

Review

The Stereochemical Basis of the Genetic Code and the (Mostly) Autotrophic Origin of Life

Juan C. Fontecilla-Camps^{1,2,3}

¹ Metalloproteins Unit, University Grenoble Alpes, Institut de Biologie Structurale, F-38044 Grenoble, France; E-Mail: juan.fontecilla@ibs.fr; Tel.: +33-0-4-57428524

² Commissariat à l'Énergie Atomique, Institut de Biologie Structurale, F-38044 Grenoble, France

³ Centre National de la Recherche Scientifique, Institut de Biologie Structurale, F-38044 Grenoble, France

External Editor: Niles Lehman

Received: 23 October 2014; in revised form: 27 November 2014 / Accepted: 11 December 2014 /

Published: 16 December 2014

Abstract: Spark-tube experiments and analysis of meteorite contents have led to the widespread notion that abiotic organic molecules were the first life components. However, there is a contradiction between the abundance of simple molecules, such as the amino acids glycine and alanine, observed in these studies, and the minimal functional complexity that even the least sophisticated living system should require. I will argue that although simple abiotic molecules must have primed proto-metabolic pathways, only Darwinian evolving systems could have generated life. This condition may have been initially fulfilled by both replicating RNAs and autocatalytic reaction chains, such as the reductive citric acid cycle. The interactions between nucleotides and biotic amino acids, which conferred new functionalities to the former, also resulted in the progressive stereochemical recognition of the latter by cognate anticodons. At this point only large enough amino acids would be recognized by the primordial RNA adaptors and could polymerize forming the first peptides. The gene duplication of RNA adaptors was a crucial event. By removing one of the anticodons from the acceptor stem the new RNA adaptor liberated itself from the stereochemical constraint and could be acylated by smaller amino acids. The emergence of messenger RNA and codon capture followed.

Keywords: primordial soup; stereochemical era; genetic code; autocatalytic cycle; low-specificity catalysts; origin of life

1. Introduction

Early propositions [1], followed by the original Miller-Urey [2] and later spark-tube experiments [3,4], along with meteorite content analyses [5,6], have been interpreted as evidence for a heterotrophic origin of life. According to this idea, commonly represented by the well-known “primordial soup” concept, abiotically generated amino acids, as well as other organic molecules, either synthesized on Earth or brought in by celestial bodies, were the first building blocks of life. According to the carbon-isotopic composition of apatite inclusions from the Isua belt in Greenland the first life on Earth appeared not later than about 3800 Myr ago, approximately 700 Myr after the planet’s formation [7]. Conversely, the earliest known microfossils, dated at 3500 Myr ago, are already structurally complex [7] and belong to microorganisms that were most likely modern in most aspects including a DNA-based genetic code. Although, as we will see below, views concerning the nature of an initial genetic coding or its absence are highly variable, it is generally agreed that the extensive involvement of RNA in contemporary protein synthesis strongly suggests it played a fundamental role in the very early stages of peptide formation. Consistent with this view RNAs are thought to have appeared very early on Earth both as self-replicators and catalysts [8]. Indeed, the discovery of ribozymes (catalytic RNAs), led W. Gilbert to propose the existence of an “RNA World”, predating the emergence of DNA and proteins [9]. The most appealing aspect of this concept is the plausible unification of informational capability and catalysis in the same molecule. Furthermore, these RNAs should have been capable of undergoing Darwinian evolution catalyzing many relevant reactions in a “RNA World” scenario [10].

2. Origin of the Genetic Code and the First Peptides

2.1. Theories about Codeless Peptide Synthesis and the Code Appearing before Translation

Some authors have proposed that the synthesis of the very first polypeptides took place in the absence of a code [11,12]. Thus, Krupkin *et al.* [12] have analyzed ribosome X-ray structures and proposed that the peptidyl transferase center (PTC) of the ribosome originally evolved to perform RNA World reactions unrelated to peptide synthesis and was subsequently recruited to catalyze non-coded amino acid polymerization [11,13]. Conversely, Johnson and Wang [14], who have analyzed amino acid side chain-ribonucleotide interactions in the ribosome, concluded that the genetic code may have developed in two stages: a first stage when prebiotically available amino acids formed peptides using a code that was not influenced by anticodon-amino acid interactions and a second stage when these interactions became essential and the code expanded. This latter stage has been defined by Yarus as the “stereochemical era” of genetic code evolution [15]. Rodin *et al.* [16] have published an exhaustive analysis of the stereochemical concept and concluded that the genetic code originated before translation. Along similar lines, several authors have proposed that RNAs could have bound (abiotic) amino acids to either improve stability or expand their functional capabilities, or both [11,17–19]. This interaction

would have eventually led to the recognition of specific amino acids by RNAs predating the emergence of peptide synthesis [18] and initiating the above-mentioned “stereochemical era” of amino acid-anticodon interactions, which was not immediately related to code evolution. At this stage, (ribo)synthetase-adaptor complexes already carry anticodons that are recognized by cognate amino acids. The amino acid binds to this anticodon triplet and esterifies the adaptor, which then dissociates from the (ribo)synthetase. The modified RNA is postulated to be a proto-tRNA [20] or it uses the amino acid as a functional handle [18]. Alternatively, Wong has proposed that initial peptidation (or aminoacylation) of RNAs was not specific and that amino acid-codon (not anticodons) interactions appeared later in evolution [21].

2.2. *Where Did the First Peptides' Amino Acids Come from?*

Regardless which hypothesis about the origin of early polypeptide synthesis is considered it is generally accepted in most of the relevant studies published so far that the first biologically polymerized amino acids were of abiotic origin. For instance, based on the concept of co-evolution of the genetic code with amino acid biosynthesis, Wong [21] has defined two types of amino acids depending on whether they were initially supplied by the environment (Phase 1) or were biosynthetically produced (Phase 2). This author classified Gly, Ala, Ser, Asp, Glu, Val, Leu, Ile, Pro and Thr as Phase 1 amino acids and Phe, Tyr, Arg, His, Trp, Asn, Gln, Lys, Cys and Met as Phase 2 amino acids. At some point during life evolution metabolic pathways developed so that biosynthetically produced Phase 1 amino acids replaced abiotic ones and Phase 2 amino acids emerged and were incorporated into proteins. At this point the genetic code expanded through co-evolution and codon capture by these new amino acids and eventually reached its present state [22]. Attractive as they may appear at first sight these ideas are not without conceptual difficulties because the use of abiotic amino acids in early functional peptides or in complex with RNAs is difficult to rationalize. If we consider spark-tube experiments [2–4] and analysis of meteorites [5] as valid models for the genesis of building blocks at life origins, then small simple amino acids, such as Gly and Ala -the latter as a racemic mixture- should have been by and large the most abundant in the primordial oceans and other possible early settings. However, binding of these small molecules to RNAs, as handles lacking large functional side chains, was unlikely to provide these nucleic acids with any significant evolutionary advantage [18]. The same reasoning applies to peptides mostly composed of these simple amino acids [21]. Conversely, more functionally relevant amino acids, such as Glu or Asp, must have been present in significantly lower concentrations in the soup [4] and unless one invokes some kind of early coding unrelated to the current one and restricted to a few amino acids [23], their covalent binding to RNAs and subsequent polymerization should have been seriously hindered by interference from, among others, the very abundant Ala and Gly. Another serious limitation to the functional diversity that these molecules could have conferred to RNAs and to early free polypeptides is the systematic absence of abiotically produced basic amino acids, such as Arg, Lys and His [24]. This absence would be especially problematic if these peptides were to interact with nucleic acids. A second source of concern is the biological retention of only a set of (L-) amino acids from large amounts of abiotic racemic mixtures of these molecules. Weber and Miller [25] have provided a rationale for the absence from proteins of some amino acids such as ornithine, which would undergo lactamization in peptides and upon acylation to RNA. However, the same approach does not work for others. For

instance, α -amino-n-butyric acid, norvaline and norleucine are abiotically found in significant amounts although they are not proteogenic [25].

2.3. The “Stereochemical Era” of Code Evolution

As may be expected, small amino acids like Gly, Ala, Pro and Ser do not generate cognate RNA anticodons in SELEX experiments carried out to test the “stereochemical era” concept of amino acid-anticodon interactions [15]. Conversely, specific and highly statistically significant anticodon-amino acid interactions have been demonstrated for Arg, Ile, His, Phe, Tyr and Trp [15]. These results are consistent with those obtained by Johnson and Wong [14] in their ribosomal RNA-protein interaction study. These authors found that Gly, Ala, Val, Pro, Ser, Glu and Thr do not bind to RNA regions containing their anticodon sequences whereas the remaining ribosomal proteins amino acid side chains, except Asn and Cys, interact with RNA sequences corresponding to their anticodons (Table 1). As mentioned above, these observations led Johnson and Wong to conclude that the genetic code may have developed first independently from amino acid-anticodon interactions and then stereochemically dependent on them. This leaves us with two options: either (i) some kind of “frozen accident” generated anticodon triplets for the small abiotic amino acids, which were not stereochemically determined and this process was followed and completed by stereochemically-assigned anticodons *of essentially the same nature* for the remaining larger amino acids; or (ii) a different primordial coding for small amino acids eventually evolved into the current triplet code with the subsequent stereochemically-determined incorporation of larger amino acids. Neither option seems especially plausible because they require either an improbable teleonomic event (the accidental emergence of a coding triplet with *future* stereochemical applications) or a very difficult to picture transition between different coding systems. Rather, the most parsimonious interpretation of these results implies that the first amino acids to be incorporated into early polypeptides *already* used a stereochemically-determined code and were for the most part the proto-metabolically generated, bulkier ones (although transcription errors were probably frequent and occasionally non-coded amino acid could have been incorporated) (Table 1). The contradiction between an initial stereochemically-determined code and the generally accepted idea that simpler amino acids were the first to be part of peptides and to enter the (modern) genetic code has already been addressed by Yarus and co-workers [15,26]. These authors have argued that RNA-based metabolic pathways could have generated the larger amino acids before the emergence of the code. Yarus also seems to disagree with the involvement of small abiotically generated amino acids as building blocks in early polypeptide synthesis [15]. Additional support for a stereochemical origin of the genetic code is the observed chirality of some biological amino acids. Indeed, a preference for L-Ile, L-Leu and L-Tyr over their D- counterparts has been observed in RNA aptamer binding experiments [15]. It is difficult to imagine a preference for L-amino acids in proteins arising from a “frozen accident” occurring in abiotic racemic mixtures of small amino acids.

Table 1. Comparison of abiotically synthesized amino acids (Miller-Urey, [2]) and amino acids that interact with their anticodons in the ribosome (Johnson-Wang, [14]).

Amino Acid	Miller-Urey	Johnson-Wang
A	+	
C		N/D
D	+	+
E	+	
F *		+
G	+	
H *		+
I *	+	+
K		+
L *	+	+
M		+
N		
P	+	
Q		+
R *		+
S	+	
T	+	
V	+	
W *		+
Y *		+

Note: * Amino acids that bind anticodons in SELEX experiments [15].

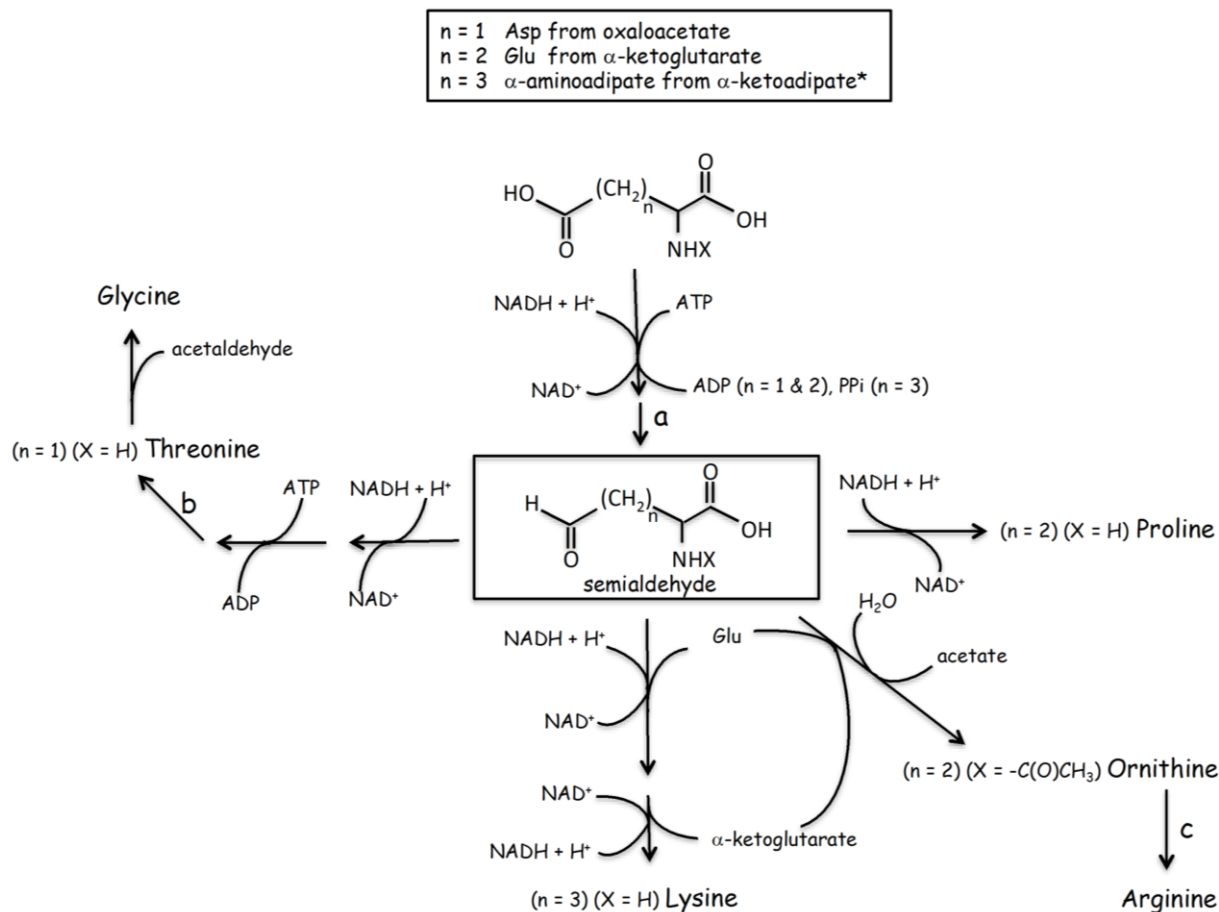
3. Discussion

3.1. Abiotic Amino Acids as unlikely Components of the First Peptides

Based on the arguments outlined above, I will argue that the concept of simple abiotic amino acids being the first components of early polypeptides is likely to be incorrect. In the “primordial soup” concept it is assumed that the first abiotic amino acid synthetic pathways were later replaced by proto-biosynthetic ones [27], an obvious requirement for life to become independent of highly contingent external factors. However, as it might be expected by the very different nature of the reaction conditions, there is no clear link between abiotic amino acid production (and abundance) and the corresponding extant biological synthetic pathways. Conversely, examination of extant metabolisms shows a clear correlation in terms of the catalyzed reactions between synthetic pathways of chemically very different amino acids [28]. As a proof of concept I will use here what I will call the “tricarboxylic acid cycle (TCA)-semialdehyde” route for amino acid synthesis (Figure 1). Semialdehydes obtained from carboxylated amino acids are intermediates in the synthesis of Pro, Thr, Lys and Arg and are formed through a set of closely related reactions including phosphorylations (ATP), hydride-based reductions (NAD(P)H) and, in the cases of Lys and Arg, transaminations (Glu). A common set of primitive catalysts with rather low and broad specificity could have catalyzed these reactions in early settings. Such commonality would have allowed for several related metabolic pathways to co-exist in a given environment using a shared

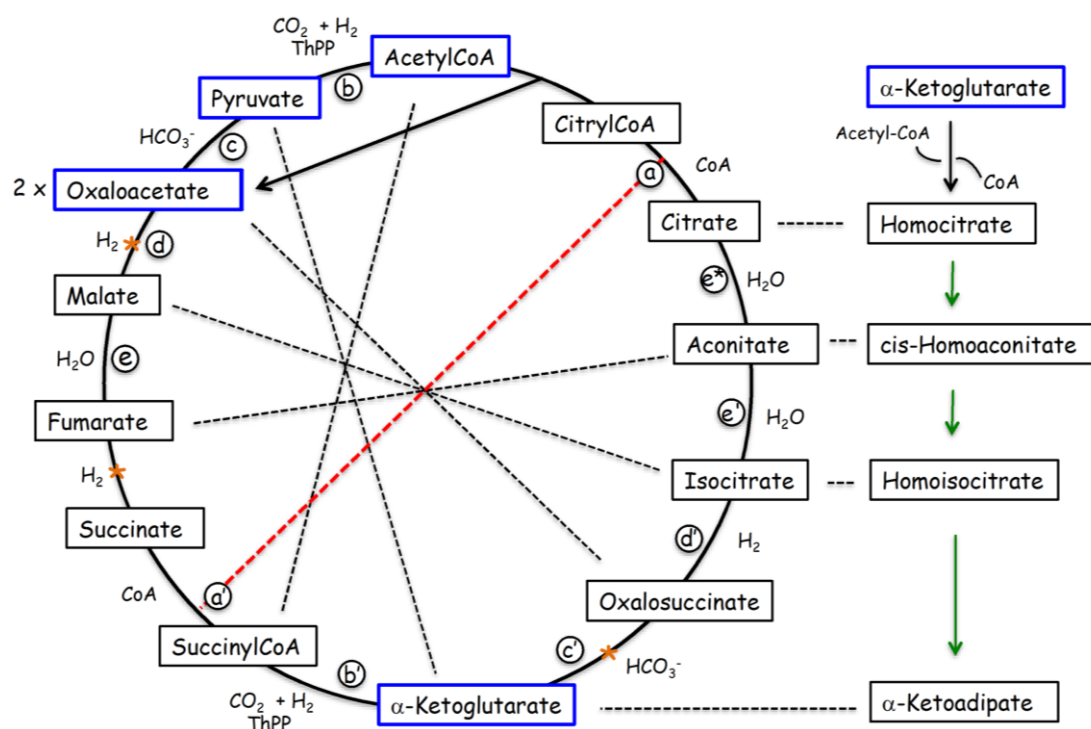
set of catalysts. Although Pro and Thr may not have been initially recognize by RNA anticodons they would have been readily available when the genetic code expanded (see below).

Figure 1. “Tricarboxylic acid cycle-semialdehyde” route for Pro, Thr, Lys and Arg synthesis. a = dephosphorylation; b = PLP-dependent dephosphorylation; c = urea cycle. The starting substrates are aminated versions of intermediates of the pathways in Figure 2.



The existence of a common set of primitive catalysts is also supported by the similarity in terms of substrates and enzymes between two halves of the TCA [29] (Figure 2) and between these halves and pathways corresponding to the synthesis of α -ketoadipate from α -ketoglutarate (Figure 1) and of α -ketoisocaproate from α -ketoisovalerate, precursors in lysine and leucine syntheses, respectively. These catalysts could have successively been inorganic mineral surfaces or globules [30,31], organic cofactors such as nucleotides [32], ribozymes [33] and finally enzymes. In its reductive version, the TCA cycle is a good candidate for an autocatalytic anabolic carbon-fixing pathway that could have operated in a pre-biotic setting [34–36]. Its importance in contemporary reactions is underscored by the central metabolic precursors that derive from it (Figure 2).

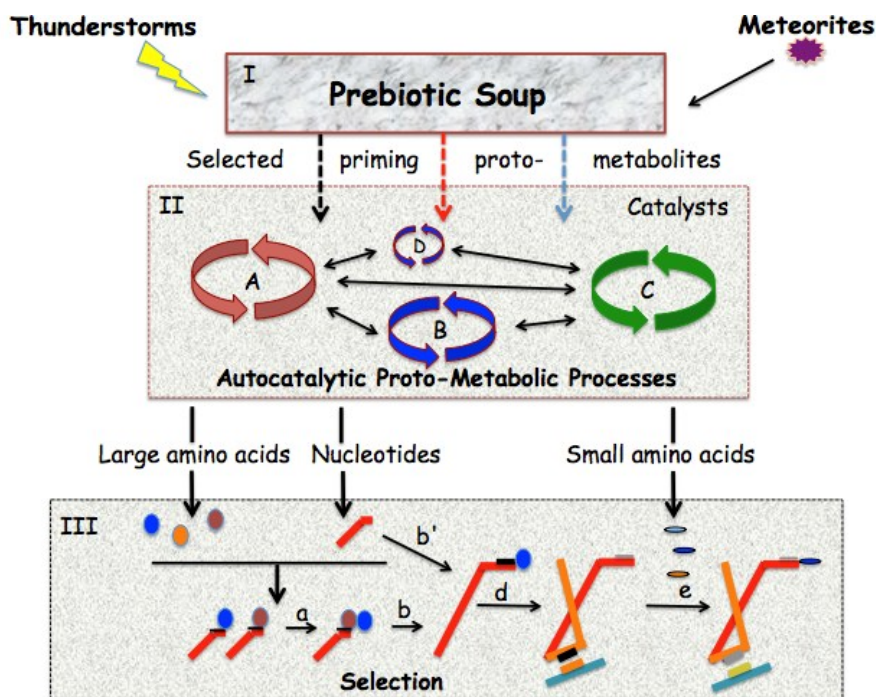
Figure 2. Tricarboxylic acid (TCA) cycle (blue frames depict central metabolites). In its reductive version (counterclockwise in the figure, [29]), carbon is fixed through the generation of two oxaloacetate molecules per cycle. One of these molecules is directly produced from citrylCoA (black arrow) and the other one through acetylCoA and pyruvate (circle). The TCA cycle can be divided into two halves (red dashed line) with approximately equivalent substrates (connected by black dashed lines) and reactions (indicated by a–d and a'–d'; e' = e *). Similar substrates and reactions are also used for the synthesis of α -ketoadipate (right side of the figure; see also Figure 1) and α -ketoisocaproate (not shown,) precursors of lysine and leucine, respectively.



Although it is argued here that abiotic amino acids were not included in the first polypeptides, other abiotically produced molecules, such as pyruvate and citrate in the case of the TCA cycle, could have primed primitive versions of various metabolic pathways [37,38]. Thus, the combination of self-replicating, catalytic RNAs, the origin of which is not discussed here, with a continuous source of amino acids and other precursors from the reductive TCA cycle and other primordial pathways might have been all it took to get life started (Figure 3).

Figure 3. Proposed sequence for the origin of coded polypeptides. **(I)** Organic molecules resulting from atmospheric and geological phenomena and meteorite and comet showers may have accumulated in some primordial Earth settings. However, the resulting mixtures were most likely too heterogeneous to allow for the emergence of proto-biological processes. Instead, it is proposed here that some organic molecules were selected, either through thermal or pH gradients and/or diffusion, or through some other uncharacterized process, to penetrate a different compartment **(II)** where primordial catalysts were available; **(III)** Some of the organic molecules reaching this compartment are proposed to have primed several interacting autocatalytic cycles (A–D), which would have been involved in essential processes

such as carbon fixation. A central idea of the hypothesis put forward here is that the initial compartment had a set of primitive catalysts with low specificity capable of catalyzing similar reactions with different but related substrates (see Figures 1 and 2). Another important notion is that this catalyst set could progressively evolve, for instance from minerals to nucleotides to peptides, still catalyzing the same or closely related reactions, without disrupting the functioning proto-metabolic chains. Thus, some products of the evolving metabolic pathways could have become, in a virtuous cycle, catalysts themselves, ever improving the efficiency and the selectivity of the reactions; **(III)** This compartment could be the same as **(II)** or be connected to it in an intimate way. Here, in the presence of many other molecules such as sugars, carboxylic acids and amines, large amino acids (colored ellipsoids) could have interacted with nucleotides (red L-shaped bars). These interactions would have eventually led to the emergence of anticodon-like structures in the nucleotides (black thin bars) destined to improve their recognition and binding of amino acids functionally important in “RNA World”-type reactions. Eventually, interactions between amino acid-loaded nucleotides could have led to the formation of oligopeptides (reaction a) in a manner similar to the reaction catalyzed by the ribosome. Formation of oligonucleotides (reactions b and b’), which may have improved ribozymes’ catalytic properties led to larger molecules still capable of carrying large functional amino acids and synthesizing oligopeptides containing them (not shown). The crucial gene duplication event, a central component of the hypothesis advanced here, (reaction d) positioned one of the anticodons away from the arm carrying the amino acid thus allowing its evolution towards the recognition of small amino acids (small colored ellipsoids; reaction e with transition from a black to a gray thin bar representing anticodon evolution). These amino acids could now also be incorporated into oligopeptides (not shown) Proto-messenger RNA (green bars) most likely appeared at this time evolving their codons (orange and yellow thin bars) to match the expansion of the anticodon repertoire including now the recognition of the small amino acid set.



3.2. Gene Duplication of Primordial RNA Adaptors and Expansion of the Genetic Code

I will now address the problem of how I think, if the first anticodons were most likely stereochemically generated only for the bulkier amino acids, eventually simpler ones were also assigned their own anticodons/codons. In order to understand this process, we need to review the current views on tRNA evolution. It is generally accepted that modern tRNAs arose from ancestors that underwent a gene duplication event [39,40]. As early as 1978, a few years after the three-dimensional structure of Phe-tRNA became available [41], Hopfield [42] already speculated that the tRNA anticodon region had been contiguous to the acceptor stem before being displaced to the other side of the molecule by gene duplication. As discussed above, the idea of having the amino acid bound to its cognate RNA anticodon, next to where the amino acylation should occur, makes sense because it would protect the aminoacyl linkage against hydrolysis [42,43]. Two aminoacyl-RNAs could have then interacted to form a dipeptide [44] not unlike contemporary acylated tRNAs do in the peptidyl transferase site of the ribosome [45]. However, direct amino acid recognition by the adaptor could only be effectual if its side chain was bulky enough to interact with the cognate bases of a possible anticodon; this proto-tRNA molecule, like a vinyl music record, was an “analog” device. By moving one of the two resulting anticodons to the other end of the adaptor molecule the gene duplication event liberated the coding process from this restrictive anticodon-amino acid direct recognition. RNA aminoacylation must have also become progressively independent of direct amino acid recognition at the acceptor stem and more dependent on anticodon-codon recognition at the anticodon region (Figure 3). At this point, the tRNA molecule became, like a Compact Disc, a “digital” device. Schimmel *et al.* [46] have argued that there is an operational RNA code embedded in the acceptor stems of contemporary tRNAs. Although this code is now very different from the corresponding amino acid anticodons Rodin *et al.* [16] have postulated that they have a common origin.

The structural revolution represented by the gene duplication of the proto-tRNA, allowed for the emergence of non stereochemically-dependent but structurally similar anticodons for the smaller amino acids. Because anticodons could now interact with cognate codons, a coding messenger RNA molecule could evolve in a way perhaps akin to the one discussed by Yarus for “Direct RNA Templates” (Figure 3) [15]. Alternatively, adaptor, messenger and proto-ribosome activities could have initially been united in one RNA molecule. The possible existence of a multifunctional ancestral RNA is supported by several observations: (i) tRNA and rRNA share significant nucleotide sequences, scattered throughout the structures [47,48]; (ii) tRNAs have a nucleotide triplets pattern reminiscent of a gene-like coding sequence, which suggests that at one point adaptor and coding functions were united in the same RNA molecule [49] and (iii) an *in vitro* selected ribozyme displays both aminoacyl-RNA and peptidyl-RNA synthetic activity [50]. From the data discussed above it may be postulated that independent mRNA and tRNA molecules and protein-based acylating activity appeared later in evolution. Once a coding relationship was established between RNAs and peptides, some of the latter may have evolved into proto-enzymes under the selective pressure exerted to enhance the synthesis of their own constituting essential amino acids and increase the chances of making more RNA-coded polypeptides. Early polypeptides were probably short, of relatively simple composition and lacked well-defined three-dimensional structures. This notion is supported by the crystal structure of the ribosome the proteins of which primarily interact with the RNA catalytic core through their extended unfolded segments [51]. Eventually, enzymatic

takeover of most metabolisms, through gene duplication and refinement of the protein synthetic machinery, must have taken place.

4. Conclusions

Life may have first appeared in rather confined settings with the contribution from abiotic sources mostly limited to the initial precursors of nucleotides and pre-metabolic pathways such as the TCA cycle. A very early requisite for Darwinian-type evolution to occur is the persistence in time and space of molecule types and their reactions. This is what makes the “RNA World” concept so attractive. However, a similar restriction applies to pre- or proto-metabolic processes because they also need to evolve. This, I think, is the major problem with the “primordial soup” hypothesis because it lacks a plausible way of providing a continuous link between the use of abiotic amino acids in early peptides and the emergence of the corresponding biosynthetic pathways. Conversely, the connection between the “RNA World” and pre- or proto-metabolic pathways makes such link possible: if the very first amino acids interacting with RNA were already synthesized through a series of rudimentarily reproducible catalyzed reactions, evolution of these pathways became possible. Without introducing any radical changes in the nature of the reactions themselves catalysts could evolve from minerals to proto-enzymes while keeping a relatively broad specificity. The similar nature of many contemporary amino acid synthetic reaction pathways, like the ones outlined above, reveals their very probable common origin in a promiscuous environment where the same catalysts were involved in several reactions with similar substrates. Besides providing a plausible link between the RNA World and the first proto-metabolic pathways the notion that the first amino acids were synthesized through a series of rudimentarily catalyzed reactions makes sense within the concept of “stereochemical era” of genetic code evolution. This is so because peptides did not have to be mostly composed of the simplest amino acids, as the “primordial soup” hypothesis will imply, but of those that could be synthesized from molecules produced by proto-metabolic pathways such as the TCA cycle, including functionally relevant ones like aspartate and lysine.

Rather than postulating that the “stereochemical era” of genetic code evolution followed some other differently-based process of code generation or no code at all, it is proposed here that stereochemistry was the original determining factor of code evolution and of the chirality of proteogenic amino acids. Because, if this hypothesis is correct, only side chains large enough to generate a cognate anticodon could have been incorporated into early peptides, smaller abiotically produced amino acids probably were not significantly involved in this process. The gene duplication event that most likely affected the proto-tRNA liberated the coding process from the stereochemical constraint that the amino acid-anticodon recognition represented and allowed for its evolution through the incorporation of new amino acids through stereochemically independent but functionally related codon capture.

Acknowledgments

The author thanks the Commissariat à l’Energie Atomique and the Centre National de la Recherche Scientifique for institutional funding.

Conflicts of Interest

The author declares no conflict of interest.

References

1. Oparin, A.I. *The Origin of Life*; MacMillan: New York, NY, USA, 1938.
2. Miller, S.L.; Urey, H.C. Organic compound synthesis on the primitive earth. *Science* **1959**, *130*, 245–251.
3. Lazcano, A.; Bada, J.L. The 1953 Stanley L. Miller experiment: Fifty years of prebiotic organic chemistry. *Orig. Life Evol. Biosph.* **2004**, *33*, 235–242.
4. 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 H₂S-rich spark discharge experiment. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 5526–5531.
5. Kvenvolden, K.; Lawless, J.; Pering, K.; Peterson, E.; Flores, J.; Ponnampuruma, C.; Kaplan, I.R.; Moore, C. Evidence for extraterrestrial amino-acids and hydrocarbons in the Murchison meteorite. *Nature* **1970**, *228*, 923–926.
6. Wolman, Y.; Haverland, W.J.; Miller, S.L. Nonprotein amino acids from spark discharges and their comparison with the Murchison meteorite amino acids. *Proc. Natl. Acad. Sci. USA* **1972**, *69*, 809–811.
7. Mojzsis, S.J.; Arrhenius, G.; McKeegan, K.D.; Harrison, T.M.; Nutman, A.P.; Friend, C.R.L. Evidence for life on Earth before 3800 million years ago. *Nature* **1996**, *384*, 55–59.
8. Unrau, P.J.; David, P.; Bartel, D.P. RNA-catalyzed nucleotide synthesis. *Nature* **1998**, *395*, 260–263.
9. Gilbert, W. Origin of life: The RNA world. *Nature* **1986**, *319*, 618.
10. Yarus, M. Darwinian behavior in a cold, sporadically fed pool of ribonucleotides. *Astrobiology* **2012**, *12*, 1–14.
11. Noller, H.F. The driving force for molecular evolution of translation. *RNA* **2004**, *10*, 1833–1837.
12. Krupkin, M.; Matzov, D.; Tang, H.; Metz, M.; Kalaora, R.; Belousoff, M.J.; Zimmerman, E.; Bashan, A.; Yonath, A. A vestige of a prebiotic bonding machine is functioning within the contemporary ribosome. *Phil. Trans. R. Soc. B* **2011**, *366*, 2972–2978.
13. Fox, G.E. Origin and evolution of the ribosome. *Cold Spring Harb. Perspect. Biol.* **2010**, *2*, doi:10.1101/cshperspect.a003483.
14. Johnson, D.B.F.; Wang, L. Imprints of the genetic code in the ribosome. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 8298–8203.
15. Yarus, M.; Widmann, J.J.; Knight, R. RNA–amino acid binding: A stereochemical era for the genetic code. *J. Mol. Evol.* **2009**, *69*, 406–429.
16. Rodin, A.S.; Szathmáry, E.; Rodin, S.N. On origin of genetic code and tRNA before translation. *Biol. Direct* **2011**, *6*, doi:10.1186/1745-6150-6-14.
17. Poole, A.M.; Jeffares, D.C.; Penny, D. The path from the RNA world. *J. Mol. Evol.* **1998**, *46*, 1–17.
18. Szathmáry, E. The origin of the genetic code: Amino acids as cofactors in an RNA World. *Trends Genet.* **1999**, *15*, 223–229.

19. Wolf, Y.I.; Koonin, E.V. On the origin of the translation system and the genetic code in the RNA world by means of natural selection, exaptation, and subfunctionalization. *Biol. Direct* **2007**, *2*, doi:10.1186/1745-6150-2-14.
20. Suga, H.; Hayashi, G.; Terasaka, N. The RNA origin of transfer RNA aminoacylation and beyond. *Phil. Trans. R. Soc. B* **2011**, *366*, 2959–2964.
21. Wong, J.T.-F. Emergence of life: From functional RNA selection to natural selection and beyond. *Front. Biosci.* **2014**, *19*, 1117–1150.
22. Wong, J.T.-F. A coevolution theory of the genetic code. *Proc. Natl. Acad. Sci. USA* **1975**, *72*, 1909–1912.
23. Lehmann, J.; Cibils, M.; Libchaber, A. Emergence of a code in the polymerization of amino acids along RNA templates. *PLoS One* **2009**, *4*, doi:10.1371/journal.pone.0005773.
24. Longo, L.M.; Blaber, M. Protein design at the interface of the pre-biotic and biotic worlds. *Arch. Biochem. Biophys.* **2012**, *526*, 16–21.
25. Weber, A.L.; Miller, S.L. Reasons for the occurrence of the twenty coded protein amino acid. *J. Mol. Evol.* **1981**, *17*, 273–284.
26. Yarus, M.; Caporaso, J.G.; Knight, R. Origins of the genetic code: The escaped triplet theory. *Annu. Rev. Biochem.* **2005**, *74*, 179–198.
27. Bada, J.L. New insights into prebiotic chemistry from Stanley Miller's spark discharge experiments. *Chem. Soc. Rev.* **2013**, *42*, 2186–2196.
28. Hernández-Montes, G.; Díaz-Mejía, J.J.; Pérez-Rueda, E.; Segovia, L. The hidden universal distribution of amino acid biosynthetic networks: A genomic perspective on their origins and evolution. *Genome Biol.* **2008**, *9*, doi:10.1186/gb-2008-9-6-r95.
29. Braakman, R.; Smith, E. Metabolic evolution of a deep-branching hyperthermophilic chemoautotrophic bacterium. *PLoS One* **2014**, *9*, doi:10.1371/journal.pone.0087950.
30. Wächtershäuser, G. Before enzymes and templates: Theory of surface metabolism. *Microbiol. Rev.* **1988**, *52*, 452–484.
31. Martin, W.; Russell, M.J. On the origins of cells: A hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes to nucleated cells. *Philos. Trans. R. Soc. Lond. Ser. B* **2003**, *358*, 59–83.
32. Jadhav, V.R.; Yarus, M. Coenzymes as coribozymes. *Biochimie* **2002**, *84*, 877–888.
33. Yarus, M. Getting past the RNA world: The initial darwinian ancestor. *Cold Spring Harb. Perspect. Biol.* **2011**, *3*, doi:10.1101/cshperspect.a003590.
34. Zhang, X.V.; Martin, S.T. Driving parts of the Krebs cycle in reverse through mineral photochemistry. *J. Am. Chem. Soc.* **2006**, *128*, 16032–16033.
35. Zhang, X.V.; Ellery, S.P.; Cynthia, M.; Friend, C.M.; Holland, H.D.; Michel, F.M.; Martin, A.A.; Schoonen, M.A.A.; Martin, S.T. Photodriven reduction and oxidation reactions on colloidal semiconductor particles: Implications for prebiotic synthesis. *J. Photochem. Photobiol. A Chem.* **2007**, *18*, 301–311.
36. Guzman, M.I.; Martin, S.T. Prebiotic metabolism: Production by mineral photoelectrochemistry of α -Ketocarboxylic acids in the reductive tricarboxylic acid cycle. *Astrobiology* **2009**, *9*, 833–842.

37. Cooper, G.; Reed, C.; Nguyen, D.; Carter, M.; Wang, Y. Detection and formation scenario of citric acid, pyruvic acid, and other possible metabolism precursors in carbonaceous meteorites. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 14015–14020.
38. Griffith, E.C.; Shoemaker, R.K.; Vaida, V. Sunlight-initiated chemistry of aqueous pyruvic acid: Building complexity in the origin of life. *Orig. Life Evol. Biosph.* **2013**, *43*, 341–352.
39. Di Giulio, M. Permuted tRNA genes of *Cyanidioschyzon merolae*, the origin of the tRNA molecule and the root of the Eukarya domain. *J. Theor. Biol.* **2008**, *253*, 587–592.
40. Sugahara, J.; Fujishima, K.; Morita, K.; Tomita, M.; Kanai, A. Disrupted tRNA gene diversity and possible evolutionary scenarios. *J. Mol. Evol.* **2009**, *69*, 497–504.
41. Kim, S.H.; Sussman, J.L.; Suddath, F.L.; Quigley, G.J.; McPherson, A.; Wang, A.H.; Seeman, N.C.; Rich, A. The general structure of transfer RNA molecules. *Proc. Natl. Acad. Sci. USA* **1974**, *71*, 4970–4974.
42. Hopfield, J.H. Origin of the genetic code: A testable hypothesis based on tRNA structure, sequence, and kinetic proofreading. *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 4334–4338.
43. Erives, A. A Model of proto-anti-codon RNA enzymes requiring L-amino acid homochirality. *J. Mol. Evol.* **2011**, *73*, 10–22.
44. Schimmel, P.; Henderson, B. Possible role of aminoacyl-RNA complexes in noncoded peptide synthesis and origin of coded synthesis. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 11283–11286.
45. Nissen, P.; Hansen, J.; Ban, N.; Moore, P.B.; Steitz, T.A. The structural basis of ribosome activity in peptide bond synthesis. *Science* **2000**, *289*, 920–930.
46. Schimmel, P.; Giegé, R.; Moras, D.; Yokoyama, S. An operational RNA code for amino acids and possible relationship to genetic code. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 8763–8768.
47. Bloch, D.P.; McArthur, B.; Widdowson, R.; Specter, D.; Guimaraes, R.C.; Smith, J. tRNA-rRNA sequence homologies: Evidence for a common evolutionary origin? *J. Mol. Evol.* **1983**, *19*, 420–428.
48. Nazarea, A.D.; Bloch, D.P.; Semrau, A.C. Detection of a fundamental modular format common to transfer and ribosomal RNAs: Second-order spectral analysis. *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 5337–5341.
49. Eigen, M.; Winkler-Oswatitsch, R. Transfer-RNA, an early gene? *Naturwissenschaften* **1981**, *68*, 282–292.
50. Illangasekare, M.; Yarus, M. A tiny RNA that catalyzes both aminoacyl-RNA and peptidyl-RNA synthesis. *RNA* **1999**, *5*, 1482–1489.
51. Hsiao, C.; Mohan, S.; Kalahar, B.K.; Williams, L.D. Peeling the onion: Ribosomes are ancient molecular fossils. *Mol. Biol. Evol.* **2009**, *26*, 2415–2425.