenses with acetaldehyde to give an intermediate that can be further reduced to *L-iso-leucine*.

Fig 8, step pathways 8a to 8c: Acetaldehyde acting as a donor condenses with L-ASA to produce an aldehyde intermediate that is reduced to an alcohol which reacts with ammonia to give *L-lysine*.

Fig 8, step pathways 9a to 9g: In a first step,  $\text{CN}^-$  condenses on the carbonyl group of L-ASA to give an  $\alpha$ -hydroxynitrile that after hydrolysis, then a reduction, gives an aldehyde intermediate. This condenses with formamide, then ammonia, to produce an imidazole ring. Dehydration then leads to a double bond formation that ultimately yields *L-histidine*. These proposed steps are similar to my suggested routes to purine rings (Fig. 2).

**The Cyanohomoalanine Family** (Fig. 9): Members of this proposed family derive from  $\gamma$ -cyano- $\alpha$ -aminobutyric acid. This five-carbon nitrile is a homologue to the four-carbon nitrile  $\beta$ -cyanoalanine. For that reason, and to be consistent, I coined the name “cyanohomoalanine” (L-HalaCN) to refer to  $\gamma$ -cyano- $\alpha$ -aminobutyric acid.

Let us examine the steps-pathways that in my view could have led to members of the proposed cyanohomoalanine family. First let's consider relevant pathways found in the living world. In organisms such as *C. violaceum*, cyanohomoalanine is synthesized from homocystine and cyanide (45). It is produced from O-acetyl-L-homoserine and cyanide by *Bacillus stearothermophilus* a thermophilic bacterium (46). Notice that this last pathway shares similarities with the route to  $\beta$ -cyanoalanine from O-acetyl-L-serine and cyanide.

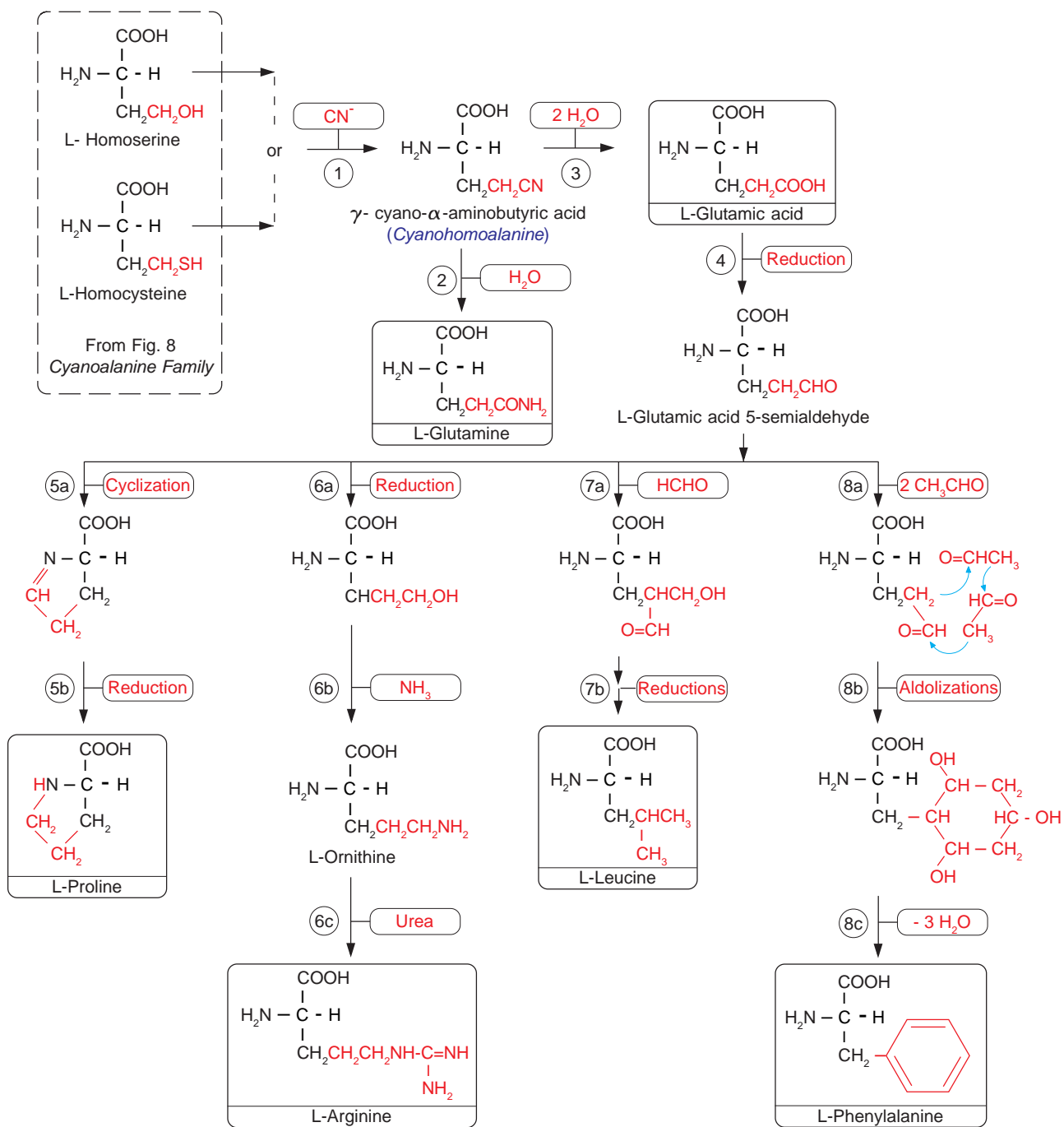


Fig 9 Proposed Cyanohomocysteine Family

Experimentalists have produced *L-glutamine* and *L-glutamic acid* from cyanohomoalanine ( $\gamma$ -cyano- $\alpha$ -aminobutyric acid) through hydrolysis (46).

In my model, the five-carbon nitrile, cyanohomoalanine ( $\gamma$ -cyano- $\alpha$ -aminobutyric acid) is produced from  $\text{CN}^-$  reacting with either of two derivatives of the cyanoalanine family: L-homocysteine or L-homoserine (Fig. 9, step 1). Hydrolysis of  $\gamma$ -cyano- $\alpha$ -aminobutyric acid gives *L-glutamine* (Fig. 9, step 2) and *L-glutamic acid* (Fig. 9, step 3). *L-glutamic acid* is reduced to L-glutamic acid 5-semialdehyde (L-GSA) (Fig 9, step 4). This five-carbon aldehyde will lead to a large group of left-handed  $\alpha$ -amino acids (Fig. 9, steps 5 through 8):

Fig 9, steps 5a and 5b: L-GSA forms an internal Schiff base and produces a cyclic intermediate that creates *L-proline* after a double-bond formation.

Fig 9, steps 6a to 6c: L-GSA is reduced to an alcohol that condenses with ammonia to give L-ornithine (Fig 9, steps 6a and 6b). Condensation of L-ornithine with urea produces *L-arginine* (Fig. 9, step 6c).

Fig 9, steps 7a and 7b: Aldol condensation between L-GSA (acting as a donor) with formaldehyde gives an intermediate that is further reduced to *L-leucine*.

Fig 9, steps 8a to 8c: Sequential aldol condensations involving two acetaldehyde (acting as donors) and L-GSA (acting as acceptors) produces a hydroxylated cyclic intermediate that, after formation of a benzene ring, turns into *L-phenylalanine*.

The chemical step-pathways just proposed illustrate how L-phosphoserine could have been the sole precursor of a whole class of left-handed- $\alpha$ -amino acids, some of which will be utilized as the building blocks of proteins. In support of this, similar pathways have been shown to exist in microorganisms that use  $\text{CN}^-$  as a source of carbon and nitrogen. Is it

possible that these present-day pathways found in cyanoorganisms are relics of ancestral routes? Present-day aspartate and glutamate families also have noticeable similarities to the suggested cyanoalanine and cyanohomoalanine families. Can we infer from this that modern-day biology mimics, and thereby maintains, the step-pathways that were efficiently at work in early metabolisms?

The possible location for the early conversions of L-phosphoserine:

Based on theoretical and experimental data, one can envision that the conversions due to the modification of side chains of L-phosphoserines took place while L-phosphoserine entities were attached to the 3' terminus of primitive tRNA structures. Contemporary tRNAs are used for converting precursor amino acids while the precursors are attached to their 3' end. This process is responsible, for example, for the conversion of aspartic acid and glutamic acid to asparagine and glutamine respectively (47). Recent data show that certain methane-producing Archaea use a tRNA-dependant pathway for the conversion of lysine into pyrrolysine which represents the 22<sup>nd</sup> genetically encoded natural amino acid (48). Important additional support for my model is found in two other examples of contemporary tRNA-dependent amino acid syntheses: (i) Selenocysteine, known as the 21<sup>st</sup> amino acid, is created from tRNA-bonded with L-phosphoserine. Phylogenetic arguments suggest that this pathway was present in the last common ancestor (49). (ii) In Methanogenic Archaea, cysteine is produced from a route involving a phosphoserine-loaded tRNA. In these methanogens this pathway can act as the sole route to cysteine (50). Interestingly, L-phosphoserine, the precursor of cysteine in this route, originates from a pathway that starts with D-3-phosphoglycerate (51). This is similar to the pathway proposed in my scenario in which the right-handed

D-3-phosphoglycerate is transformed to L-phosphoserine that is then converted to L-cysteine. I believe that these tRNA-dependant conversions of L-phosphoserine entities are vestigial remains of ancestral routes.

I further propose that the very processes that yielded the L-phosphoserine conversions concomitantly gave rise to the Genetic Code. Such chemical transformations could have taken place in the ‘late’ RNA world and would have involved ribozyme-based metabolism. Experimental evidence tends to support the plausible existence of early metabolism pathways catalyzed by ribozymes (52). (How the genetic code could have originated in this context will be discussed in another article).

Notice that in my suggested scenario, the relevant left-handed  $\alpha$ -amino acids do not originate as “free” and separate entities. When created from L-phosphoserine, as I propose, they are already attached to the 3’ end of their respective primitive tRNA structure. Consequently, I suggest that, as it does in the present-day living world, the early peptide bond formations involved assemblies of aligned tRNAs that would have brought their attached amino acids into close proximity, thus allowing them to condense. Such events would have opened the door to the Protein World, overlapping the RNA World.

## **6. Proposed Ancestral Routes to Lipids**

The cell membranes of Eukarya and Bacteria are composed of glycerol residues attached via ester or ether linkages to straight-chain fatty derivatives usually featuring even-number carbon skeletons. By contrast, the cell membranes of Archaea are composed of polyisoprenoids glycerol ethers. How straight-chain derivatives, and isoprenoid structures,

could have originated before the emergence of enzyme-catalyzed synthetic pathways has been the focus of only few experimentalists and theorists. Let us first focus on some of these experiments that attempted to synthesize straight-chain fatty derivatives under assumed prebiotic conditions.

Conducted under high pressures and high temperatures, experiments known as Fisher-Tropsch-type (FTT) reactions, attempted to simulate hydrothermal vents. Typically, these reactions were carried out for two to three days in stainless steel vessels at 175°C using very high concentrations of formic acid or oxalic acid as precursor carbon sources. They led to very complex mixtures of oxygenated and non-oxygenated products showing no predominance for even-number carbon skeletons. The yields of synthesized compounds having chain lengths greater than 10 were very low (less than 0.1 % of the starting carbon substrate) (53). It seems improbable that FTT reactions model the conditions of the prebiotic world because they require extreme conditions of pressure and temperature; unreasonably high concentrations of the starting materials; the products are very complex mixtures, and; they don't yield the correct biomolecules.

An alternative route to the straight-chain derivatives that I now propose, involves pathways using aldol condensation as an efficient way to produce successive C-C bonds for linear-chain formation. This is consistent with the chemical processes that are brought to light in the full-picture model developed in this paper. Adding plausibility, my suggested routes for the linear-chain formations take place under mild conditions.

I suggest that polyunsaturated linear aldehydes of the general formula,  $\text{CH}_3(\text{CH}=\text{CH})_n\text{-CHO}$ , were the source of the first straight-chains fatty derivatives that were

involved in the formation of early lipids. Before being more specific, I want to mention that polyunsaturated linear aldehydes, known as polyenals, have been found in the living world. Experimental studies of many species of parrots have shown that polyenals, such as tetradecahexenal ( $C_{14}$ ), hexadecaheptenal ( $C_{16}$ ), octadecaoctenal ( $C_{18}$ ) and eicosanonenal ( $C_{20}$ ), are used as red pigments in their feathers (54). Authors supposed that the biosynthetic pathways leading to polyenals might involve  $C_2$  (acetate) units. However, the mechanism of such syntheses are not yet really well understood (55). Is it possible that the polyenals found in the parrots are vestigial remains of ancient pathways? If so, what prebiotic processes could have yielded polyunsaturated linear aldehydes such as these?

To understand this, let us first examine some experimental data. It has been shown that when acetaldehyde adsorbs on  $TiO_2$ , lattice oxygen, acting as a base, catalyzes a sequential aldol condensation process. As a result of this condensation, high molecular weight compounds containing conjugated  $C=C$  bonds are produced (56). Another known direct way to produce polyenals is through sequential aldol condensations of  $C_4$  units brought by crotonaldehyde. Crotonaldehyde can be created from a base-catalyzed aldol condensation between two acetaldehyde molecules. This process has also been shown to take place in living-systems. For example, results of one study indicate that a polyamine, such as spermidine, directly reacts with acetaldehyde to generate crotonaldehyde, most likely via an enamine aldol condensation mechanism (57). Crotonaldehyde can self-condense to produce a series of polyenals. Such a process was first described in 1937 by organic chemists who used piperidine, a secondary amine, as a catalyst (58). As illustrated in Fig. 10A and 10B, sequential  $\gamma$ -aldolizations can explain the formation of  $C_8$ ,  $C_{12}$ , and  $C_{16}$  polyenals.

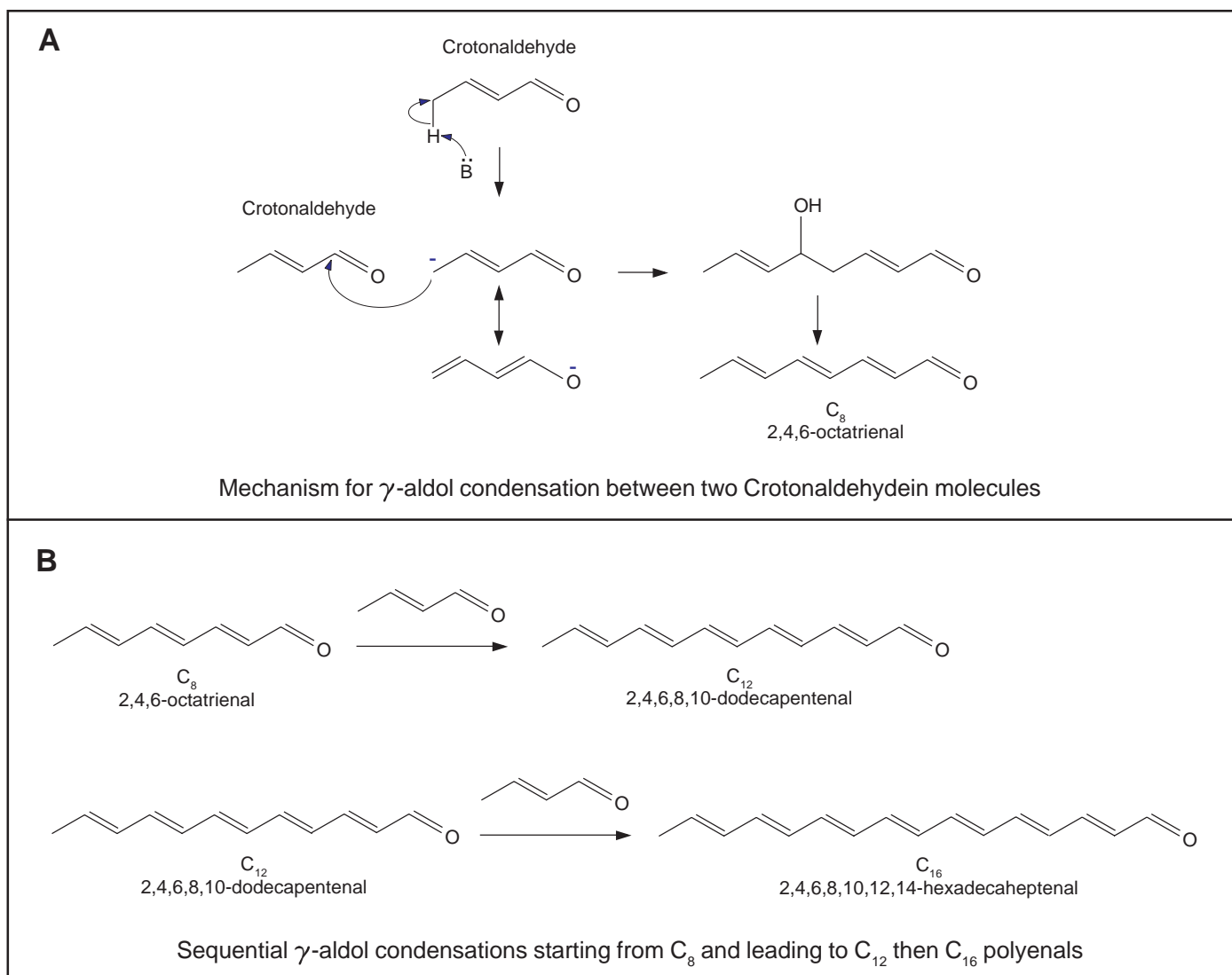


Fig. 10

Condensation of these last compounds with acetaldehyde ( $\text{C}_2$ ) units will produce  $\text{C}_{10}$ ,  $\text{C}_{14}$ , and  $\text{C}_{18}$  polyenals.

If ancient metabolism pathways led to polyenals via sequential aldol condensations of acetaldehyde or crotonaldehyde units it would have first required that acetaldehyde was a prevalent component of the early Earth. Acetaldehyde can be created through several different

pathways. For example, studies show that acetaldehyde is produced after thermal degradation of serine being decarboxylated, then deaminated (59). Acetaldehyde can also be produced by decarboxylation of pyruvic acid, an  $\alpha$ -keto acid. Pyruvic acid has also been identified among products of thermal degradation of serine (59).

Assuming that acetaldehyde units led to the formation of a series of polyunsaturated linear aldehydes with even-carbon numbers as described above, what would have been the fate of those straight-chain compounds? Early polyenals could have provided the carbon skeleton for the formation of saturated or unsaturated linear compounds such as fatty aldehydes, fatty acids, and fatty alcohols. Fatty aldehyde could have condensed to a phosphorylated glycerol via a vinyl ether linkage to give rise to early plasmalogen lipids. Fatty acid or fatty alcohol, if available, would condense via ester or ether linkages. It is worth mentioning that experimental data shows that, under prebiotic conditions, condensations of glycerol with fatty aldehydes give better results than condensations with fatty acids, and fatty alcohols form little, if any, glycerides (60). Ultimately, fatty derivatives linked to a glycerol backbone would have led to the creation of protocell membranes in Bacteria and Eukaria.

Let's compare the biosynthetic pathways to straight-chain derivatives with my suggested pathways. In the living world, fatty acids derive from sequential condensation of  $C_2$  units brought by acetate (acetylCoA). In my suggested pathway, I propose that acetate (acetylCoA) replaced the  $C_2$  units that were brought, in early metabolisms, by acetaldehyde. It's worth mentioning that, in the living world, acetylCoA derives directly from acetaldehyde.

Let us now focus on isoprenoids. What was the primeval C<sub>5</sub> unit that was used to build those early compounds, and what kind of mechanism was responsible for the chain-elongation processes? In attempts to answer these questions some researchers suggested that an archaic C<sub>5</sub> unit of isoprenoid was created via an acid-catalyzed Prins reaction from the condensation of isobutene with formaldehyde (61). According to those authors, such a reaction would have led to isopentenol entities that, after being attached to a phosphate mineral surface, would have produced polyprenol phosphates. However, the validity of this suggested synthesis has been questioned because isobutene is not considered a plausible prebiotic reactant (62).

What I propose for the primeval C<sub>5</sub> unit is 3-methyl-2-butenal, also known as methylcrotonaldehyde, or prenal. This is based in part on retrosynthetic analysis of some compounds belonging to the isoprenoid/carotenoid families, specifically *all-trans* dehydrocitral and 3-dehydroretinal. The first compound has been found, for example, in the leaf oil of *Ambrosia confertiflora* (63), and in mites (64). The second compound is a chromophore found, for example, in the eyes of various species of fresh-water crayfish (65). Using retrosynthetic analysis on dehydrocitral and 3-dehydroretinal, it is possible to conclude that those two compounds can be built from the respective dimer and tetramer of prenal, my proposed C<sub>5</sub> unit. It is worth mentioning that prenal is produced as a biogenic compound in, for example, root material of wild ginger (66) and wild varieties of blackberries (67).

It is my view that isoprenoid aldehydes such as *all-trans* dehydrocitral and 3-dehydroretinal are molecular fossils of an early time when prenal entities were building-block units that led to the creation of the early carotenoids. If indeed prenal residues played such roles, what kind of mechanisms can explain the oligomerization of these C<sub>5</sub> units?

In 1935, chemists showed that base-catalyzed, successive condensations of prenal led to the formation, in good yield, of dehydrocitral, and to the creation of the C<sub>15</sub> aldehyde 3,7,11-trimethyl-2,4,6,8,10-dodecapental (tetrahydrofarnesal) in moderate yield. Small amounts of an aldehyde of higher molecular weight (red in color) have also been detected. This compound was most likely the acyclic C<sub>20</sub> aldehyde, 3,7,11,14-tetramethyl-2,4,6,8,10,12,14-hexadecaheptenal (dehydroacycloretinal) (68). As illustrated in Fig. 11A and 11B, such syntheses are the result of sequential  $\gamma$ -aldol condensations of prenal units.

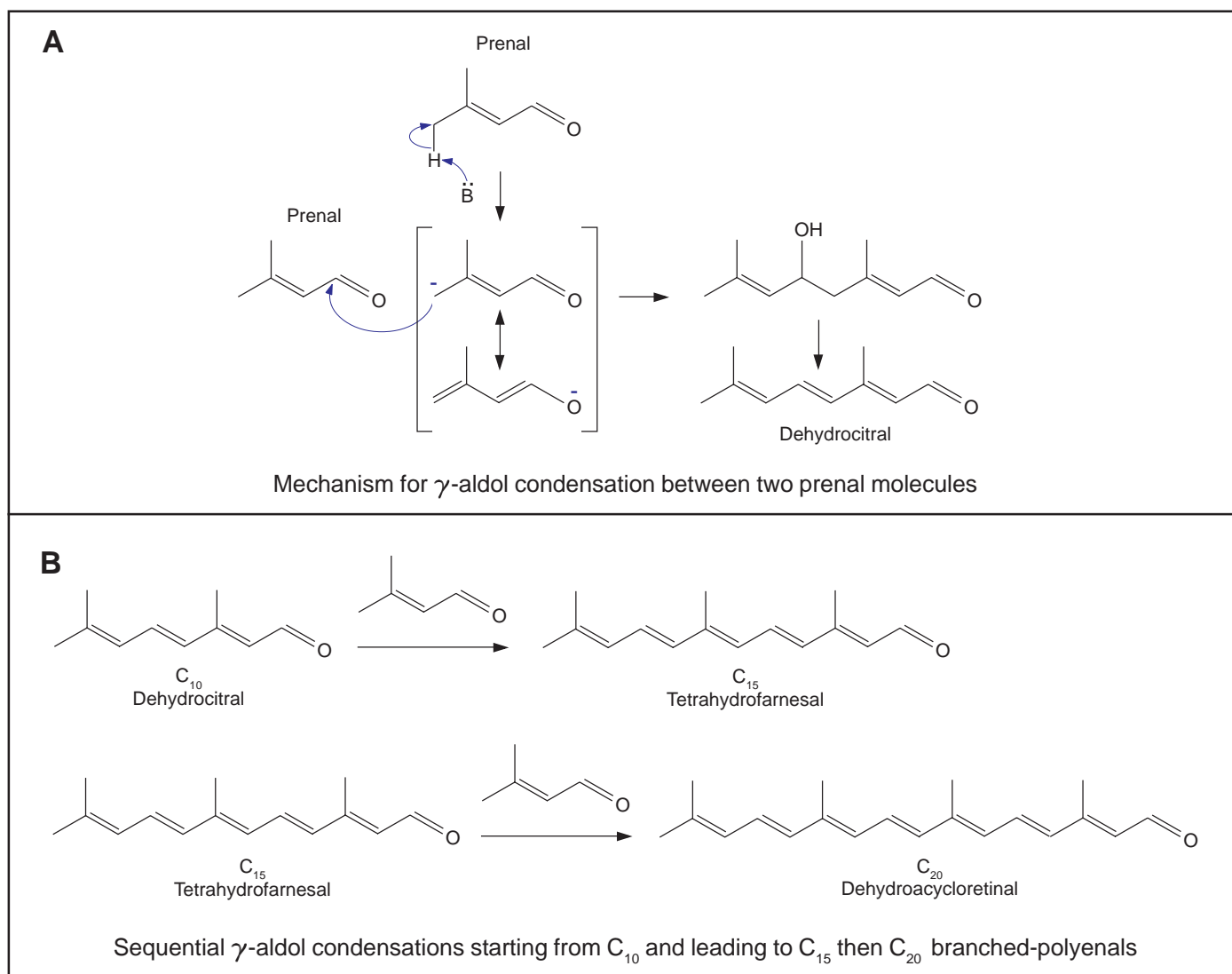


Fig. 11

This mechanism is quite similar to the one responsible for self-condensation of crotonaldehyde discussed above. Reductive coupling of dehydroacetylcrotonal that can form from self-condensation of four prenal units could lead to bisdehydrolycopene. This C<sub>40</sub> compound is a natural carotenoid that has been found, for example, in Valencia oranges (69).

Assuming that prenal was the early isoprenoid C<sub>5</sub> unit, an important question that now arises is how the prenal entities could have formed in the first place. One obvious direct route seems to be the cross-aldol condensation between acetone, acting as an acceptor, and acetaldehyde, acting a donor. However such a reaction is not favorable because acetone would have been the donor, and acetaldehyde the acceptor.

Another route to prenal, that I now suggest, may have involved the cross-aldol condensation of pyruvic acid, acting as donor, and acetone acting as acceptor. The first step produces a six-carbon intermediate (CH<sub>3</sub>)<sub>2</sub>C(OH)CH<sub>2</sub>COCOOH. Decarboxylation of this α-keto acid can give rise to (CH<sub>3</sub>)<sub>2</sub>C(OH)CH<sub>2</sub>CHO which, after dehydration, yields prenal. The acetone, which is required to condense with pyruvic acid, can be produced from two acetaldehyde units through different routes. Self condensation of two acetaldehydes leads to a four-carbon aldol, CH<sub>3</sub>CH(OH)CH<sub>2</sub>CHO, which can be oxidized to acetoacetic acid. Decarboxylation of this last unstable compound creates acetone. An alternative route to acetone can also involve the four-carbon aldol, CH<sub>3</sub>CH(OH)CH<sub>2</sub>CHO. This aldol can be converted via H-transfer to the keto form CH<sub>3</sub>COCH<sub>2</sub>CH<sub>2</sub>OH which, after reverse aldol condensation, can produce acetone and formaldehyde. Acetone can also be produced from decarboxylation/deamination of threonine.

If we accept that prenal entities were part of early metabolisms, then, besides being the building-block units of early retinoids/carotenoids, such as, for example, dehydrocitrinal, acycloretinal, or 3-dehydroretinal, they could have led to a series of branched-chain polyenals. Derivatives of such polyunsaturated aldehydes could have been attached via ether linkages to sn-glycerol-1-phosphate, thus creating protocell membranes in Archaea.

Making the sequence of events suggested in this section compelling is that the described chemical steps are completely compatible, and in concert, with the other reactions that take place within the local microenvironment that is a part of the full-picture model presented in this paper. In this ‘one-pot’ reaction system, molecules such as acetaldehyde, acetone and pyruvic acid would have formed *in situ* from precursors such as serine, or glyceric acid, and their interactions would have opened the door to the Lipid World, overlapping the Protein and RNA Worlds.

## Conclusion

The author has presented a full-picture model of the chemical pathways and processes that gave rise to the biologically relevant nucleobases, D-sugars, L- $\alpha$ -amino acids, straight-chain fatty acids of even-number carbons, and branched-chain isoprenoid compounds. The model also accounts for the origin of biological homochirality. The proposed scenario takes place in a local sheltered micro environment of early Earth. It starts with a reaction between cyanide ions and glycolaldehyde phosphate entities that produce an enantiopure right-handed, three-carbon cyanohydrin that was most likely the result of an asymmetric amplification that propagated from a small initial imbalance. The author has suggested that this chiral,

three-carbon cyanohydrin, acting as ‘Premier Precursor’ leads to a set of chiral, three-carbon derivatives that, in the proposed scenerio, further lead to relevant biofamilies: (i) The three-carbon, D-glyceraldehyde-3-phosphate, yields D-sugars (such as D-ribose-3,5-bisphosphate); (ii) the three-carbon, D-3-phosphoglyceric acid leads to L-phosphoserine (L-Sep), with the creation of the amino group occurring with a chiral inversion. L-phosphoserine (L-Sep) then acts as a precursor to a whole class of relevant left-handed  $\alpha$ -amino acids; (iii) the three-carbon, D-glyceraldehyde-3-phosphate gives rise to D-glycerol-3-phosphate which then acts as a backbone of lipids; (iv) a group of three-carbon derivatives that originate from the ‘Premier Precursor’ plays an important part in the building of nucleobases, U, C, G, and A. Continuing interrelated cascading chemical events that occur within the proposed ‘one-pot’ reaction system open the door to a preRNA/Phosphoserine World in which primitive tRNAs play a major role in converting L-Seps into a variety of L- $\alpha$ -amino acids while starting the build-up of the genetic code\*. This ultimately yields all of the entities that give rise to the RNA/Protein/Lipid World.

In summary, Figure 12 depicts the proposed ancestral routes to the relevant biofamilies. It is the author’s hope that this model provides new insights into our understanding of how, billions of years ago, living systems emerged in a prebiotic environment from just a few small molecules.

\* The author will propose a scenario for the emergence of the genetic code in a separate article.

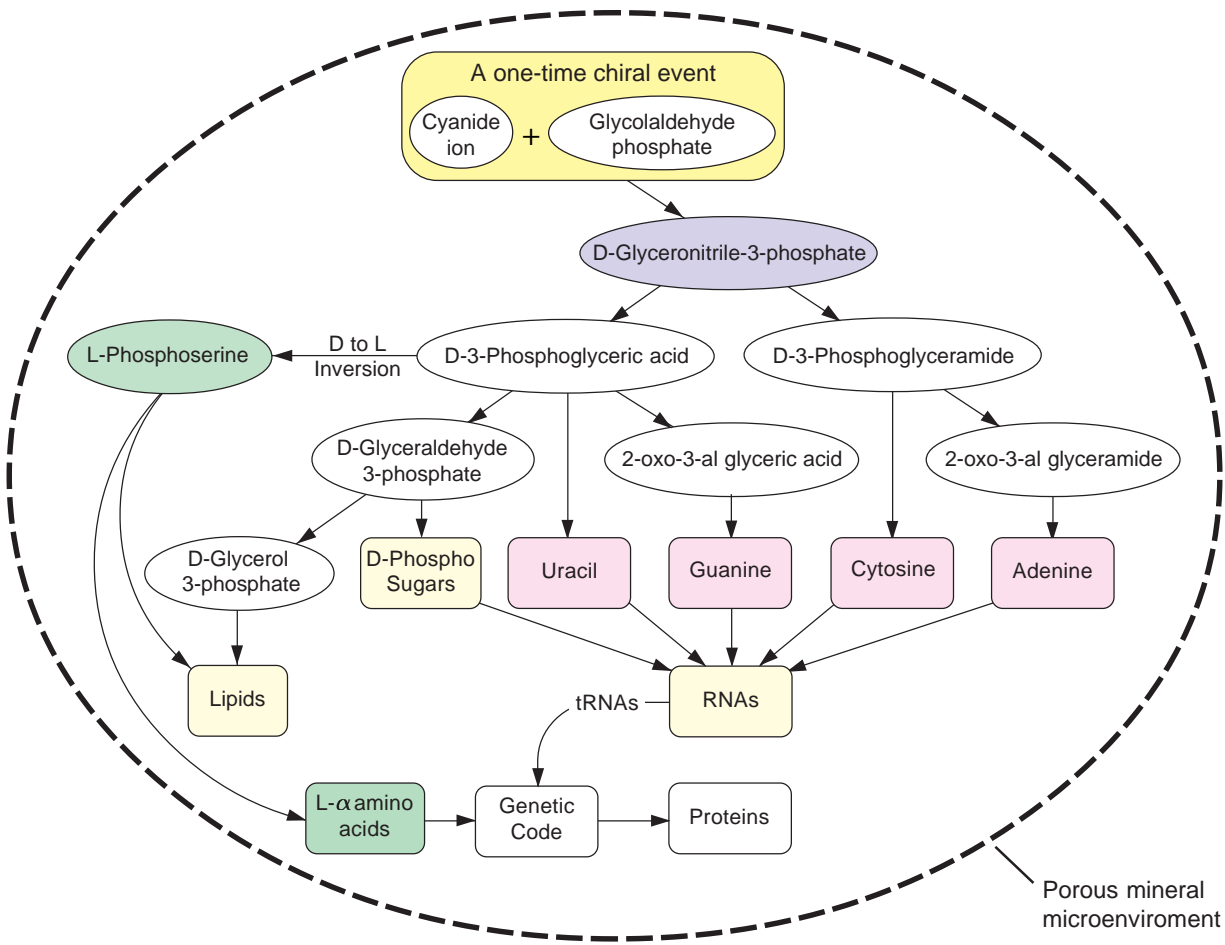


Fig. 12. Summary of proposed early chemical pathways that yielded the RNA/Protein/Lipid World.

## References

1. H. Kiliani, *Ber* **20**, 339 (1887)
2. A. S. Seriani, J. Pierce, R. Baker, *Biochemistry* **18**, 1192 (1979)
3. P. A. Gerakines, M. H. Mooores, R. J. Hudson, *Icarus* **170**, 202 (2004)
4. C. F. Chyba *et al.*, *Science* **249**, 366 (1990)
5. L. M. Mukhin, *Nature* **251**, 50 (1974)
6. J. P. Ferris *et al.*, *J. Mol. Evol.* **11**, 293 (1978)
7. C. Huber, G. Wächtershäuser, *Science* **314**, 630 (2006)
8. J. M. Hollis, F. J. Lovas, P. R. Jewell, *Astrophysical J. Lett.* **540**, 107 (2000)
9. A. S. Seriani, E. L. Clark, R. Baker, *Carbohydrate Res.* **72**, 79 (1979)
10. R. Krishnamurthy, G. O. Arrhenius, A. Eschenmoser, *Orig. Life Evol. Biosphere* **29**, 333 (1999)
11. G. Costanzo *et al.*, *J. Biol. Chem.* **282**, 16729 (2007)
12. R. Lohmann, L. E. Orgel, *Science* **171**, 490 (1971)
13. N. Lahav, *Heterogeneous Chemical Review* **1**, 159 (1994)
14. M. Schoonen, A. Smirnov, C. Cohn, *Ambio* **33**, 539 (2004)
15. G. O. Arrhenius, *Helv. Chim. Acta.* **86**, 1569 (2003)
16. I. Parson, M. R. Lee, J. V. Smith, *Proc. Natl. Acad. Sci. USA* **95**, 15163 (1998)
17. J. P. Ferris *et al.*, *Orig. Life Evol. Biosphere* **5**, 153 (1974)
18. M. P. Roberston, S. L. Miller, *Nature* **375**, 772 (1995)
19. L. E. Orgel, *Orig. Life Evol. Biosphere* **32**, 279 (2002)
20. J. Oro, *Biochem. Biophys. Res. Commun.* **2**, 407 (1960)
21. L. E. Orgel, *Crit. Rev. Biochem. Mol. Biol.* **39**, 99 (2004)
22. A. W. Schartz, H. Joosten, A. B. Voet, *Biosystems* **15**, 191 (1982)
23. G. E. Folsome *et al.*, *Nature* **232**, 108 (1971)
24. S. J. Mojzsis, R. Krishnamurthy, G. Arrhenius, *The RNA World, Second Edition Cold Spring Harbor Laboratory Press*, Chapt. 1 p 15 (1999)

25. M. North, *Synlett* 807 (1993)
26. J.-M. Brunel, I. P. Holmes, *Angew. Chem. Int. Ed.* **43**, 2752 (2004)
27. W. R. Jackson *et al.*, *Tetrahedron Letts.* **31**, 1447 (1990)
28. T. Ziegler, B. Horsch, F. Effenberger, *Synthesis* 575 (1990)
29. F. Effenberger, U. Stelzer, *Angew. Chem. Int. Ed.* **30**, 873 (1991)
30. K. Soai *et al.*, *Nature* **378**, 767 (1995)
31. D. G. Blackmond, *PNAS* **101**, 5732 (2004)
32. E. I. Klabunovskii, *Astrobiology* **1**, 127 (2001)
33. R. M. Hazen, D. S. Sholl, *Nature Materials* **2**, 367 (2003).
34. R. Krishnamurthy, S. Pitsch, A. Eschenmoser, *Orig. Life Evol. Biosphere* **29**, 129 (1999)
35. A. L. Weber, S. Pizzarello, *PNAS* **103**, 12713 (2006)
36. A. Cordova, W. Notz, C. F. Barbas, *J. Org. Chem.* **67**, 301 (2001)
37. G. Joyce, L. E. Orgel, *The RNA World, Second Edition, Cold Spring Harbor Laboratory Press*, p 68 (1999)
38. R. A. Sanchez, L. E. Orgel, *J. Mol. Biol.* **47**, 531 (1970)
39. A.-A. Ingar *et al.*, *ChemBioChem* **4**, 504 (2003)
40. J. P. Ferris, W. J. Hagan, Jr, *Tetrahedron* **40**, 1093 (1984)
41. A. McAdam, C. Knowles, *Biochem Biophys Acta.* **786**, 123 (1984)
42. H. Hendrickson, E. Conn, *J. Biol. Chem.* **244**, 2632 (1969)
43. D. Meyers, S. Ahmad, *Biochem Biophys Acta.* **1075**, (1991)
44. M. Piotrowski, S. Sabine, E. W. Weiler, *J. Biol. Chem.* **276**, 2616 (2001)
45. C. Ressler *et al.*, *Biochemistry* **12**, 5369 (1973)
46. H. Omura *et al.*, *J. Biosci. Bioeng.* **96**, 53 (2003)
47. G. Raczniak *et al.*, *J. Biol. Chem.* **276**, 45862 (2001)
48. G. Srinivasan, C. M. James, J. A. Krzycki *Science* **296**, 1459 (2002)
49. J. Yuan *et al.*, *PNAS* **103**, 18923 (2006)

50. A. Sauerwald *et al.*, *Science* **307**, 1969 (2005)
51. S. Helgadottir *et al.*, *J. Bacteriol.* **189**, 575 (2007)
52. X. Chen, N. Li, A. D. Ellington, *Chemistry & Biodiversity* **4**, 633 (2007)
53. T. M. McCollom, G. Ritter, B. R. T. Simonet, *Orig. Life Evol. Biosphere* **29**, 153 (1999)
54. K. J. McGraw, M. C. Nogare, *Biol. Lett.* **1**, 38 (2004)
55. R. Stradi, E. Pini, G. Celentano, *Comp. Biochem. Physiol. B* **130**, 57 (2001)
56. S. Luo, J. L. Falconer, *Catal. Lett.* **57**, 89 (1999)
57. J. A. Theruvathu *et al.*, *Nucleic Acids Res.* **33**, 3513 (2005)
58. R. Kuhn, C. Grundmann, *Chem. Ber.* **70**, 1318 (1937)
59. V. A. Yaylayan, A. Keyhani, A. Wnorowski, *J. Agric. Food Chem.* **48**, 636 (2000)
60. W. R. Hargreaves, S. J. Mulvihill, D. W. Deamer, *Nature* **266**, 78 (1977)
61. G. Ourisson, Y. Nakatani, *Chem. Biol.* **1**, 11 (1999)
62. D. W. Deamer, *Microbiol. Mol. Biol. Rev.* **61**, 239 (1997)
63. J. Kumamoto, R. W. Scora, W. W. Payne, *J. Agric. Food Chem.* **23**, 123 (1975)
64. H. Hiroyuki *et al.*, *Biosci. Biotechnol. Biochem.* **65**, 2749 (2001)
65. T. Susuki, E. Eguchi, *Cell. Molec. Life Sciences* **43**, 1111 (1987)
66. M. G. Motto, N. J. Secord, *J. Agric. Food Chem.* **33**, 789 (1985)
67. D. N. Georgilopoulos, A. N. Gallois, *Z. Lebensm. Unters. Forsch.* **185**, 299 (1987).
68. F. G. Fisher, K. Hultsch, *Ber.* **68**, 1726 (1935)
69. A. Winterstein, A. Struder, R. Ruder, *Chem. Ber.* **93**, 2951 (1960)
70. Acknowledgement: I want to express my deepest gratitude to William M. Simon for his contributions in designing all the graphics, helping me find the right words and sentences, and also for his continuous support and encouragement.