J Mol Evol (2002) 55:616–622
DOI: 10.1007/s00239-002-2357-6

JOURNAL OF MOLECULAR EVOLUTION

Genetic Code Origin: Are the Pathways of Type Glu-tRNA^{Gln} → Gln-tRNA^{Gln} Molecular Fossils or Not?

Massimo Di Giulio

International Institute of Genetics and Biophysics, CNR, Via G. Marconi 10, 80125 Naples, Italy

Received: 28 December 2001 / Accepted: 8 March 2002

Abstract. A logical-evolutionary analysis is conducted to clarify whether or not pathways of type Glu-tRNA^{Gln} → Gln-tRNA^{Gln} are molecular fossils of the mechanism that gave rise to the evolutionary organization of the genetic code. The result of this analysis is that these pathways are most likely a manifestation of this mechanism. This provides strong evidence in favor of the coevolution theory of genetic code origin, as this theory is based on the amino acid biosynthetic transformation taking place on tRNA-like molecules which imprinted the genetic code structuring. Comments on the different interpretations of these pathways found in the literature are also provided.

Key words: Precursor/product amino acids — Unusual biosynthetic pathways — Molecular fossils — Coevolution theory — genetic code origin — RNA world

The Coevolution Theory of Genetic Code Origin and the Aims of the Analysis

The coevolution theory of genetic code origin (Wong 1975) postulates that during code evolution there was a phase in which only a few amino acids (the precursors) were codified. As the other amino acids (the products) evolved from these through biosynthetic pathways, part or all of the codon domain of the precursor amino acids was conceded to the product amino acids (Wong 1975). Therefore, the theory suggests an evolutionary origin of genetic code organization, determined by an imprinting of the biosynthetic pathways linking the amino acids on the code organization (Wong 1975).

Although the statistical foundations of this theory have recently been questioned (Ronneberg et al. 2000), they nevertheless seem to be robust, as has been shown in another analysis (Di Giulio 2001a). Moreover, the quantity of data which seem to corroborate the coevolution theory is so large (Dillon 1973; Wong 1975, 1976, 1980, 1988; McClendon 1986; Wachtershauser 1988; Danchin 1989; Taylor and Coates 1989; de Duve 1991; Di Giulio 1992b, 1993, 1996, 1997a–c, 1999, 2000; Di Giulio et al. 1994; Di Giulio and Medugno 1998, 1999, 2000, 2001; Morowitz 1992; Miseta 1989; Edwards 1996; Chaley et al. 1999; Bermudez et al. 1999; Tumbula et al. 2000) that we are seriously led to consider this theory as one of the best hypotheses at our disposal to explain the origin of genetic code structuring.

This theory claims that the mechanism leading the precursor amino acids to concede part or all of their codon domain to the product amino acids was based on tRNA-like molecules on which the theory envisages that the precursor–product amino acid biosynthetic transformation must have taken place (Wong 1975). In other words, if the precursor–product amino acid transformation did take place on tRNA-like molecules, it would not be difficult to imagine how the precursor amino acid was able to concede some codons from its domain to the product amino acid, because the latter was automatically charged on

Correspondence to: Dr. M. Di Giulio; *email:* digiulio@iigb.na.cnr.it

Table 1. The pathways currently taking place on tRNAs and transforming one amino acid into another, together with their phylogenetic distribution[a]

Pathway	Phylogenetic distribution
Glu-tRNA^{Gln} → Gln-tRNA^{Gln}	Bacteria and Archaea
Asp-tRNA^{Asn} → Asn-tRNA^{Asn}	Bacteria (present in a minority) and Archaea
Ser-tRNA^{Sec} → Sec-tRNA^{Sec}	Bacteria, Archaea, and Eucarya
Met-tRNA^{fMet} → fMet-tRNA^{fMet}	Bacteria and organelles

[a] For the original references, see Ibba et al. (1997, 2000) and Tumbula et al. (2000).

a tRNA that recognized codons belonging to the precursor.

We are therefore somewhat surprised to find that today's organisms contain pathways of the type Glu-tRNA^{Gln} → Gln-tRNA^{Gln} (Ibba et al. 1997, 2000; Tumbula et al. 2000) (Table 1), which indeed seem to be a manifestation of the mechanism on which the coevolution theory is founded (Wong 1976, 1988; Danchin 1989; Di Giulio 1993, 1997a–c, 1999, 2000). Establishing whether or not these pathways actually are molecular fossils of the mechanism that gave rise to genetic code structuring is a critically important point. This is because, as molecular fossils, these pathways would provide evidence of inestimable value in that these fossils would have a high content of the history of the early phases of life on earth. Therefore the aim of the present paper is to conduct a logical-evolutionary analysis to ascertain whether or not the pathways of type Glu-tRNA^{Gln} → Gln-tRNA^{Gln} are molecular relics of the mechanism that gave rise to genetic code organization and comment on some interpretations of the meaning of these pathways.

The Pathways Transforming One Amino Acid into Another and Taking Place on a tRNA

All these pathways (Table 1) share the fact that an amino acid is charged onto a tRNA specific for another amino acid and that one biosynthetic step [in the case of selenocysteine (Sec) more than one biosynthetic step (Baron and Bock 1995)] transforms the first amino acid into the second. After this biosynthetic transformation the tRNAs charged in this unusual way can take part in protein synthesis.

Given the analogy among all these transformations (Table 1), it seems clear that they should be the manifestation of a single mechanism which determined them. This seems clear if we consider the alternative hypotheses that these pathways are the result of a number of different mechanisms, that is, they were produced at different times and under different selective pressures, as all four of these pathways (Table 1) ultimately affect a single fundamental process: protein synthesis. In other words, it would be extremely curious if all these pathways were not the manifestation of a single mechanism which gave rise to them, as they have a single final effect, namely, that of affecting protein structure.

Pathways Encountering Difficulty in Their Evolution: How Did They Ever Evolve?

Clearly there must have been a very important reason why these pathways (Table 1) came to exist. That is, the selective pressure leading to the evolution of these pathways must have been very strong because, otherwise, evolving these four pathways independently without a more than plausible reason would have been practically impossible. Yet this important reason apparently seems to elude us because, normally, the task of charging a given amino acid onto a specific tRNA is universally performed by an aminoacyl-tRNA synthetase. Therefore, although we expect to find a clear evolutionary reason for the existence of these pathways, it is not at all evident, if not indeed obscure, and we are unable to find an immediate justification for it (Di Giulio 2000).

However, Poole et al. (1998) suggest an interesting explanation for the Glu-tRNA^{Gln} → Gln-tRNA^{Gln} pathway based on the acquisition of thermophily. They claim that as Gln is unstable at high temperatures, this should limit the rate of protein synthesis, and the solution for thermophiles/hyperthermophiles was to charge Glu on tRNA^{Gln} and therefore transaminate the Glu immediately prior to incorporating Gln into proteins. Clearly an identical consideration could be made for the pathway Asp-tRNA^{Asn} → Asn-tRNA^{Asn}, as Asn is also unstable at high temperatures (Greenstein and Winitz 1961). Poole et al. (1998) consider these two pathways, and therefore the mischarging, as a biochemical adaptation to extreme conditions and not as an ancestral trait. This hypothesis makes a simple prediction: if these pathways actually were the result of an adaptation to high temperatures, they should not be present in mesophile organisms phylogenetically far from thermophiles. This does not appear to be the case, as, for instance, the Glu-tRNA^{Gln} → Gln-tRNA^{Gln} pathway is present throughout the Bacteria and Archaea domains and not just in thermophile organisms (Table 1). Hence, it is only under the hypothesis that the root of the uni-

versal tree is situated in the Eukarya domain that the authors' hypothesis would be true, and even then, only with a 50% probability. In the other two possible cases of universal tree rooting, this pathway would be a trait belonging to the last universal common ancestor (LUCA) and, therefore ancestral and not derived as Poole et al. (1998) claim; that is, unless this pathway evolved twice and independently in the Bacteria and Archaea domains, which is highly improbable. More generally, this hypothesis would not explain the presence of the pathways for the other two amino acids, which seem to be stable at high temperatures, or the involvement of different tRNAs charged with different amino acids in the biosynthesis of different molecules (see Table 1 of Di Giulio 1997a). All this therefore defines these pathways as probable molecular fossils (Di Giulio 1997a).

The entire issue becomes even more complex if we consider that under the hypothesis that these pathways are an acquired trait (Poole et al. 1998; Ardell and Sella 2001), their evolution would have been greatly hindered (Di Giulio 2000). If the first step in evolving these pathways consisted of charging an amino acid onto a tRNA specific for another amino acid, this operation would produce a mischarged tRNA. And this would be lethal if it went into ribosomes because it would insert one amino acid instead of another into all the proteins and would therefore be highly selected against. Even if the evolution of these pathways followed a retrograde evolution, there would in any case be the appearance of a tRNA charged with an inappropriate amino acid, the effects of which would almost certainly be lethal. Therefore, the evolution of these pathways in protein synthesis as we know it would be extremely difficult, if not impossible.

We thus come to an apparently paradoxical conclusion: How did pathways encountering difficulty in their evolution ever evolve? (Di Giulio 2000). It is clear that to remove these difficulties we need simply postulate that these pathways reflect an ancestral metabolic state. In other words, these pathways are not an acquired trait (Poole et al. 1998; Ardell and Sella 2001) but a very ancient feature, i.e., they are molecular fossils of a metabolic stage (Wong 1976, 1988; Wachtershauser 1988; Danchin 1989; Benner et al. 1989; de Duve 1991; Di Giulio 1992a, 1993, 1997a, b, 1999, 2000; Edwards 1996).

Phylogenetic Considerations

The presence of the Ser-tRNA^{Sec} $\rightarrow$ Sec-tRNA^{Sec} pathway in all three main lines of divergence (Table 1) makes it possible to attribute this pathway to the LUCA. It is acknowledged that Sec (selenocysteine) is the 21st amino acid codified in the genetic code (Bock et al. 1991b); that is, that there was probably a time when Sec was commonly incorporated in proteins. However, the emission of oxygen into the atmosphere is thought to have created a strong selective pressure aiming to eliminate Sec from proteins, given its strong susceptibility to oxidation (Bock et al. 1991a, b). Therefore, we must logically conclude that this pathway could be related to genetic code origin because (1) phylogenetic considerations attribute the existence of this pathway to the LUCA, therefore attesting to its clear antiquity; moreover, given that (2) a mechanism affecting the genetic code is involved, it is natural to think that this might be the expression of the mechanism promoting genetic code structuring; and, finally, (3) in light of the RNA world hypotheses (White 1976; Gilbert 1986; Benner et al. 1989; Gibson and Lamond 1990; Di Giulio 1997a, c, 2000; Poole et al. 1998), this pathway can certainly be attributed to certain phases of the origin of life and therefore to genetic code origin (Di Giulio 1997a, c).

The Glu-tRNA^{Gln} $\rightarrow$ Gln-tRNA^{Gln} pathway is present in the Archaea and Bacteria domains (Table 1) and it is therefore more likely that the LUCA also had this pathway, although if the universal tree were rooted in the Eukarya domain, the LUCA might or might not have had this pathway with an equal probability. A substantially equivalent conclusion can be reached for the pathway Asp-tRNA^{Asn} $\rightarrow$ Asn-tRNA^{Asn}, although this pathway has been identified only in the Archaea domain and in few organisms of the Bacteria domain (Table 1). However, this pathway should be refractory to horizontal transfer and it is thus as if it were present throughout the Bacteria domain, given the implausibility of its independent origin from the same pathway present in the Archaea domain. Finally, as the pathway Met-tRNA^{fMet} $\rightarrow$ fMet-tRNA^{fMet} is present only in the Bacteria domain, it is most probably a particular feature of this domain, also in consideration of the major differences among the three domains with regard to the initiation of translation (Guarlerzi and Pon 1990; Pain 1996; Kozak 1999; Di Giulio 2001b). We must therefore conclude that this pathway might not have been present in the LUCA. Nevertheless, it was certainly present in the ancestor of the Bacteria domain and is thus very ancient and might still reflect the mechanism that led to the structuring of the genetic code, although expressed in a more advanced phase (Di Giulio 2001b).

In conclusion, the phylogenetic considerations on these pathways are, on the whole, such as to imply their presence in the LUCA, thus attesting to their certain antiquity, which, in addition to the mechanism through which they operate and the difficulties that would derive under the hypothesis of their recent origin, would truly seem to imply an involvement of

these pathways in the origin of genetic code structuring.

Conclusions

On the basis of the comments made in the previous sections we must reasonably conclude that these pathways (Table 1) had something to do with genetic code origin. This conclusion, although debatable, is the best we have to rationalize the evolution of these pathways. The alternative hypothesis, which sees these pathways as derived traits (Poole et al. 1998; Ardell and Sella 2001), is certainly untrue for the Ser-tRNA^{Sec} → Sec-tRNA^{Sec} pathway (Table 1), as it is for the pathways involving the two amides (Table 1), because all these pathways were most likely present in the LUCA (see previous sections) and cannot therefore, by definition, be a derived trait. However, the hypothesis which sees these pathways as being unrelated to genetic code origin, while acknowledging their antiquity (Cavalier-Smith 2001), clashes with (1) the intrinsic difficulty in evolving these pathways in a protein synthesis apparatus which is already complex and (2) the unclear immediacy of a strong selective pressure aiming to trigger their evolution, as well as (3) the mechanism through which these pathways operate, which is such as naturally to suggest its connection to genetic code origin, as, moreover, envisaged by the coevolution theory (Wong 1975). All these difficulties are simultaneously removed if we postulate that these pathways are molecular fossils of a metabolic state, i.e., they are a manifestation of the mechanism through which the genetic code structuring was originated.

Comments on the Different Evolutionary Interpretations of These Pathways

Ardell and Sella (2001, p. 271) recently echoed the assertions of Poole et al. (1998) regarding these pathways, that is, "This particular case is almost certainly a derived rather than an ancestral condition." In light of the comments made in the previous sections, we can confidently say that this is simply untrue. Moreover, the authors (Ardell and Sella 2001, p. 271) say that "if misacylation is a reasonable model for a proximal mechanism of code change through duplication and divergence, then this evidence would seem to favor the encoding of novel amino acids that are stereochemically or physicochemically related, rather than metabolically related to an ancestral ligand." This is simply paradoxical. These pathways (Table 1) are the most evident manifestation of the mechanism on which the coevolution theory is founded (Wong 1975, 1988; Di Giulio 2000) because they establish a link between the codon domain of the precursor amino acid and the codons of the product (see also the introductory section). Even if physicochemically similar amino acids are involved in these transformations, this is no evidence in favor of either the stereochemical or the physicochemical theory. The stereochemical theory (Woese 1965, 1967; Nelsestuen 1978; Balasubramanian et al. 1980; Shimizu 1982; Yarus 1998) suggests that genetic code origin took place through direct interaction between anticodons (or codons) and the corresponding amino acids. Therefore, this interaction is clearly contradicted by these pathways because, if we assume that they manifest the mechanism that structured the genetic code, they do not contain any interaction between anticodons and amino acids, as envisaged by the stereochemical hypothesis, but only a metabolic transformation of one amino acid into another, which contradicts this interaction. The postulates of the physicochemical theory (Woese et al. 1966; Fitch and Upper 1987) claim that the physicochemical properties of amino acids must be highly optimized inside the genetic code table in that similar amino acids must be codified by codons that are as similar as possible, which does not currently seem to be the case (Wong 1980; Di Giulio et al. 1994; Di Giulio and Medugno 1998, 1999, 2001). It is also worth recalling that the coevolution theory is compatible, i.e., it does not deny that similar amino acids can have somehow similar codons (Wong 1980). What the latter theory does deny is that the optimization level of the physicochemical properties of amino acids in the genetic code is very high (Wong 1980), which, as we have already said, does not currently seem to be the case.

Cavalier-Smith (2001, p. 585) says that "there is no evidence from the structure of the code that glutamine or asparagine captured glutamate or aspartate codons. Phylogenetic data also do not support such an explanation for either amino acid." This is simply not true. If the codons of Gln and Asn codified, as envisaged by the coevolution theory, in a phase of code development for Glu and Asp, respectively, then this simple postulate is such as to remove, in the genetic code, all the noncontiguities between amino acids in precursor–product relationships (Wong 1975, p. 1909). Therefore, the opposite of what Cavalier-Smith (2001) says is true: the coevolution theory is based on this very fact and the genetic code structure provides support in favor of this interpretation. It is therefore highly significant for the coevolution theory that these two pathways have been identified for these two pairs of amino acids in precursor–product relationships and not for other pairs. The sentence implying phylogenetic arguments is also very dubious, if not simply untrue. This is because, as we have seen, the pathways involving the two amides were most likely present in the LUCA, and therefore, regardless

of the phylogenetic considerations of Cavalier-Smith on the distribution of glutaminyl-tRNA synthetase and asparaginyl-tRNA synthetase, these pathways are very ancient and might well have been involved in the structuring of the genetic code. Moreover, the presence of glutaminyl-tRNA synthetase in the LUCA is not certain on the basis of current data (Lamour et al. 1994; Handy and Doolittle 1999) and it might have originated in the Eukarya domain and only later have horizontally transferred to the Bacteria domain (Lamour et al. 1994; Handy and Doolittle 1999). The latter point thus stresses that the pathway involving Gln might have had something directly to do with the origin of the genetic code, as the corresponding aminoacyl-tRNA synthetase might have been absent during this origin.

Cavalier-Smith (2001, p. 584) says, "Selenocysteine almost certainly first became encoded by the modification of serine while bonded to one of its tRNAs; it is possible the only amino acid added to the code in this way which Wong (1975) postulated was the predominant method." And also, "There is no reason to suppose that the UGA codon was ever used by serine: serine simply provided a tRNA that could mutate to selenocysteinyl-tRNA to capture it" (Cavalier-Smith 2001, p. 586). There is something in this that is simply incomprehensible, as it does not want to accept, in the absence of a truly alternative hypothesis, that this example strongly corroborates the coevolution theory (Wong 1976, 1988; Di Giulio 1992a, 1993, 1997b, 1999, 2000). The presence of the pathway leading to Ser-tRNA^{Sec} $\rightarrow$ Sec tRNA^{Sec} in the three main phyletic lines (Table 1) indicates that this was present in the LUCA. Moreover, on the basis of the comments made in the previous sections, the pathways involving Gln and Asn seem to be equally ancient as the pathway involving Sec, and this suggests that the same mechanism that allowed Gln and Asn to enter the genetic code was most probably also responsible for the entry of Sec; and since in the case of Gln and Asn there are no termination codons involved, this indicates that the codon UGA might possibly, but not necessarily, have codified for another amino acid (i.e., for Ser) before codifying for Sec and thus for termination. If all this is hypothetical, then Cavalier-Smith's arguments are equally hypothetical, but we cannot consider it as such because it is inserted in the framework of a theory (Wong 1975) that is supported by an enormous number of data (Dillon 1973; Wong 1975, 1976, 1980, 1988; McClendon 1986; Wachtershauser 1988; Danchin 1989; Taylor and Coates 1989; de Duve 1991; Di Giulio 1992b, 1993, 1996, 1997a–c, 1999, 2000; Di Giulio et al. 1994; Di Giulio and Medugno 1998, 1999, 2000, 2001; Morowitz 1992; Miseta 1989; Edwards 1996; Chaley et al. 1999; Bermudez et al. 1999; Tumbula et al. 2000). Moreover, as shown here, these pathways (Table 1) seem to have all the requisites of molecular fossils (Di Giulio 1997a) and so my interpretation is not really so hypothetical. More generally, here we find ourselves facing a problem often encountered in evolutionary biology, namely, that our convictions are based more on feeling that on truly rigorous proofs.

Cavalier-Smith (2001, p. 586) says, "The codon correlations with biosynthetic pathways are robust (Taylor and Coates 1989; Di Giulio and Medugno 2000) but, if as I have argued, this correlation is simply caused by conservation of amino acid recognition patterns during codon capture, it would be misleading to refer to it as being explained by the coevolution theory." Such claims are simply false. The conservation of amino acid recognition patterns can explain neither the correlation nor, above all, the specificity of the correlation existing between the biosynthetic pathways linking amino acids and the attribution to them of specific codons in the genetic code (Wong 1975; Taylor and Coates 1989; Di Giulio 2001a). This is because these patterns should be in relation to the amino acids belonging to the columns of the genetic code, as it is here that physicochemical properties of amino acids are better allocated (Nelsestuen 1978; Wolfenden et al. 1979; Sjostrom and Wold 1985; Di Giulio 1989; Taylor and Coates 1989), whereas the biosynthetic relationships are mostly distributed over the rows of the code (Taylor and Coates 1989). This therefore defines a generally opposite behavior between these conservation patterns and the biosynthetic relationships between amino acids, which suggests that Cavalier-Smith's explanation is most probably untrue because, otherwise, this behavior should have been at least parallel. And even if it is known that amino acids in precursor–product relationships have, on average, similar physicochemical properties (Di Giulio 1992b, 1996), this cannot explain the relationship between biosynthetic pathways linking amino acids and the organization of the genetic code because this relationship is, moreover, too specific and must therefore be the result of an equally specific mechanism (like the one envisaged by the coevolution theory), not unspecific as predicted by the amino acid recognition patterns through the mechanism based on the codon capture hypothesis. The latter mechanism (Osawa and Jukes 1988) cannot explain the relationship between the biosynthetic pathways between amino acids and genetic code organization because it establishes no real connection between the codons of amino acids in a biosynthetic relationship. In other words, the logic of progressive codon capture used by Cavalier-Smith (2001) cannot explain the coevolution between the biosynthetic pathways of amino acids and the genetic code

because it is not based on a mechanism which establishes a true connection between the codons of the amino acids in biosynthetic relationships and cannot therefore imply that these relationships are then reflected in the genetic code. More specifically, a product amino acid evolving along a biosynthetic pathway having another amino acid as its precursor must not occupy, on the basis of the codon capture hypothesis, a codon contiguous to the precursor, but it can take any codon that became available at that evolutionary moment, thus not necessarily in relation to the codons of that precursor. It therefore seems to me quite literally impossible that the codon capture mechanism can take into account the correlation and, above all, the specificity of the relationship between the biosynthetic pathways linking amino acids and the organization of the genetic code, as claimed by Cavalier-Smith (2001).

References

Ardell DH, Sella G (2001) On the evolution of redundancy in the genetic codes. J Mol Evol 53:269–281

Balasubramanian R, Seetharamulu P, Raghunathan G (1980) A conformational rationale for the origin of the mechanism of nucleicacid-directed protein synthesis of 'living' organisms. Origins Life 10:15–30

Baron C, Bock A (1995) The selenocysteine-inserting tRNA species: Structure and function. In: Soll D, RajBhandary UL (ed) tRNA structure, biosynthesis, and function. ASM Press, Washington, DC, pp 529–544

Benner SA, Ellington AD, Tauer A (1989) Modern metabolism as a palimpsest of the RNA world. Proc Natl Acad Sci USA 86:7054–7058

Bermudez CI, Daza EE, Andrade E (1999) Characterization and comparison of *Escherichia coli* transfer RNAs by graph theory based on secondary structure. J Theor Biol 197:193–205

Bock A, Forchhammer K, Heider J, Baron C (1991a) Selenoprotein synthesis: an expansion of the genetic code. Trends Biochem Sci 16:463–467

Bock A, Forchhammer K, Heider J, Leinfelder W, Sawers G, Veprek B, Zinoni (1991b) Selenocysteine: The 21st amino acid. Mol Microbiol 5:515–520

Cavalier-Smith T (2001) Obcells as proto-organisms: Membrane heredity, lithophosphorilation, and the origins of the genetic code, the first cell, and photosynthesis. J Mol Evol 53:555–595

Chaley MB, Korotkov EV, Phoenix DA (1999) Relationships among isoacceptor tRNA seems to support the coevolution theory of the origin of the genetic code. J Mol Evol 48:168–177

Danchin A (1989) Homeotopic transformation and the origin of translation. Prog Biophys Mol Biol 54:81–86

de Duve C (1991) Blueprint for a cell: The nature and origin of life. Neil Patterson, Carolina Biological Supply Company, Burlington, NC, pp 175–181

Di Giulio M (1989) Some aspects of the organization and evolution of the genetic code. J Mol Evol 29:191–201

Di Giulio M (1992a) On the origin of the genetic code. Trends Ecol Evol 7:176–178

Di Giulio M (1992b) The evolution of aminoacyl-tRNA synthetases, the biosynthetic pathways of amino acids and the genetic code. Origins Life Evol Biosph 22:309–319

Di Giulio M (1993) Origin of glutaminyl-tRNA synthetase: An example of palimpsest? J Mol Evol 37:5–10

Di Giulio M (1996) The β-sheets of proteins, the biosynthetic relationships between amino acids, and the origin of the genetic code. Origins Life Evol Biosph 26:589–609

Di Giulio M (1997a) On the origin of the genetic code. J Theor Biol 187:573–581

Di Giulio M (1997b) The origin of the genetic code. Trends Biochem Sci 22:49

Di Giulio M (1997c) On the RNA world: Evidence in favor of an early ribonucleopeptide world. J Mol Evol 45:571–578

Di Giulio M (1999) The coevolution theory of the origin of the genetic code. J Mol Evol 48:253–254

Di Giulio M (2000) The RNA world, the genetic code and the tRNA molecule. Trends Genet 16:17–18

Di Giulio M (2001a) A blind empiricism against the coevolution theory of the origin of the genetic code. J Mol Evol 53:724–732

Di Giulio M (2001b) The non-universality of the genetic code: The universal ancestor was a progenote. J Theor Biol 209:345–349

Di Giulio M, Medugno M (1998) The historical factor: the biosynthetic relationships between amino acids and their physicochemical properties in the origin of the genetic code. J Mol Evol 46:615–621

Di Giulio M, Medugno M (1999) Physicochemical optimization in the genetic code origin as the number of codified amino acids increases. J Mol Evol 49:1–10

Di Giulio M, Medugno M (2000) The robust statistical bases of the coevolution theory of the genetic code. J Mol Evol 50:258–263

Di Giulio M, Medugno M (2001) The level and landscape of optimization in the origin of the genetic code. J Mol Evol 52:372–382

Di Giulio M, Capobianco MR, Medugno M (1994) On the optimization of the physicochemical distances between amino acids in the evolution of the genetic code. J Theor Biol 186:43–51

Dillon LS (1973) The origins of the genetic code. Bot Rev 39:301–345

Edwards MR (1996) Metabolite channeling in the origin of life. J Theor Biol 179:313–322

Fitch W, Upper K (1987) The phylogeny of tRNA sequences provides evidence for ambiguity reduction in the origin of the genetic code. Cold Spring Harbor Symp Quant Biol 52:759–767

Gibson TJ, Lamond AI (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

Greenstein JP, Winitz M (1961) Glutamic acid and glutamine. In: Chemistry of amino acids, Vol 3. John Wiley, New York, pp 1929–1954

Guarlerzi CO, Pon CL (1990) Initiation of mRNA translation in prokaryotes. Biochemistry 29:5881–5889

Handy J, Doolittle RF (1999) An attempt to pinpoint the phylogenetic introduction of glutaminyl-tRNA synthetase among bacteria. J Mol Evol 49:709–715

Ibba M, Curnow AW, Soll D (1997) Aminoacyl-tRNA synthesis: Divergent routes to a common goal. Trends Biochem Sci 22:39–42

Ibba M, Becker HD, Stathopoulos C, Tumbula DL, Soll D (2000) The adaptor hypothesis revisited. Trends Biochem Sci 25:311–316

Kozak M (1999) Initiation of translation in prokaryotes and eukaryotes. Gene 234:187–208

Lamour V, Quevillon S, Diriong S, Guyen VC, Lipinski M, Mirande M (1994) Evolution of the Glx-tRNA synthetase family: The glutaminyl enzyme as a case of horizontal gene transfer. Proc Natl Acad Sci USA 91:8670–8674

McClendon JH (1986) The relationship between the origins of the biosynthetic paths to the amino acids and their coding. Origins Life 16:260–270

Miseta A (1989) The role of protein associated amino acid precursor molecules in the organization of genetic codons. Physiol Chem Phys Med NMR 21:237–242

Morowitz HJ (1992) Beginnings of cellular life: Metabolism recapitulates biogenesis. Yale University, Vail-Ballou Press, Binghamton, NY, pp 160–171

Nelsestuen GL (1978) Amino acid-directed nucleic acid synthesis. A possible mechanism in the origin of life. J Mol Evol 11:109–120

Osawa S, Jukes TH (1988) Evolution of the genetic code as affected by anticodon content. Trends Genet 4:191–198

Pain VM (1996) Initiation of protein synthesis in eukaryotic cells. Eur J Biochem 236:747–771

Poole AM, Jeffares DC, Penny D (1998) The path from the RNA world. J Mol Evol 46:1–17

Ronneberg TA, Landweber LF, Freeland SL (2000) Testing a biosynthetic theory of the genetic code: Fact or artifact? Proc Natl Acad Sci USA 97:13690–13695

Shimizu M (1982) Molecular basis for the genetic code. J Mol Evol 18:297–303

Sjostrom M, Wold S (1985) A multivariate study of the relationship between the genetic code and the physico-chemical properties of amino acids. J Mol Evol 22:272–277

Taylor FJR, Coates D (1989) The code within the codons. BioSystems 22:177–187

Tumbula DL, Becker HD, Chang WZ, Soll D (2000) Domain-specific recruitment of amide amino acids for protein synthesis. Nature 407:106–110

Wachtershauser G (1988) Before enzymes and templates: theory of surfage metabolism. Microbiol Rev 52:452–484

White III HB (1976) Coenzymes as fossils of an earlier metabolic state. J Mol Evol 7:101–104

Wong JT (1975) A co-evolution theory of the genetic code. Proc Natl Acad Sci USA 72:1909–1912

Wong JT (1976) The evolution of a universal genetic code. Proc Natl Acad Sci USA 73:2336–2340

Wong JT (1980) Role of minimization of chemical distances between amino acids in the evolution of the genetic code. Proc Natl Acad Sci USA 77:1083–1086

Wong JT (1988) Evolution of the genetic code. Microbiol Sci 5:174–182

Woese CR (1965) On the origin of the genetic code. Proc Natl Acad Sci USA 54:1546–1552

Woese CR (1967) The genetic code. Harper & Row, New York

Woese CR, Dugre DH, Dugre SA, Kondo M, Saxinger WC (1966) On the fundamental nature and evolution of the genetic code. Cold Spring Harbor Symp Quant Biol 31:723–736

Wolfenden RV, Cullis PM, Southgate CCF (1979) Water, protein folding, and the genetic code. Science 206:575–577

Yarus M (1998) Amino acids as RNA ligands: A direct-RNA-template theory for the code's origin. J Mol Evol 47:109–117