THE EARLY PHASES OF GENETIC CODE ORIGIN: CONJECTURES ON THE EVOLUTION OF CODED CATALYSIS

MASSIMO DI GIULIO

Institute of Genetics and Biophysics Adriano Buzzati Traverso, CNR, Naples, Napoli, Italy
(* author for correspondence, e-mail: digiulio@iigb.na.cnr.it, Fax: +39 081 5936123)

Abstract. A review of the most significant contributions on the early phases of genetic code origin is presented. After stressing the importance of the key intermediary role played in protein synthesis, by peptidyl-tRNA, which is attributed with a primary function in ancestral catalysis, the general lines leading to the codification of the first amino acids in the genetic code are discussed. This is achieved by means of a model of protoribosome evolution which sees protoribosome as the central organiser of ancestral biosynthesis and the mediator of the encounter between compounds (metabolite-pre-tRNAs) and catalysts (peptidyl-pre-tRNAs). The encounter between peptidyl-pre-tRNA catalysts in protoribosome is favoured by metabolic pre-mRNAs and later resulted (given the high temperature at which this evolution is supposed to have taken place) in the evolution of mRNAs with codons of the type GNS. These mRNAs codified only for those amino acids that the coevolution theory of genetic code origin sees as the precursors of all other amino acids. Some aspects of the model here discussed might be rendered real by the transfer-messenger RNA molecule (tmRNA) which is here considered a molecular fossil of ancestral protein synthesis.

Keywords: catalysis evolution, peptidyl-pre-tRNA, protein synthesis origin, pre-tRNA, protoribosome, tmRNA

1. Introduction: The Logic Underlying the Hypothesis

Many papers have been written on the origin of the genetic code (for reviews see: Szathmáry, 1993, 1999; Di Giulio, 1997a; Knight *et al*., 1999) but it has perhaps never been explicitly claimed that a potentially important definition of the genetic code is that this origin must simply represent the evolution of coded catalysis. If we neglect the genetic complexity of actual organisms and consider the main role played today by proteins, i.e. catalysis, we come to the simple equation: genetic code equals coded enzymatic catalysis. This must have been even more likely for the early phases of genetic code evolution. That the origin of the genetic code must correspond to that of the evolution of enzymatic catalysis becomes particularly clear when we consider the other roles now played by proteins, i.e. the structural and regulatory roles. It is certainly true that the origin of the genetic code cannot be related to the role played by proteins during regulation, in the most general sense, because protein-mediated regulation in ancestral entities cannot have played a fundamental role. Furthermore, we can see that the origin of the genetic code can only

possibly have been related to the structural role played by proteins if we consider that this might have been truly important only if ancestral proteins were called upon to perform fundamental functions, for example that of the main constituents of cell membranes. However, the important structural or regulatory functions possibly or actually performed by proteins do not seem so general, numerous and, above all, more important than the role that proteins play in catalysis. Hence, although the intervention of proteins might have been precocious and important, such as in the structures of primordial entities, this was most probably only a secondary effect and not the driving force behind genetic code evolution.

It therefore seems automatically obvious that the intervention of proteins in primordial catalysis must have been the driving force behind genetic code origin. In the present paper I attempt to analyse the early phases of genetic code origin in the light of the hypothesis that the origin was mainly determined by the evolution of catalysts which eventually resulted in coded catalysis.

2.   The 'Cellular' Entity in which Genetic Code Origin began

I assume that the origin of the genetic code began in 'cellular' entities in which only a crude metabolism had evolved. This assumption therefore falls into the class of theories hypothesising that metabolism originated prior to replication (De Duve, 1991; Dyson, 1985). Here it is not relevant whether the origin of life was heterotrophic or autotrophic in nature but the only important thing is that there were entities capable of achieving a crude metabolism, such as the autotrophic one described by Wachtershauser (1988, 1992). Moreover, I assume that these entities lived at a high temperature.

In these entities there was absolutely no genetic apparatus, even of a type much simpler than those we now know, and reproduction was ensured by a crude sharing of the entity constituents among the sister entities (see for example: Dyson, 1985; Segré and Lancet, 1999). At this stage of evolution, catalysis was performed by ions and by small organic molecules (for references see Di Giulio, 1997b) but as far as we are concerned here, this must have also been performed by amino acids as such, di-amino acids and tri-amino acids and therefore also by brief peptides (De Duve, 1991; Bar-Nun et al., 1994; Di Giulio, 1997b). The other crucial assumption in the model investigated here is that amino acids were assisted in catalysis by a nucleotide component (White, 1976, 1982; Wong, 19991; Szathmáry, 1993; Di Giulio, 1997b; Wong and Xue, 2002). This is the simplest stage in which, perhaps for the first time, amino acids encountered nucleotides (Di Giulio, 1997b). This is the starting point for my commentary.

3. The Evolution of Coded Catalysis: Towards Protein Synthesis

We are thus in an evolutionary stage in which catalysis was performed by amino acids, small peptides, nucleotides, amino acids conjugated with nucleotides (coenzymes) or brief nucleotide chains, in addition to ions and small organic compounds (for references see Di Giulio, 1997b). I assume that at this stage 'protein synthesis' had a lot in common with the synthesis of peptide antibiotics (Lipmann, 1965, 1971; De Duve, 1987, 1991; Danchin, 1989; Kleinkauf and Dohren, 1990; Di Giulio, 1996b). Therefore, the first peptide synthesis might have taken place on molecules of phospho-pantetheine (Lipmann, 1965) or on pantetheine attached to a phosphoserine (Miller and Schlesinger, 1993) in a thioester world (De Duve, 1988, 1991) and/or on a variant of the CoA (Reanney, 1977) in an RNA world (Gilbert, 1986; Joyce, 1989). Peptide synthesis thus progressed on tRNA-like molecules evolving from CoA-like molecules (Reannet, 1977) towards the actual protein synthesis (Orgel, 1989; Wong, 1991; Schimmel and Henderson, 1994; Di Giulio, 1994b, 1997b). Moreover, I assume, as Wong (1991) has, that any model for the origin of protein synthesis, and thus of the genetic code, must explain the evolution of the key intermediary of this origin, namely peptidyl-tRNA. As already suggested (Wong, 1991), the evolution of this complex must have been the fundamental step that resulted in the origin of protein synthesis in that, as the peptidyl-tRNA no longer performs a function per se. It evidently must have in itself performed a fundamental role in catalysis at evolutionary times that were crucial for genetic code development (Wong, 1991). In other words, it seems implicitly clear that peptidyl-tRNA-like structures performed all the catalysis needed for survival at certain times because the equivalent complex now has no function per se and so it must almost certainly have had such a function in the distant past because we would otherwise be unable to recognise any selective pressure aiming to get the peptidyl-tRNA to evolve (Wong, 1991).

This argument thus helps to establish that there must have been an evolutionary stage in which all catalysis, or most of it, was performed by peptidyl-tRNA-like molecules. This constitutes a first benchmark in our investigation. It also implies a first evolutionary transition of ancestral catalysts in that many of those present at the beginning of the origin of life were already replaced or supported by peptidyl-tRNA-like molecules in catalysis at this new evolutionary stage.

Another assumption that I consider more than reasonable and extremely likely is that in some stages of protein synthesis, the interaction between two different peptidyl-tRNA-like molecules must have been favoured by an RNA template (Orgel, 1968, 1989; Lahav, 1991; Di Giulio, 1994a). In other words, this interaction between two or more peptidyl-tRNA-like molecules must have been the direct template although not the direct specific sequence (Lahav, 1991; Di Giulio, 1994a).

More specifically, the interaction between two or more peptidyl-tRNA-like molecules was mediated by an RNA template which, at this stage, had the sole task of favouring this interaction by simply bringing together the two or more peptidyl-

tRNA-like molecules and linking them by means of hydrogen bonds between the bases. This link was only partially specific in that the whole population of peptidyl-tRNA-like molecules could interact with just a handful of RNA templates and the resulting synthesis produced catalysts that could be used by ancestral entities only at a low or very low yield.

We must therefore discuss the nature of the ribosome and how the genetic code originated from this evolutionary stage.

4. Ancestral Ribosome

Protoribosome might, in principle, have performed either one specific function or a number of different functions (Campbell, 1991; Gibson and Lamond, 1990). In accordance with the hypothesis being analysed here, I tend to believe that protoribosome became an aggregate of catalysts, perhaps multimers (De Duve, 1991), at a certain stage of these ancestral entities' evolution and later became an aggregate of ribo-nucleopeptides of the type peptidyl-RNA. The function of this aggregate of catalysts was to enable an almost generalised catalysis on a single particle, or more probably on different types of particles.

A selective pressure that favoured the evolution of this type of particle must have kept numerous catalysts physically united in a single aggregate and thus limited their dispersion due to the high temperature at which the process is assumed to have taken place, both generally and in particular during 'cellular' division. Moreover, the physically united catalysts were able to more prolifically catalyse the synthesis of correlated molecules crucial for survival. Furthermore, if as suggested by various theories (Wachtershauser, 1988; Russell *et al.*, 1998), the ancestral entities evolved at a high temperature, then a strong selective pressure aiming to stabilise peptidyl-tRNA-like molecules or, more generally, RNAs (considering their lability at high temperatures) led to the aggregation of these molecules in complexes. Protoribosome might thus have evolved to stabilise peptide-RNA complexes (or RNAs) since this particle would certainly have limited the damage deriving, for instance, from the hydrolysis of certain bonds. This is because in these particles the activity of these RNA molecules should not be lost since the molecular structure should remain mostly spatially intact and in this they are favoured by the forces of aggregation of the particles.

This view of ancestral ribosome's origin explains why ribosome (i) is itself an aggregate of two subunits which might have worked separately in the distant past; (ii) is formed of about sixty mostly small proteins and not of a mere few proteins, and this cannot be easily explained only by the structural role and may bear witness to a multiplicity of past functions; (iii) is formed of a few types of rRNAs, thus further attesting that the large and the small ribosomal subunits might have worked separately and implying an ancient functional heterogeneity of the two ribosomal subunits. This is once again difficult to explain given the single role now played by

rRNAs but it is understandable if the latter were pre-mRNAs in the particles (see below). These interpretations suggest looking for catalytic activities regarding the ancestral entities in the small ribosomal proteins.

Finally, it should be noted that at this evolutionary stage we are not yet in the presence of a genetic code but only of its embryonic form: peptidyl-tRNA-like molecules which were made to interact by means of one or a few RNAs in the protoribsome.

5. Genetic Code Evolution

What must have evolved in protoribosome was the ability to 'read' RNA. Initially, as the metabolites were charged on pre-tRNA (Gibson and Lamond, 1990; Di Giulio, 1994a, 1997a, b) these were placed in a correct apposition by means of interactions with protoribosome's internal RNA (Gibson and Lamond, 1990; Di Giulio, 1994a, 1997a, b). Subsequently protoribosome must have acquired the ability to read external RNA. On these RNAs there must have been written the succession with which the various metabolite-pre-tRNA complexes and the catalysts (peptidyl-pre-tRNAs) were to interact (Tyagi, 1981; Crothers, 1982; Cedergren and Grosjean, 1987; Edwards, 1989, 1996; Gibson and Lamond, 1990; Lamond and Gibson, 1990; Di Giulio, 1997b). These RNAs (pre-mRNAs) on which the successions of steps on the biosynthetic pathways leading to the synthesis of the various compounds were written, are postulated as being the oldest form of mRNA (Gibson and Lamond, 1990). Here we catch a glimpse of the equivalence between genetic code organisation and the biosynthetic pathways of amino acids (Wong, 1975). In other words, at this evolutionary stage there was a code for metabolism whose evolution resulted in the genetic code (Gibson and Lamond, 1990; Lamond and Gibson, 1990; Di Giulio, 1997b).

For the origin of the genetic code, we must show how these pre-mRNAs led to mRNA proper.

Therefore, at this evolutionary stage we have a protoribosome that can read pre-mRNAs on which is written the succession of interactions between pre-tRNAs, bearing not only amino acids but also other metabolites and peptidyl-pre-tRNAs (Gibson and Lamond, 1990; Di Giulio, 1997a, b) to which the model attributes a primary role in catalysis (Wong, 1991; Di Giulio, 1997b). Moreover, here peptidyl-pre-tRNA is seen as a simple and natural evolution of aminoacyl-pre-tRNA (Di Giulio, 1997b).

But how exactly did these pre-mRNAs give rise to mRNA and hence the genetic code? The main selective pressure must have been the one which improved catalysts, namely peptidyl-pre-tRNA molecules, taking them to complete codification. To obtain this improvement, synthesis of peptidyl-pre-tRNA molecules had to change from template-guided to direct sequence synthesis (Orgel, 1989; Lahav, 1991; Di Giulio, 1994a). At this stage the reproducibility of peptidyl-pre-tRNA

molecules is therefore clearly inefficient. In order to increase this efficiency, some changes must have been made. (i) The protein part of peptidyl-pre-tRNA molecules must have gradually assumed the primary role in catalysis. The structural themes of proteins that were most probably being selected were the β-turns (Jurka and Smith, 1987a, b) and the β-sheets (Orgel, 1975, 1977; Brack and Orgel, 1975; Di Giulio, 1996a, 1997b). (ii) The interaction between peptidyl-pre-tRNA molecules taking place on the pre-mRNAs must have become physically contiguous in that the two or more peptidyl-pre-tRNAs that could have paired up with distant regions of the template RNA now require these regions to be contiguous. This evidently improved the reading efficiency of pre-mRNAs. The recombination, in the most general sense of the word, of pre-mRNA molecules or parts of them might have played a fundamental role in this stage. That is to say, a new population of pre-mRNAs was created and then it at least partially replaced the old messages. Clearly, these changes must have provided a strong positive feedback on the whole system.

We have thus reached an evolutionary stage at which the interaction of the peptidyl-pre-tRNA molecules is favoured by contiguous 'codons'. Here 'codons' does not imply that only a triplet of bases favoured recognition between peptidyl-pre-tRNA and the template, but this might have entailed pairings longer than more than three bases.

For the transition towards the genetic code, a subpopulation of pre-mRNAs must have evolved. Then, by means of sequences of codons starting with G and a three-base reading module, the system must have begun to decode these messages with the amino acids: Asp, Glu, Ser, Ala and Gly. But how did this take place? This is the crucial step in genetic code origin. We must thus explain how a system that decoded pre-mRNAs codifying both successions of interactions between metabolite-pre-tRNAs, aminoacyl-pre-tRNAs and peptidyl-pre-tRNAs and interactions between peptidyl-pre-tRNAs progressed to a system that finally preferred to decode sequences with codons beginning with G.

The evolution of RNAs rich in codons beginning with G, or more generally, sequences rich in G and C might be the consequence of the high temperature at which the process is assumed to have taken place. Sequences of RNA rich in G and C should be more resistant than sequences rich in A and U at high temperatures because they should, by means of their more rigid secondary and tertiary structures, keep the molecules' active parts intact even if some of the bonds in the molecules are broken. This might explain how G and C rich messages evolved and why codons having G as their first base evolved and codified for amino acids that were the first to be inserted into the genetic code (Wong, 1975; Trifonov, 2000; Ikehara *et al.*, 2002).

I find it extremely hard to see how a pre-mRNA codifying for a succession of biosynthetic steps became an mRNA codifying for an albeit short protein. Nevertheless, I feel that explaining this in a different way from the one I use below might be extremely important in order to understand the profound implications of the coevolution theory of genetic code origin (Wong, 1975). One possibility is

that pre-mRNAs which, as has been said, also specified the successions of interactions between peptidyl-pre-tRNA molecules represented the only true precursors of the new mRNAs that hence evolved from these. Whereas, pre-mRNAs codifying for the successions of the steps of biosynthetic pathways (by means of interactions between metabolite-pre-tRNAs) simply became extinct. In other words, real mRNAs evolved only from pre-mRNAs which specified the successions of interactions between molecules of peptidyl-pre-tRNA. Clearly the fine tuning of the reading module of pre-mRNAs which contained high percentages of G and C, under the selective pressure of the improvement of catalysis achieved by peptidyl-pre-tRNA molecules, must have resulted in sequences containing codons of the type GNS (N = any base, and S = G or C) (Ikehara *et al.*, 2002). This type of codon in the actual genetic code specifies the amino acids Ala, Val, Asp, Glu and Gly. These amino acids include most of what are held to be the oldest to be codified (Wong, 1975; Di Giulio, 1997a; Trifonov, 2000; Ikehara *et al.*, 2002). In any case, since this list of amino acids does not include Ser, I am led to believe that either mRNAs with GNS codons still codified ambiguously at this stage, for instance the GNC or GNG codons codified for Ser and Gly (biosynthetically interconvertible), or it is only when these mRNAs begin to use all the other codons that Ser entered the genetic code. The latter stage thus defined the synonymous codon domains of all the precursor amino acids as envisaged by the coevolution theory of the origin of the genetic code (Wong, 1975; Di Giulio and Medugno, 1999). And from this point on, it seems to me that the further evolution of the genetic code is well described by this theory (Wong, 1975).

6. Experimental Evidence: tmRNA, a Probable Molecular Fossil of Ancient Protein Synthesis

Transfer-messenger RNA (tmRNA) is a molecule that can act both as a transfer RNA and as a messenger RNA (Muto *et al.*, 1998). As tmRNAs have a tRNA-like structure at the 3′ end with the alanine identity determinants, they are charged with this amino acid. Charged in this way, they can interact with ribosome when the reading of a truncated mRNA is unable to continue the synthesis of the polypeptide chain. In this case, tmRNA is able to migrate to ribosome. We thus have transfer of the polypeptide chain to alanine on tmRNA. Protein synthesis then continues by reading a sequence inside the tmRNA. A tag sequence is thus attached to the protein's C-terminal end which will route the protein to the degradation pathway (Muto *et al.*, 1998).

tmRNA is present in the whole Bacteria domain (Gillet and Felden, 2001). This shows that it is a very ancient molecule. However, its presence in the last universal common ancestor is not certain since the Eukarya and Archaea domains do not seem to possess the tmRNA molecule (Gillet and Felden, 2001). Nevertheless,
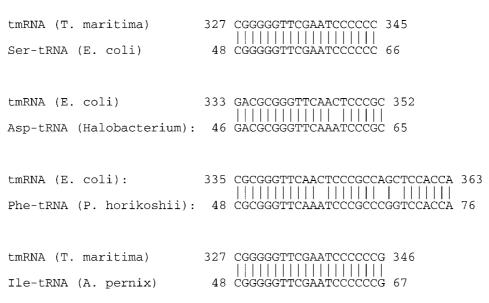
```
tmRNA (T. maritima)          327 CGGGGGGTTCGAATCCCCCC 345
                                 |||||||||||||||||||||
Ser-tRNA (E. coli)            48 CGGGGGGTTCGAATCCCCCC 66
```

```
tmRNA (E. coli)              333 GACGCGGGTTCAACTCCCGC 352
                                 ||||||||||||||| ||||||
Asp-tRNA (Halobacterium):     46 GACGCGGGTTCAAATCCCGC 65
```

```
tmRNA (E. coli):             335 CGCGGGTTCAACTCCCGCCAGCTCCACCA 363
                                 |||||||||||| ||||||| | |||||||
Phe-tRNA (P. horikoshii):     48 CGCGGGTTCAAATCCCGCCCGGTCCACCA 76
```

```
tmRNA (T. maritima)          327 CGGGGGGTTCGAATCCCCCCG 346
                                 |||||||||||||||||||||||
Ile-tRNA (A. pernix)          48 CGGGGGGTTCGAATCCCCCCG 67
```

Figure 1. Examples of the similarity between some sequences of genes of the tmRNA molecules and those of the tRNAs.

I believe that tmRNA is a fossil of ancient protein synthesis for the following reasons.

If tmRNA was an acquired and not an ancestral trait, then its phylogenetic distribution should not have affected an entire domain but should have been more fragmentary. This is because the onset of the three main phyletic lines still seems to be subject to what appears to be a rapid Darwinian evolution (Doolittle and Brown, 1994; Woese, 1998, 2000; Di Giulio, 2001) (for instance, supported by the diversity of the beginning of protein synthesis in the three domains (Kyrpides and Woese, 1998a, b; Di Giulio, 2001)) that would not have been able to give rise to an adaptation like that mediated by tmRNA. The action performed by tmRNA seems to be that of a sophisticated adaptation which is the characteristic feature of organisms that are genetically fully mature and should not therefore be found in the Bacteria domain's ancestor which should still have many rapidly evolving features (Woese, 1998, 2000; Di Giulio, 2001). More directly, the action performed by tmRNA cannot evolve in organisms in which it is protein synthesis itself which is still not fully defined (Di Giulio, 2001) because its action would interfere with the completion of evolution of protein synthesis. Therefore, I conclude that tmRNA was probably a molecule belonging to the last universal common ancestor and hence probably involved in the origin of protein synthesis.

But how is the tmRNA molecule's action mechanism related to the model discussed in this paper?

The homology of tmRNA with the tRNA molecule is clear. For instance, the tmRNA identity determinant which allows alanine to be charged on this molecule is

the same as that of tRNA for alanine (Muto *et al.*, 1998). Moreover, the similarities between the tRNA-like region of tmRNA and the equivalent region of tRNAs are clear (Figure 1) and they are thus homology indices. This indicates a strict evolutionary relationship between tmRNA and tRNA molecules and the characteristic hairpin structure of the 3′ end of the tmRNA molecule might have been the ancestor of the tRNA molecule (Hopfield, 1978; Di Giulio, 1992, 1995).

Therefore, if tmRNA is a molecular fossil of ancient protein synthesis, as I firmly believe, then tmRNA must represent certain phases of its evolution. The hypothetical molecules of aminoacyl-pre-tRNA (or even those of peptidyl-pre-tRNA) which, according to the above discussions interacted with the protoribosome's internal RNA, might be made real by the tmRNA molecule as this also seems to interact with the 16S rRNA (Muto *et al.*, 1998) thus supporting the claims made in the present paper. In other words, the tmRNA molecule might represent the molecular fossil of the hypothetical aminoacyl-pre-tRNA and peptidyl-pre-tRNA molecules which have been extensively discussed here.

Acknowledgement

References

Bar-Nun, A., Kochavi, E. and Bar-Nun, S.: 1994, Assemblies of Amino Acids as Possible Prebiotic Catalysts, *J. Mol. Evol.* **39**, 116−122.

Brack, A. and Orgel, L. E.: 1975, *β*-Structures of Alternating Polypeptides and their Possible Prebiotic Significance, *Nature* **256**, 383−387.

Campbell, J. H.: 1991, An RNA Replisome as Ancestor of Ribosome, *J. Mol. Evol.* **32**, 3−5.

Cedergren, R. and Grosjean, H.: 1987, On the Primacy of Primordial RNA, *Biosystems* **20**, 175−180.

Crothers, D. M.: 1982, Nucleic Acid Aggregation Geometry and the Possible Evolutionary Origin of Ribosomes and the Genetic Code, *J. Mol. Biol.* **162**, 379−391.

Danchin, A.: 1989, Homeotopic Transformation and the Origin of Translation, *Prog. Biophys. Mol. Biol.* **54**, 81−86.

De Duve, C.: 1987, Selection by Differential Molecular Survival: A Possible Mechanism of Early Chemical Evolution, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 8253−8256.

De Duve, C.: 1988, Did God make RNA? *Nature* **336**, 209−210.

De Duve, C.: 1991, *Blueprint for a Cell: The Nature and Origin of Life*, Neil Patterson Publishers, Carolina Biological Supply Company, Burlington, NC.

Di Giulio, M.: 1992, On the Origin of the Transfer RNA Molecule, *J. Theor. Biol.* **159**, 199−214.

Di Giulio, M.: 1994a, On the Origin of Protein Synthesis: A Speculative Model Based on Hairpin RNA Structures, *J. Theor. Biol.* **171**, 303−308.

Di Giulio, M.: 1994b, On the Origin of Protein Synthesis: A Speculative Model based on Hairpin RNA Structures, *J. Theor. Biol.* **171**, 303−308.

Di Giulio, M.: 1995, Was it an Ancient Gene Codifyng for a Hairpin RNA that, by Means of Direct Duplication, gave Rise to the Primitive tRNA Molecule? *J. Theor. Biol.* **177**, 95−101.

Di Giulio, M.: 1996a, The β-Sheets of Proteins, the Biosynthetic Relationships between Amino Acids, and the Origin of the Genetic Code, *Orig. Life Evol. Biosph.* **26**, 589−609.

Di Giulio, M.: 1996b, The Origin of Protein Synthesis: On some Molecular Fossils Identified through Comparison of Protein Sequences, *Biosystems* **39**, 159−169.

Di Giulio, M.: 1997a, On the Origin of the Genetic Code, *J. Theor. Biol.* **187**, 573−581.

Di Giulio, M.: 1997b, On the RNA World: Evidence in Favor of an Early Ribonucleopeptide World, *J. Mol. Evol.* **45**, 571−578.

Di Giulio, M.: 2001, The Non-universality of the Genetic Code: The Universal Ancestor was a Progenote, *J. Theor. Biol.* **209**, 345−349.

Di Giulio, M. and Medugno, M.: 1999, Physicochemical Optimization in the Genetic Code Origin as the Number of Codified Amino Acids Increases, *J. Mol. Evol.* **49**, 1−10.

Doolittle, W. F. and Brown, J. R.: 1994, Tempo, Mode, the Progenote, and the Universal Root, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 6721−6728.

Dyson, F.: 1985, *Origins of Life*, Cambridge University Press, Cambridge.

Edwards, M. R.: 1996, Metabolite Channeling in the Origin of Life, *J. Theor. Biol.* **179**, 313−322.

Edwards, M. R.: 1989, A Possible Origin of RNA Catalysis in Multienzyme Complexes, *Orig. Life Evol. Biosph.* **19**, 69−72.

Gibson, T. J. and Lamond, A. I.: 1990, Metabolic Complexity in the RNA World and Implications for the Origin of Proteins Synthesis, *J. Mol. Evol.* **30**, 7−15.

Gilbert, W.: 1986, The RNA World, *Nature* **319**, 618.

Gillet, R. and Felden, B.: 2001, Emerging Views on tmRNA-Mediated Protein Tagging and Ribosome Rescue, *Mol. Microbiol.* **42**, 879−885.

Ikehara, K., Omori, Y., Arai, R. and Hirose, A.: 2002, A Novel Theory on the Origin of the Genetic Code: A GNC-SNS Hypothesis, *J. Mol. Evol.* **54**, 530−538.

Hopfield, J. J.: 1978, Origin of the Genetic Code: A Testable Hypothesis Based on tRNA Structure, Sequence and Kinetic Proofreading, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 4334−4338.

Joyce, G. F.: 1989, RNA Evolution and the Origins of Life, *Nature* **338**, 217−224.

Jurka, J. and Smith, T. F.: 1987a, β-Turn-driven Early Evolution: the Genetic Code and Biosynthetic Pathways, *J. Mol. Evol.* **25**, 15−19.

Jurka, J. and Smith, T. F.: 1987b, β-Turns in Early Evolution: Chirality, Genetic Code, and Biosynthetic Pathways, *Cold Spring Harbor Symp. Quant. Biol.* **52**, 407−410.

Kleinkauf, H. and Von Dohren, H.: 1990, Nonribosomal Biosynthesis of Peptide Antibiotics, *Eur. J. Biochem.* **192**, 1−15.

Knight, R. D., Freeland, S. J. and Landweber, L. F.: 1999, Selection, History and Chemistry: The Three Faces of the Genetic Code. *Trends Biochem. Sci.* **24**, 241−247.

Kyrpides, N. C. and Woese, C. R.: 1998b, Universally Conserved Translation Initiation Factors, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 224−228.

Kyrpides, N. C. and Woese, C. R.: 1998b, Archaeal Translation Initiation Revisited: The Initiation Factor 2 and Eukaryotic Initiation Factor 2N $\alpha - \beta - \delta$ Subunit Families. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 3726−3730.

Lahav, N.: 1991, Prebiotic Co-Evolution of Self-Replication and Translation Or RNA World? *J. Theor. Biol.* **151**, 531−539.

Lamond, A. I. and Gibson, T. J.: 1990, Catalitic RNA and the Origin of Genetic Systems, *Trends Genet* **6**, 145−149.

Lipmann, F., 1965, *In Origins of Prebiological Systems and of their Molecular Matrices*, S. Fox (ed.), Academic, New York, pp. 259−280.

Lipmann, F.: 1971, Attempts to Map a Process Evolution of Peptide Biosynthesis, *Science* **173**, 875−884.

Miller, S. L. and Schlesinger, G.: 1993, Prebiotic Syntheses of Vitamin Coenzymes: II. Pantoic Acid, Pantothenic Acid and the Composition of Coenzyme A. *J. Mol. Evol.* **36**, 308−314.

Muto, A., Ushida, C. and Himeno, H.: 1998, A Bacterial RNA that Functions as both a tRNA and an mRNA, *Trends Biochem. Sci.* **23**, 25−29.

Orgel, L. E.: 1968, Evolution of the Genetic Apparatus, *J. Mol. Biol.* **38**, 381−393.

Orgel, L. E.: 1975, Prebiotic Polynucleotides and Polypeptides, *Isr. J. Chem.* **14**, 11−16

Orgel, L. E.: 1977, '*β*-Turns and the Evolution of Protein Synthesis', in E. M. Bradbury and K. Javaherian (eds), *The Organization and Expression of the Eukaryotic Genome: Proceedings of the International Symposium*, 3−6 May 1976, Teheran. Academic Press, London, pp. 499−504.

Orgel, L. E.: 1989, The Origin of Polynucleotide-Directed Protein Synthesis, *J. Mol. Evol.* **29**, 465−474.

Russell, J., Daia, D. E. and Hall, A. J.: 1998, 'The Emergence of Life From FeS Bubbles at Alkaline Hot Springs in an Acid Ocean', in J. Wiegel and W. W. Adams (eds), *Thermophiles: The Keys to Molecular Evolution and the Origin of Life*, Taylor and Francis, pp. 77−110.

Reanney, D. C.: 1977, Aminoacyl Thiol Esters and the Origins of Genetic Specificity, *J. Theor. Biol.* **65**, 555−569.

Schimmel, P. and Henderson, B.: 1994, Possible Role of Aminoacyl-RNA Complex in Non-coded Peptide Synthesis and Origin of Coded Synthesis, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 11283−11286.

Segré, D. and Lancet, D.: 1999, A Statistical Chemistry Approach to the Origin of Life, *Chemtracts-Biochem. Mol. Biol.* **12**, 382−397.

Szathmáry, E.: 1993, Coding Coenzyme Handles: A Hypothesis for the Origin of the Genetic Code, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 9916−9920.

Szathmáry, E.: 1999, The Origin of the Genetic Code: Amino Acids as Cofactors in an RNA World, *Trends Genet.* **15**, 223−229.

Tyagi, S.: 1981, Origin of Translation: The Hypothesis of Permanently Attached Adaptors, *Origins Life* **11**, 343−351.

Trifonov, E. N.: 2000, Consensus Temporal Order of Amino Acids and Evolution of the Triplet Code, *Gene* **261**, 139−151.

Wachtershauser, G.: 1988, Before Enzymes and Templates: Theory of Surface Metabolism, *Microbiol. Rev.* **52**, 452−484.

Wachtershauser, G.: 1992, Groundworks for an Evolutionary Biochemistry: The Iron-Sulphur World, *Prog. Biophys. Mol. Biol.* **58**, 85−201.

White III, H. B.: 1976, Coenzymes as Fossils of an Earlier Metabolic State, *J. Mol. Evol.* **7**, 101−104.

White, H. B.: 1982, 'Evolution of Coenzymes and the Origin of Pyridine Nucleotides', in J. Everse, B. Anderson and K. S. You (eds), *The Pyridine Nucleotide Coenzymes*, Academic Press, pp. 1−17.

Woese, C. R.: 1998, The Universal Ancestor, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6854−6859.

Woese, C. R.: 2000, Interpreting the Universal Phylogenetic Tree, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 8392−8396.

Wong, J. T.: 1975, A Co-evolution Theory of the Genetic Code, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 1909−1912.

Wong, J.T.: 1991, Origin of Genetically Encoded Protein Synthesis: A Model Based On Selection for RNA Peptidation, *Orig. Life Evol. Biosph.* **21**, 165−176.

Wong, J. T. and Xue, H.: 2002, 'Self-perfecting Evolution of Heteropolymer Building Blocks and Sequences as the Basis of Life', in *Fundamentals of Life*, Editions Scientifiques et Medicales Elsevier SAS.