Acta Biochimica Polonica

Vol. 48 No. 2/2001

313–321

QUARTERLY

*Review*

# Genomics and the evolution of aminoacyl-tRNA synthesis★✪

Benfang Ruan, Ivan Ahel, Alex Ambrogelly, Hubert D. Becker, Shipra Bunjun, Liang Feng, Debra Tumbula-Hansen, Michael Ibba[1], Dragana Korencic, Hiroyuki Kobayashi, Clarisse Jacquin-Becker, Nina Mejlhede[1], Bokkee Min, Gregory Raczniak, Jesse Rinehart, Constantinos Stathopoulos, Tong Li and Dieter Söll[2,3]✉

*Departments of Molecular Biophysics and Biochemistry; [2]Chemistry; and, [3]Molecular, Cellular and Developmental Biology; Yale University, New Haven, Connecticut, U.S.A.;*
*[1]Center for Biomolecular Recognition, Department of Medical Biochemistry and Genetics, Laboratory B, The Panum Institute, Copenhagen, Denmark*

Received: 14 February, 2001

**Key words:** aminoacyl-tRNA, evolution, protein synthesis, tRNA, translation

**Translation is the process by which ribosomes direct protein synthesis using the genetic information contained in messenger RNA (mRNA). Transfer RNAs (tRNAs) are charged with an amino acid and brought to the ribosome, where they are paired with the corresponding trinucleotide codon in mRNA. The amino acid is attached to the nascent polypeptide and the ribosome moves on to the next codon. Thus, the sequential pairing of codons in mRNA with tRNA anticodons determines the order of amino acids in a protein. It is therefore imperative for accurate translation that tRNAs are only coupled to amino acids corresponding to the RNA anticodon. This is mostly, but not exclusively, achieved by the direct attachment of the appropriate amino acid to the 3′-end of the corresponding tRNA by the aminoacyl-tRNA synthetases. To ensure the accurate translation of genetic information, the aminoacyl-tRNA synthetases must display an extremely high level of substrate specificity. Despite this highly conserved**

★Presented at the International Conference on "Molecular Architecture of Evolution, Primary and Secondary Determinants" Poznań, Poland, October 29–31, 2000.

✪Work in the authors' laboratories is supported by grants from the National Institute of General Medical Sciences (GM 22854 and GM 55674, DS) and the European Commission (QLG-CT-99-00660, MI). M. I. gratefully acknowledges the support of the Alfred Benzon Foundation.

✉Correspondence to: Dr. Dieter Söll, Department of Molecular Biophysics and Biochemistry, Yale University, P.O. Box 208114, 266 Whitney Avenue, New Haven, CT 06520-8114, U.S.A., tel: (203) 432 6200; fax: (203) 432 6202; e-mail: soll@trna.chem.yale.edu

**Abbreviations:** AARSs, aminoacyl-tRNA synthetases; AdT, amidotransferase.

**function, recent studies arising from the analysis of whole genomes have shown a significant degree of evolutionary diversity in aminoacyl-tRNA synthesis. For example, non-canonical routes have been identified for the synthesis of Asn-tRNA, Cys-tRNA, Gln-tRNA and Lys-tRNA. Characterization of non-canonical aminoacyl-tRNA synthesis has revealed an unexpected level of evolutionary divergence and has also provided new insights into the possible precursors of contemporary aminoacyl-tRNA synthetases.**

The ribosomal synthesis of proteins from a messenger RNA (mRNA) template is one of the defining features of the central dogma of molecular biology (Crick, 1970). Proteins are made by the sequential translation of codons into their corresponding amino acids, resulting in the synthesis of a polypeptide whose sequence corresponds to that defined in the respective mRNA. Amino acids are delivered for protein synthesis as aminoacyl-tRNA : translation factor complexes. The identity of an amino acid inserted at a particular position in a nascent polypeptide is determined by two key molecular recognition events: the interaction of the aminoacyl-tRNA anticodon with an appropriate codon in mRNA, and the correct pairing of amino acid and tRNA anticodon in the aminoacyl-tRNA. As a result, the fidelity of protein synthesis is dependent on the presence in the cell of a complete set of correctly aminoacylated tRNAs (reviewed in Ibba & Söll, 1999).

Given that protein synthesis generally makes use of a repertoire of twenty amino acids, the existence of twenty corresponding aminoacyl-tRNA synthetases (AARSs), as predicted by Crick in the adaptor hypothesis (Crick, 1958), was to be expected. This expectation was confirmed by the discovery of the twenty AARSs during the 1960s and early 1970s (reviewed in Söll & Schimmel, 1974; Ibba & Söll, 2000). During this time the first alternative route of aminoacyl-tRNA synthesis was also reported in bacteria (Wilcox & Nirenberg, 1968) and subsequently in archaea (White & Bayley, 1972) and eukaryal organelles (Schön *et al*., 1988). This pathway, which circumvents glutaminyl-tRNA synthetase during glutaminyl-tRNA$^{Gln}$ synthesis (see below), was generally assumed to be no more than a rare evolutionary curiosity. Similarly, the discovery of co-translational insertion of selenocysteine upon suppression of certain in-frame UGA codons (reviewed in Commans & Böck, 1999) was held to be no more than another rare exception to the near-universal use of the 20 canonical AARSs. Recent studies, principally driven by the availability of numerous complete genome sequences, have now shown that contrary to all expectations, numerous organisms do not use a full complement of twenty canonical AARS enzymes to synthesize aminoacyl-tRNAs for protein synthesis. In such cases, a reduced number of canonical AARSs (the minimal known complement is 16; Bult *et al*., 1996; Smith *et al*., 1997) are complemented by a variety of novel enzymes and pathways to provide the full range of aminoacyl-tRNAs necessary for protein synthesis. Here we will describe recent advances in our understanding of these non-canonical routes of aminoacyl-tRNA synthesis. The implications of these findings for our understanding of the evolution of aminoacyl-tRNA synthesis will also be discussed.

## tRNA-DEPENDENT TRANSAMIDATION AS A ROUTE TO Asn-tRNA AND Gln-tRNA

Recent studies reveal that the most common divergence from canonical aminoacyl-tRNA synthesis is exhibited by the tRNA-dependent amino acid transformation pathways (Eqns. 1–4). These two-step, indirect routes to glutaminyl-tRNA$^{Gln}$ (Gln-tRNA$^{Gln}$) and asparaginyl-tRNA$^{Asn}$ (Asn-tRNA$^{Asn}$) generally exist when glutaminyl-tRNA synthetase

(GlnRS) or asparaginyl-tRNA synthetase (AsnRS), respectively, is absent. During Gln-tRNA$^{Gln}$ synthesis (Eqns. 1–2), tRNA$^{Gln}$ is

$$Glu + tRNA^{Gln} + ATP \Leftrightarrow Glu\text{-}tRNA^{Gln} + AMP + PP_i \quad (1)$$

$$Glu\text{-}tRNA^{Gln} + Gln + ATP \Leftrightarrow Gln\text{-}tRNA^{Gln} + Glu + ADP + P_i \quad (2)$$

$$Asp + tRNA^{Asn} + ATP \Leftrightarrow Asp\text{-}tRNA^{Asn} + AMP + PP_i \quad (3)$$

$$Asp\text{-}tRNA^{Asn} + Gln + ATP \Leftrightarrow Asn\text{-}tRNA^{Asn} + Glu + ADP + P_i \quad (4)$$

first misaminoacylated with glutamate by a non-discriminating (relaxed tRNA specificity) glutamyl-tRNA synthetase (GluRS), which, in addition to generating Glu-tRNA$^{Glu}$, can also synthesize Glu-tRNA$^{Gln}$. The resulting mischarged tRNA is then specifically recognized by glutamyl-tRNA$^{Gln}$ amidotransferase (GluAdT, Curnow *et al.*, 1997) and converted into Gln-tRNA$^{Gln}$. Similarly, Asn-tRNA$^{Asn}$ is formed (Eqns. 3–4) *via* a non-discriminating aspartyl-tRNA synthetase (AspRS) and an aspartyl-tRNA$^{Asn}$ amidotransferase (AspAdT, Curnow *et al.*, 1996). For both GluAdT and AspAdT activity, bacteria and archaea use a single, heterotrimeric enzyme encoded by *gatCAB* (Curnow *et al.*, 1998; Tumbula *et al.*, 2000). In addition, archaea possess a unique GluAdT (Tumbula *et al.*, 2000). Thus, the amide aminoacyl-tRNA pathways of the three kingdoms (Bacteria, Archaea and Eukarya)

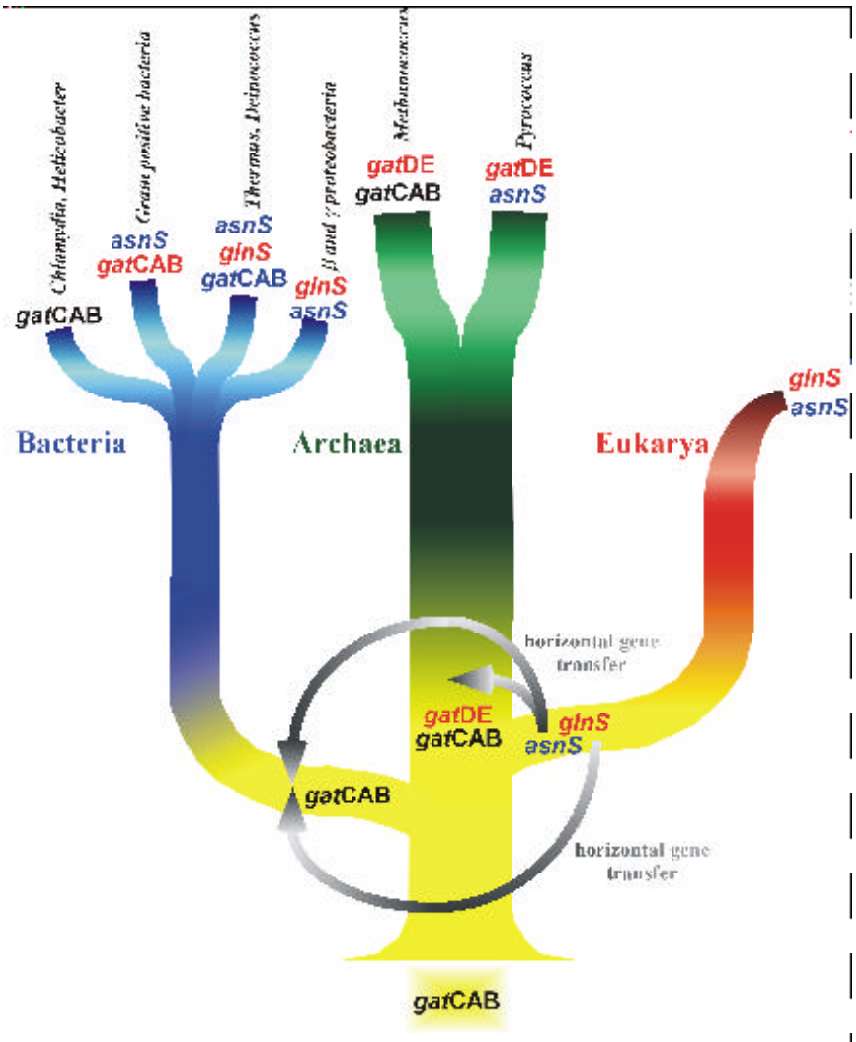


**Figure 1. Phylogenetic distribution of the indirect and direct pathways of amide aminoacyl-tRNA synthesis.**

The common ancestor roots the schematic evolutionary tree. Presence of the direct or indirect routes of asparaginyl-tRNA (asn-tRNA) or glutaminyl-tRNA (gln-tRNA) formation is indicated by the presence of the genes encoding the enzymes responsible for these pathways. *gat*CAB, *gat*DE, tRNA-dependent amidotransferase and indirect route; *asnS*, *glnS*, asparaginyl-tRNA synthetase, glutaminyl-tRNA synthetase and direct route. Blue, Asn-tRNA synthesis. Red Gln-tRNA synthesis. Green, Asn-tRNA and Gln-tRNA synthesis. The different existing combinations of enzymes responsible for amide aminoacyl-tRNA formation are indicated on the top of the different branches of the tree, representative organisms are given.

use different enzymes and mechanisms (Fig. 1). Similarly, kingdom-specific tRNA-dependent amino acid transformation pathways are also responsible for synthesizing selenocysteinyl-tRNA (from Ser-tRNA$^{Sec}$, Commans & Böck, 1999) and formylmethionyl-tRNA (from Met-tRNA$_i^{Met}$, RajBhandary, 1994).

Characterization of the indirect synthetic pathways to Gln-tRNA$^{Gln}$ and Asn-tRNA$^{Asn}$ indicates that they evolved as distinct systems in the three kingdoms. All known examples of eukaryal cytoplasmic protein synthesis use exclusively GlnRS and AsnRS. In contrast, bacteria and eukaryal organelles use predominantly GluAdT and AsnRS (reviewed in Tumbula *et al.*, 2000). Most of the exceptions to this rule seem to result from horizontal gene transfer (i.e., transfer of genes between different organisms). For example, GlnRS has been described only in some proteobacteria (reviewed in Brown & Doolittle, 1999), in the *Thermus*/*Deinococcus* group (Becker & Kern, 1998; Curnow *et al*., 1998) and in the mitochondria of trypanosomatids (Nabholz *et al*., 1997). The current lack of sequence data on this last group restricts speculation on the origin of its GlnRS. Phylogenetic analyses of bacterial GlnRS sequences consistently suggest a recent gene transfer from the eukarya (e.g., Brown & Doolittle, 1999). This, in turn, is suggestive of loss of the indirect pathway in some of these organisms, or for recruitment of GatCAB to Asn-tRNA$^{Asn}$ formation as seen in *Deinococcus radiodurans* and *Thermus thermophilus* (Becker & Kern, 1998; Curnow *et al*., 1998). Most strikingly, the archaea use the indirect transamidation pathway almost exclusively (Tumbula *et al.*, 1999), as GlnRS activity and the corresponding gene have never been found in this kingdom. Only a few archaea have AsnRS genes (Tumbula *et al.*, 2000) (phylogenetic analysis again suggesting horizontal transfer, Woese *et al.,* 2000), whereas the vast majority use the AspAdT pathway characterized in *Haloferax volcanii* (Curnow *et al*., 1996) and *Methanobacterium thermoautotrophicum* (Tumbula *et al*., 2000). While some archaea could potentially use GatCAB for both GluAdT and AspAdT function, each archaeal genome also encodes a second GluAdT enzyme. This heterodimeric enzyme, encoded by the *gatD* and *gatE* genes, is strictly archaeal and not found elsewhere. The purified *M. thermoautotrophicum* GatDE enzyme has GluAdT activity *in vitro* and is unable to form Asn-tRNA (Tumbula *et al*., 2000). The respective roles of the two archaeal GluAdT enzymes have yet to be determined. However, the existence of GatDE in every archaeal genome, even in the presence of GatCAB, suggests a critical function.

## CLASS I-TYPE LYSYL-tRNA SYNTHETASES

The aminoacyl-tRNA synthetases can be divided into two classes (I and II) of ten members each based on the presence of mutually exclusive amino-acid sequence motifs (reviewed in Arnez & Moras, 1997). This division reflects structurally distinct topologies within the active site, class I AARSs containing a Rossmann fold and class II a unique anti-parallel $\beta$ fold. In addition, it has been observed that class I enzymes bind the acceptor helix of tRNA on the minor groove side whereas class II enzymes bind the major groove side. An AARS of particular substrate specificity will always belong to the same class regardless of its biological origin, reflecting the ancient evolution of this enzyme family. The only known exceptions to this rule are the lysyl-tRNA synthetases which are class I enzymes in certain archaea and bacteria but are otherwise members of class II (Ibba *et al.* 1997a; 1997b). This surprising finding arose from studies based upon analysis of the complete genome sequences of the archaea *Methanococcus jannaschii* and *M. thermoautotrophicum*. Sequence-homology searches indi-

cated that only 16 identifiable open reading frames in each of these genomes encoded recognizable AARSs, with AsnRS, cysteinyl-tRNA synthetase (CysRS), GlnRS and lysyl-tRNA synthetase (LysRS) apparently absent from both organisms. Whereas the widespread absence of AsnRS- and GlnRS-encoding genes could be readily explained by their functional replacement by the corresponding transamidation pathways (see above), the absence of CysRS and LysRS was initially more problematic. For LysRS the answer proved to be relatively straightforward, the majority of archaea and a scattering of bacteria being found to contain a previously unidentified class-I-type LysRS. This is in contrast to all previously identified LysRS enzymes which belong to class II, thereby violating the "class rule" of AARS classification. Despite their lack of structural homology, both class I and class II LysRSs are able to recognize the same amino acid and highly similar tRNA substrates, providing an example of functional convergence by divergent enzymes (Ibba *et al*., 1999). Furthermore, the similarity in the tRNA identity sets for the class I and II enzymes may also indicate that $tRNA^{Lys}$ itself predates at least one of the LysRS families (Ribas de Pouplana *et al*., 1998; Ibba *et al*,. 1999).

## A DUAL-SPECIFICITY ProCysRS CONTRADICTS THE ACCEPTED DEFINITION OF AN AARS

The genome sequences of the thermophilic archaea *M. jannaschii* and *M. thermoautotrophicum* do not contain any identifiable genes encoding CysRS proteins, in contrast to the genomes of more than 40 other organisms from all the three kingdoms, which encode canonical class I CysRS enzymes. This apparent discrepancy was resolved by biochemical and genetic studies that showed the enzyme responsible for the formation of Cys-$tRNA^{Cys}$ to be a class II enzyme, prolyl-tRNA synthetase (ProRS) (Stathopoulos *et al*., 2000 and references therein).

The ProRS of *M. jannaschii* (referred to as ProCysRS) can synthesize both Cys-$tRNA^{Cys}$ and Pro-$tRNA^{Pro}$, but not Cys-$tRNA^{Pro}$ or Pro-$tRNA^{Cys}$, *in vitro* and *in vivo*. To date, this is the only known example of a single AARS that can specify two different amino acids in translation. While no organisms outside of the archaea have yet been found to lack a gene encoding a canonical CysRS, the dual function ProRS is not confined to archaea. Molecular phylogenies of ProRS amino-acid sequences suggested that the deep-rooted eukaryon *Giardia lamblia* might also contain a ProRS with CysRS activity. This possibility was subsequently confirmed experimentally, raising the possibility that ProCysRS enzymes may be present in other organisms (Bunjun *et al*., 2000).

The basis of the dual substrate specificity of ProCysRS is related to differences in the mechanisms by which the two aminoacyl-tRNA products are synthesized (Fig. 2). While ProCysRS does not require the presence of $tRNA^{Pro}$ for prolyl-adenylate synthesis, the activation of cysteine is observed only in the presence of $tRNA^{Cys}$ (Stathopoulos *et al*., 2001). The binding and activation of proline facilitate $tRNA^{Pro}$ binding, perhaps at different sites than used by $tRNA^{Cys}$, while simultaneously preventing $tRNA^{Cys}$ binding perhaps by stearic hindrance. In a similar way, when $tRNA^{Cys}$ is bound on the enzyme, activation of proline is blocked (possibly by an allosteric effect) and thus only cysteine can be activated. Although the structural basis for this activity is currently unclear, mutagenesis of active site residues suggests that the binding sites for cysteine and proline overlap (Stathopoulos *et al*., 2001).

How widespread is the distribution of dual specificity enzymes? A recent study of transposon mutagenesis of *Mycoplasma genitalium*, the smallest free-living organism, reported that two aminoacyl-tRNA synthetases may be dispensable as cells were found con-

tain ing in ser tions in the cor re spond ing genes (Hutchison *et al*., 1999). This suggested that in these cells other en zy matic ac tiv i ties sub sti-

knockout strains shows that all 20 canonical AARS in volved in cy to plas mic pro tein syn the sis are es sen tial. Thus, the like li hood of wide

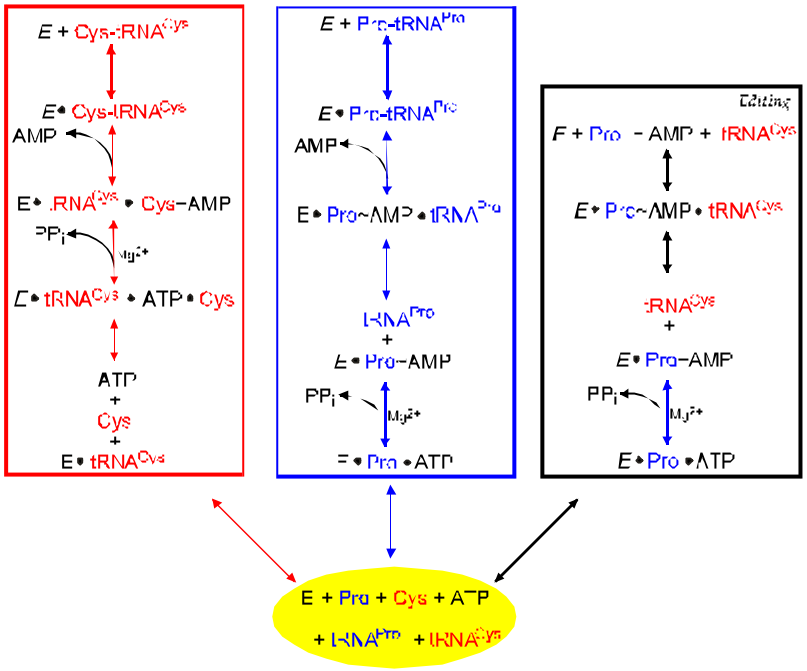


**Fig ure 2. Sche matic rep re sen ta tion of the en zy matic mech a nism of *M. jannaschii* ProCysRS.**

ProCysRS is ab bre vi ated **E**, and the dots rep re sent com plexes be tween the en zyme and any of the sub strates or in ter me di ates of the re ac tion. The red and blue boxes out line the steps lead ing to Cys-tRNA$^{Cys}$ and Pro-tRNA$^{Pro}$ for ma tion re spec tively. The steps out lined by the black box rep re sent the pu ta tive ed it ing mech a nism that we sug gest, to ex plain ab sence of re lease of Pro-tRNA$^{Cys}$ by *M. jannaschii* ProCysRS.

tute for tyrosyl-tRNA synthetase (TyrRS) and isoleucyl-tRNA synthetase (IleRS) function. We there fore at tempted to com ple ment an *E. coli tyrS*$^{ts}$ strain with the other nine teen AARS genes. Since this was not suc cess ful, we took a closer look at the transposon in ser tion data; this analysis suggested that if translation re-initiation after the transposon block oc curred, then gene products of both in ter rupted genes would very likely be functional. TyrRS would be truncated at the N-terminal end by only 22 amino acids but still contain the cru cial HIGH re gion, while the po si tion of transposon in ser tion in *ileS* would give rise to two polypeptide frag ments of an ap prox i mate size (160 and 732 amino acids) that were shown to be active in *trans* complementation studies of *E. coli ileS* fragments (Shiba & Schimmel, 1992). In addition, an examina tion of the phenotypes of the yeast genome

spread occurrence of dual-specificity AARSs is not high.

## EVOLUTION OF CONTEMPORARY AMINO-ACYL-tRNA SYNTHESIS FROM THE RNA WORLD

The need for accurate aminoacyl-tRNA syn thesis during interpretation of the genetic code sug gests an early or i gin for the pro cess dur ing the evo lu tion of con tem po rary gene ex pres sion. Con se quently, the pro cess by which aminoacyl-tRNA synthesis arose and devel oped into one of the most accurate and spe cific func tions in the cell has been the sub ject of much speculation (e.g. Schimmel & Kelly, 2000). The experimental description of nu mer ous cat a lytic RNAs (both nat u ral and syn thetic) and the subsequent formulation and

refinement of the "RNA World" hypothesis (reviewed in Gesteland *et al.*, 1999) has provided the basis for much of this speculation. For example, the description of RNAs able to catalyze their own aminoacylation suggested an obvious origin for the evolution of contemporary aminoacyl-tRNA synthesis (Illangasekare *et al.*, 1995), although other experimental support for such ideas has remained scarce. For instance, the proposal that tRNA (or its precursor) preceded the evolution of aminoacyl-tRNA synthesis (Maizels & Weiner, 1994) only recently found indirect support from biochemical and phylogenetic studies of class I LysRSs (Ribas de Pouplana *et al.*, 1998; Ibba *et al.*, 1999). More recently, further direct support for primordial RNA-based aminoacyl-tRNA synthesis was provided by the synthesis of a ribozyme able to recognize an activated amino acid specifically and transfer it to the 3' end of a tRNA (Lee *et al.*, 2000). Such ribozymes can effectively act *in trans* as *bona fide* AARSs, providing a potential link between the contemporary "protein world" and a primordial "RNA world". Perhaps even more significantly, these two examples would seem to provide a "prologue" (*trans*-acting ribozymes) and an "epilogue" (AARSs that arose after tRNAs) for many of the postulated scenarios for the evolution of contemporary aminoacyl-tRNA synthesis. One such scenario is summarized below. In the RNA world, ribozymes existed that could catalyze aminoacylation using a limited repertoire of amino acid substrates. The specificity of these enzymes then expanded with the recruitment of protein moieties, which allowed the activation of an ever wider range of amino acids. Within such ribonucleoprotein complexes the protein moieties would, by providing an enhanced range of functional groups, gradually take over the catalytic function of aminoacylation. This would then give rise to the first primarily protein-based AARSs, each of which could probably specify more than one amino acid in a primitive protein synthesis machinery (Delarue, 1995). Duplication and diversification of these primitive synthetases would then form the basis for the evolutionary radiation that gave rise to the contemporary AARSs. Until recently, the final transition in this scheme, from synthetases of broad to narrow substrate specificity, while suggested by AARS phylogenies (Nagel & Doolittle, 1995), was not supported by any experimental findings. However, recent data support the validity of this proposed transition. For example, studies in archaea have shown that, at least the canonical cysteinyl-tRNA synthetase is not essential for viability (I. Anderson, W. Whitman, C. Stathopoulos and T. Li, unpublished), suggesting that particular AARS enzymes may synthesize more than one aminoacyl-tRNA. One such enzyme has now been identified, the archaeal genre ProCysRS, and its further characterization can be expected to provide unprecedented insights into the evolution of extant aminoacyl-tRNA synthetases.

---

We thank Clyde Hutchison and Scott Peterson for discussions.

## REFERENCES

Arnez, J.G. & Moras, D. (1997) Structural and functional considerations of the aminoacylation reaction. *Trends Biochem. Sci.* **22**, 211–216.

Becker, H.D. & Kern, D. (1998) *Thermus thermophilus*: a link in evolution of the tRNA-dependent amino acid amidation pathways. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 12832–12837.

Brown, J.R. & Doolittle, W.F. (1999) Gene descent, duplication, and horizontal transfer in the evolution of glutamyl- and glutaminyl-tRNA synthetases. *J. Mol. Evol.* **49**, 485–495.

Bult, C.J., White, O., Olsen, G.J., Zhou, L., Fleischmann, R.D., Sutton, G.G., Blake, J.A., FitzGerald, L.M., Clayton, R.A., Gocayne, J.D., Kerlavage, A.R., Dougherty, B.A., Tomb,

J.F.,Adams, M.D., Reich, C.I., Overbeek, R., Kirkness, E.F., Weinstock, K.G., Merrick, J.M., Glodek, A., Scott, J.L., Geoghagen, N.S.M. & Venter, J.C. (1996) Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science* **273**, 1058–1073.

Bunjun, S., Stathopoulos, C., Graham, D., Min, B., Kitabatake, M., Wang, A.L., Wang, C.C., Vivarès, C.P., Weiss, L.M. & Söll, D. (2000) A dual-specificity aminoacyl-tRNA synthetase in the deep-rooted eukaryote *Giardia lamblia*. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 12997–13002.

Commans, S. & Böck, A. (1999) Selenocysteine inserting tRNAs: An overview.*FEMS Microbiol. Rev.* **23**, 335–351.

Crick, F.H.C. (1958) On protein synthesis. *Symp. Soc. Exp. Biol.* **12**, 138–163.

Crick, F.H.C. (1970) Central dogma of molecular biology. *Nature* **227**, 561–563.

Curnow, A. W., Ibba, M. & Söll, D. (1996) tRNA-dependent asparagine formation. *Nature* **382**, 589–590.

Curnow, A.W., Hong, K.W., Yuan, R., Kim, S.I., Martins, O., Winkler, W., Henkin, T.M. & Söll, D. (1997) Glu-tRNA$^{Gln}$ amidotransferase: A novel heterotrimeric enzyme required for correct decoding of glutamine codons during translation. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 11819–11826.

Curnow, A.W., Tumbula, D.L., Pelaschier, J.T., Min, B. & Söll, D. (1998) Glutamyl-tRNA$^{Gln}$ amidotransferase in *Deinococcus radiodurans* may be confined to asparagine biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 12838–12843.

Delarue, M. (1995) Partition of aminoacyl-tRNA synthetases in two different structural classes dating back to early metabolism: Implications for the origin of the genetic code and the nature of protein sequences. *J. Mol. Evol.* **41**, 703–711.

Gesteland, R.F., Cech, T.R. & Atkins, J.F. (1999) *The RNA World*. Cold Spring Harbor Laboratory Press.

Hutchison, C.A., Peterson, S.N., Gill, S.R., Cline, R.T., White, O., Fraser, C.M., Smith, H.O. & Venter, J.C. (1999) Global transposon mutagenesis and a minimal *Mycoplasma* genome. *Science* **286**, 2165–2169.

Ibba, M., Bono, J.L., Rosa, P.A. & Söll, D. (1997a) Archaeal-type lysyl-tRNA synthetase in the Lyme disease spirochete *Borrelia burgdorferi*. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 14383–14388.

Ibba, M., Morgan, S., Curnow, A.W., Pridmore, D. R., Vothknecht, U. C., Gardner, W., Lin, W., Woese, C. R. & Söll, D. (1997b) A euryarchaeal lysyl-tRNA synthetase: Resemblance to class I synthetases. *Science* **278**, 1119–1122.

Ibba, M., Losey, H.C., Kawarabayasi, Y., Kikuchi, H., Bunjun, S. & Söll, D. (1999) Substrate recognition by class I lysyl-tRNA synthetases: A molecular basis for gene displacement. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 418–423.

Ibba, M. & Söll, D. (1999) Quality control mechanisms during translation. *Science* **286**, 1893–1897.

Ibba, M. & Söll, D. (2000) Aminoacyl-tRNA synthesis. *Annu. Rev. Biochem.* **69**, 617–650.

Illangasekare, M., Sanchez, G., Nickles, T. & Yarus, M. (1995) Aminoacyl-RNA synthesis catalyzed by an RNA. *Science* **267**, 643–647.

Lee, N., Bessho, Y., Wei, K., Szostak, J.W. & Suga, H. (2000) Ribozyme-catalyzed tRNA aminoacylation. *Nat. Struct. Biol.* **7**, 28–33.

Maizels, N. & Weiner, A.M. (1994) Phylogeny from function: Evidence from the molecular fossil record that tRNA originated in replication, not translation. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 6729–6734.

Nabholz, C.E., Hauser, R. & Schneider, A. (1997) *Leishmania tarentolae* contains distinct cytosolic and mitochondrial glutaminyl-tRNA synthetase activities. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 7903–7908.

Nagel, G.M. & Doolittle, R.F. (1995) Phylogenetic analysis of the aminoacyl-tRNA synthetases. *J. Mol. Evol.* **40**, 487–498.

RajBhandary, U.L. (1994) Initiator transfer RNAs. *J. Bacteriol.* **176**, 547–552.

Ribas de Pouplana, L., Turner, R.J., Steer, B.A. & Schimmel, P. (1998) Genetic code origins:

tRNAs older than their synthetases? *Proc. Natl. Acad. Sci. U.S.A.* **95**, 11295–11300.

Schimmel, P. & Kelley, S.O. (2000) Exiting an RNA world. *Nat. Struct. Biol*. **7**, 5–7.

Schön, A., Kannangara, C.G., Gough, S. & Söll, D. (1988) Protein biosynthesis in organelles requires misaminoacylation of tRNA. *Nature* **331**, 187–190.

Shiba, K. & Schimmel, P. (1992) Functional assembly of a randomly cleaved protein. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 1880–1884.

Smith, D.R., Doucette-Stamm, L.A., Deloughery, C., Lee, H., Dubois, J., Aldredge, T., Bashirzadeh, R., Blakely, D., Cook, R., Gilbert, K., Harrison, D., Hoang, L., Keagle, P., Lumm, W., Pothier, B., Qiu, D., Spadafora, R., Vicaire, R., Wang, Y., Wierzbowski, J., Gibson, R., Jiwani, N., Caruso, A., Bush, D. & Reeve, J.N. (1997) Complete genome sequence of *Methanobacterium thermoautotrophicum* ΔH: Functional analysis and comparative genomics. *J. Bacteriol.* **179**, 7135–7155.

Söll, D. & Schimmel, P.R. (1974) Aminoacyl-tRNA synthetases. *The Enzymes* **10**, 489–538.

Stathopoulos, C., Li, T., Longman, R., Vothknecht, U.C., Becker, H.D., Ibba, M. & Söll, D. (2000) One polypeptide with two aminoacyl-tRNA synthetase activities. *Science* **287**, 479–482.

Stathopoulos, C., Jacquin-Becker, C., Becker, H.D., Li, T., Ambrogelly, A. Longman, R. & Söll, D. (2001) *Methanococcus jannaschii* prolyl-cysteinyl-tRNA synthetase possesses overlapping amino acid binding sites. *Biochemistry* **40**, 46–52.

Tumbula, D.L., Vothknecht, U.C., Kim, H.-S., Ibba, M., Min, B., Li, T., Pelaschier, J., Stathopoulos, C., Becker, H.D. & Söll, D. (1999) Archaeal aminoacyl-tRNA synthesis: Diversity replaces dogma. *Genetics* **152**, 1269–1276.

Tumbula, D.L., Becker, H.D., Chang, W.-Z. & Söll, D. (2000) Domain-specific recruitment of amide amino acids for protein synthesis. *Nature* **407**, 106–110.

White, B.N. & Bayley, S.T. (1972) Further codon assignments in an extremely halophilic bacterium using a cell-free protein-synthesizing system and a ribosomal binding assay. *Can. J. Biochem.* **50**, 600–609.

Wilcox, M. & Nirenberg, M. (1968) Transfer RNA as a cofactor coupling amino acid synthesis with that of protein. *Proc. Natl. Acad. Sci. U.S.A.* **61**, 229–236.

Woese, C.R., Olsen, G., Ibba, M. & Söll, D. (2000) Aminoacyl-tRNA synthetases, the genetic code, and the evolutionary process. *Microbiol. Mol. Biol. Rev.* **64**, 202–236.