Chapman University

# Chapman University Digital Commons

Biology, Chemistry, and Environmental Sciences Faculty Articles and Research

Science and Technology Faculty Articles and Research

2004

# Aminoacyl-tRNAs: Setting the Limits of the Genetic Code

Michael Ibba
*Chapman University*, ibba@chapman.edu

Dieter Söll
*Yale University*

Follow this and additional works at: https://digitalcommons.chapman.edu/sees_articles

Part of the Amino Acids, Peptides, and Proteins Commons, Biochemistry Commons, Cellular and Molecular Physiology Commons, Molecular Biology Commons, Nucleic Acids, Nucleotides, and Nucleosides Commons, and the Other Biochemistry, Biophysics, and Structural Biology Commons

# Aminoacyl-tRNAs: Setting the Limits of the Genetic Code

REVIEW

# Aminoacyl-tRNAs: setting the limits of the genetic code

## Michael Ibba[1,3] and Dieter Söll[2,4]

[1]Department of Microbiology, The Ohio State University, Columbus, Ohio 43210-1292, USA; [2]Departments of Molecular Biophysics and Biochemistry, and Chemistry, Yale University, New Haven, Connecticut 06520-8114, USA

Aminoacyl-tRNAs (aa-tRNAs) are simple molecules with a single purpose—to serve as substrates for translation. They consist of mature tRNAs to which an amino acid has been esterified at the 3′-end. The 20 different types of aa-tRNA are made by the 20 different aminoacyl-tRNA synthetases (aaRSs, of which there are two classes), one for each amino acid of the genetic code (Ibba and Söll 2000). This would be fine if it were not for the fact that such a straightforward textbook scenario is not true in a single known living organism. aa-tRNAs lie at the heart of gene expression; they interpret the genetic code by providing the interface between nucleic acid triplets in mRNA and the corresponding amino acids in proteins. The synthesis of aa-tRNAs impacts the accuracy of translation, the expansion of the genetic code, and even provides tangible links to primary metabolism. These central roles vest immense power in aa-tRNAs, and recent studies show just how complex and diverse their synthesis is.

## How many aa-tRNAs are there?

The aa-tRNAs typically found in the cell can be divided into three main types (Table 1). The largest group comprises correct pairings of an amino acid and the corresponding tRNA, and they serve as substrates for ribosomal translation. Among these, the canonical elongator tRNAs are by far the most common. They are usually made directly by the corresponding aaRS (Woese et al. 2000; O'Donoghue and Luthey-Schulten 2003), and then screened for correctness by elongation factor Tu (EF-Tu), which also delivers them to the ribosome (LaRiviere et al. 2001). In much the same way, initiator aa-tRNAs in archaea and eukaryotes are made directly by methionyl-tRNA synthetase (MetRS), and then bind to initiation factors that deliver them to the ribosome. In bacteria and organelles the situation is more complex, as the Met-tRNA$^{fMet}$ produced by MetRS must first be formylated by a specific formyltransferase before it can be used to initiate protein synthesis. The remainder of the translation substrate aa-tRNAs comprises species used for non-canonical elongation, such as, for example, the seleno-cysteinyl-tRNA$^{Sec}$ that decodes UGA stop codons as selenocysteine (Sec). The second major type of aa-tRNAs found in the cell comprises misacylated translation substrates, tRNAs where the amino acid esterified to the 3′-end does not correlate with the anticodon sequence as defined by the genetic code. Misacylated aa-tRNAs fall into two groups, those deliberately made by the cell as precursors for pretranslational modification, and those resulting from errors made by aaRSs. Rapid turnover by aaRSs and kinetic proofreading by EF-Tu ensure that the generation of misacylated translation substrates does not compromise the fidelity of cellular protein synthesis. The remaining type of aa-tRNAs found in the cell consists of correctly charged species used for cellular processes other than translation. Examples are the porphyrin 'precursor' Glu-tRNA$^{Glu}$ (Schauer et al. 2002), the cell wall peptidoglycan precursors Gly-tRNA$^{Gly}$ (Bumsted et al. 1968) and Ser-tRNA$^{Ser}$ (Petit et al. 1968), and the lysyl-phosphatidylglycerol precursor Lys-tRNA$^{Lys}$ (Lennarz et al. 1966; Oku et al. 2004). When these numerous aa-tRNAs are all taken into account in *Escherichia coli*, where the picture is perhaps most complete, it can be seen that a normally growing cell might contain 25 or more different types of aa-tRNAs. Here we will review recent advances in our understanding of aa-tRNA synthesis and function, with particular emphasis on those types outside the standard set of 20 isoacceptor families.

## 20, 21, 22 . . .

In bacteria, mitochondria, and chloroplasts, protein synthesis is initiated by formylmethionyl-tRNA$^{fMet}$, an aa-tRNA outside the standard set of 20. Even though both eukaryotic (tRNA$_i^{Met}$) and bacterial/organellar (tRNA$^{fMet}$) initiator tRNAs already contain common features that ensure that neither act in elongation (Stortchevoi et al. 2003), formylation is important for optimal growth in systems using tRNA$^{fMet}$ (Vial et al. 2003). The requirement for formylation to specifically mark initiator aa-tRNAs in these systems is perhaps best illustrated in the mitochondria of *Trypanosoma brucei*. This trypanosome does not synthesize tRNAs in its mitochondria, which instead import cytosolic nuclear-encoded tRNAs. These include the eukaryotic elongator tRNA$^{Met}$ (but not tRNA$_i^{Met}$), all of which is then charged with methio-

**Table 1.** *Aminoacyl-tRNAs found in bacterial cells*

| Aminoacyl-tRNA | Interacting partners | Fate of aminoacyl moiety | Number of aa-tRNA sets or tRNA species |
|---|---|---|---|
| Translation substrates | | | |
| Canonical elongator aa-tRNAs | EF-Tu/EF-1a | Protein synthesis | 20 |
| Initiator aa-tRNAs | Initiation factors | Protein synthesis | fMet-tRNA$^{fMet}$ |
| Noncanonical elongator aa-tRNAs | Other elongation factors (e.g., SelB) | Protein synthesis | Sec-tRNA$^{Sec}$ |
| Misacylated translation substrates | | | |
| Deliberately non-cognate aa-tRNA | Pretranslational modification pathway | Modification then protein synthesis | 4 (Asp-tRNA$^{Asn}$, Glu-tRNA$^{Gln}$, Met-tRNA$^{fMet}$, Ser-tRNA$^{Sec}$) |
| Mistakenly noncognate aa-tRNA | Aminoacyl-tRNA synthetases D-amino acid deacylases | Free amino acid | Varies D-Tyr-tRNA$^{Tyr}$ and others |
| Nontranslation substrates | | | |
| Canonical elongator aa-tRNAs | Porphyrin biosynthesis | Nonprotein macromolecules | Glu-tRNA$^{Glu}$ |
| | Peptidoglycan biosynthesis | | Gly-tRNA$^{Gly}$, Ser-tRNA$^{Ser}$ |
| | Lysylphosphatidylglycerol synthesis | | Lys-tRNA$^{Lys}$ |

nine. A fraction of this Met-tRNA$^{Met}$ is then formylated to yield what effectively becomes formylmethionyl-tRNA$^{fMet}$, allowing initiation of protein synthesis (Tan et al. 2002). Even though formylmethionine is an amino acid outside the canonical 20 inserted during protein synthesis, it has not historically been regarded as the 21st amino acid. This probably relates to the fact that the N-formyl group is removed from mature proteins by peptide deformylase (Giglione and Meinnel 2001), leaving the canonical methionine as the mature N terminus. In contrast, the reactive (Hatfield and Gladyshev 2002) noncanonical amino acid selenocysteine retains its native structure in mature proteins and is generally regarded as the 21st genetically encoded amino acid.

The pathway by which selenocysteine is inserted during protein synthesis in response to a UGA codon in a specific mRNA sequence context has been most completely described to date in *E. coli* (for review, see Böck et al. 2004). Briefly, a UGA-recognizing tRNA species (tRNA$^{Sec}$) is misacylated by seryl-tRNA synthetase to Ser-tRNA$^{Sec}$, and then converted to Sec-tRNA$^{Sec}$ by selenocysteine synthase (SelA). Sec-tRNA$^{Sec}$ subsequently forms a complex with the selenocysteine-specific elongation factor SelB and GTP, and this in turn binds a specific structure (the selenocysteine insertion element, or SECIS) in ribosome-bound mRNA. The resulting quaternary complex then directs the cotranslational insertion of selenocysteine at particular in-frame UGA codons. The synthesis of selenocysteine in *E. coli* is carefully balanced with availability of the trace element selenium via a SECIS-like element in the 5′ nontranslated region of the *selAB* operon that encodes SelA and SelB. When excess quaternary complexes are available in the cell, they bind the SECIS-like element and repress synthesis of SelA and SelB (Thanbichler and Böck 2002). Comparable pathways for selenocysteine insertion have been found in archaea (Rother et al. 2003), animals (e.g., the human genome encodes 25 selenoproteins; Kryukov et al. 2003), and most recently plants (Novoselov et al.

2002; Rao et al. 2003). The exact mechanism of insertion in archaea and eukaryotes differs from bacteria in two main aspects. In bacteria the SECIS is normally located adjacent to the UGA, but is in the 3′ untranslated region in archaea and eukaryotes. As a consequence of these differing locations, the SelB protein also differs in its function. In bacteria SelB binds both Sec-tRNA$^{Sec}$ and the SECIS element to facilitate insertion, whereas in other cases Sec-tRNA$^{Sec}$ binds SelB, whose association with the SECIS is mediated by SECIS-binding proteins (Driscoll and Copeland 2003; Zavacki et al. 2003). Continuing progress in understanding the mechanism of selenocysteine insertion has now allowed for the first time a fairly complete description of the occurrences of selenoproteins in the living kingdom (Kryukov et al. 2003; Castellano et al. 2004), which may eventually explain why, even though selenoproteins are widespread, they are by no means ubiquitous. One factor in the rather idiosyncratic distribution of selenoproteins may be their broader substrate specificity and their efficiency in a wider range of microenvironmental conditions compared to equivalent proteins that instead contain cysteine in their active sites (Gromer et al. 2003).

Current knowledge of pyrrolysine (Pyl), the 22nd cotranslationally inserted amino acid encoded by UAG, suggests that it is far less broadly distributed than selenocysteine, being restricted to a handful of archaea and bacteria. Genome analyses predict that in-frame UAG codons are fairly widespread in the *Methanosarcineae* and may also be present in at least one bacterium, *Desulfitobacterium hafniense* (Srinivasan et al. 2002). The first hints of the existence of a 22nd genetically encoded amino acid came from studies of monomethylamine methyltransferase in the archaeon *Methanosarcina barkeri*. It was found that an in-frame UAG stop codon in this gene was translated in vivo as lysine or a lysine derivative (James et al. 2001). Subsequent crystallographic data showed that this residue was in fact the novel amino acid pyrrolysine (Hao et al. 2002). The dis-

covery in *M. barkeri* of a specialized suppressor tRNA containing the required CUA anticodon (tRNA$^{Pyl}$, encoded by the *pylT* gene) suggested that pyrrolysine might be cotranslationally inserted just as selenocysteine is (Srinivasan et al. 2002); yet posttranslational modification of lysine also remains a possibility.

## Pyrrolysine insertion

Cotranslational insertion of pyrrolysine would be initiated by the attachment of lysine to tRNA$^{Pyl}$ to generate Lys-tRNA$^{Pyl}$ (Fig. 1). The exact mechanism of Lys-tRNA$^{Pyl}$ synthesis has been a matter of recent debate; whereas initial work suggested that the aaRS-like protein PylS was responsible (Srinivasan et al. 2002), more recent studies point to tRNA$^{Pyl}$ being charged by the class II lysyl-tRNA synthetase (Polycarpo et al. 2003; Theobald-Dietrich et al. 2004), with the class I lysyl-tRNA synthetase also playing an as yet undefined, and perhaps nonessential, role in the process. By analogy to selenocysteinyl-tRNA synthesis, the subsequent step in pyrrolysine insertion is expected to be the conversion of Lys-tRNA$^{Pyl}$ into pyrrolysyl-tRNA$^{Pyl}$. Although there is to date no direct evidence of this conversion, a number of proteins encoded by genes in the *pyl* operon, including PylS, have been proposed as candidates to catalyze formation of the amino acid's pyrrole group (Srinivasan et al. 2002; Polycarpo et al. 2003). The lack of experimental evidence to date also allows the possibility that direct

acylation of pyrrolysine, and not lysine, onto tRNA$^{Pyl}$ is accomplished (see Fig. 1). This would be in contrast to the pretranslational modification of other amino acids charged to tRNA (Asp-tRNA$^{Asn}$ and Glu-tRNA$^{Gln}$ [see below], and Met-tRNA$^{fMet}$ and Ser-tRNA$^{Sec}$ [see above]). This alternative scheme demands that pyrrolysine be a normal stable metabolite in *Methanosarcina* with high specificity during aminoacylation, given that PylS and the canonical lysyl-tRNA synthetases all attach lysine very efficiently to tRNA$^{Pyl}$ (Srinivasan et al. 2002; Polycarpo et al. 2003).

From comparisons to selenoprotein synthesis, the final steps in pyrrolysine insertion are expected to be the binding of pyrrolysyl-tRNA$^{Pyl}$ to a specialized translation elongation factor, delivery to the ribosome, and insertion of pyrrolysine into the nascent polypeptide in response to an in-frame UAG codon. Current knowledge of the mechanism of selenocysteine insertion, indirect experimental evidence, and comparative genomic analyses provide some insight into how these final steps might be accomplished (Fig. 1). The requirements for elongation factor binding during cotranslational insertion are complex, as is evident from extensive work with selenocysteine. Foremost among the prerequisites is the need for the "precursor" Lys-tRNA$^{Pyl}$ not to bind elongation factor, as this could lead to the inadvertent translation of some UAG codons as lysine rather than pyrrolysine. Interestingly, an anticodon variant of Lys-tRNA$^{Pyl}$ is in fact a substrate for the bacterial elongation
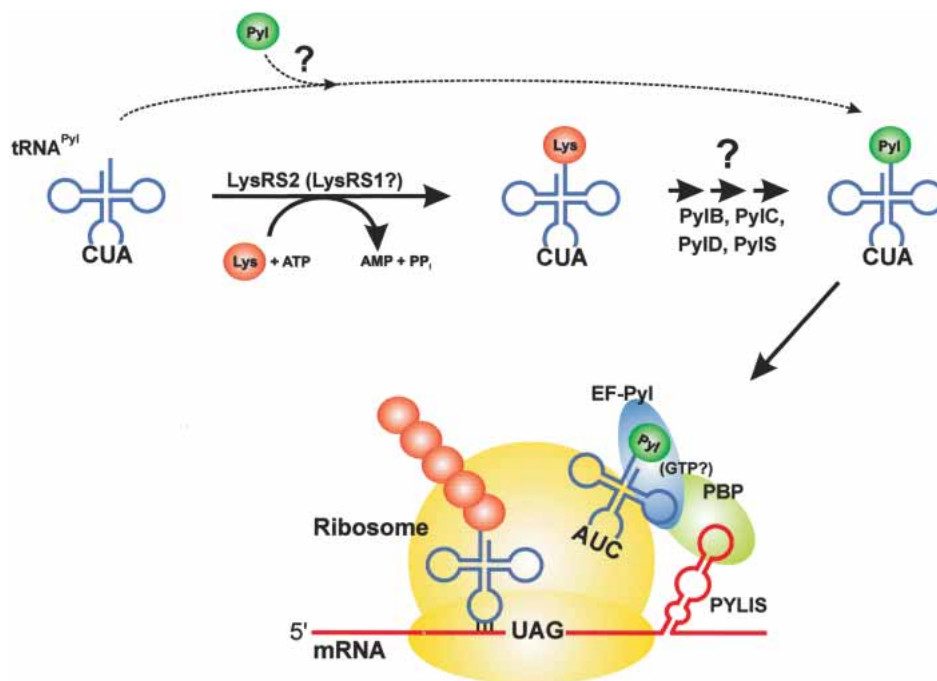


**Figure 1.** Hypothetical scheme for the cotranslational insertion of pyrrolysine in response to a context-dependent UAG codon. tRNA$^{Pyl}$ (anticodon CUA) is either first aminoacylated with lysine, and subsequently modified to generate Pyl-tRNA$^{Pyl}$, or alternatively Pyl-tRNA$^{Pyl}$ is directly synthesized by an unknown route. Pyl-tRNA$^{Pyl}$ would then bind the SelB-like elongation factor EF-Pyl (see text for details of candidate EF-Pyl proteins). The Pyl-tRNA$^{Pyl}$:EF-Pyl complex would associate with the pyrrolysine insertion sequence (PYLIS) via a putative PYLIS binding protein (PBP), thereby directing delivery of Pyl-tRNA$^{Pyl}$ to the ribosomal A site when the decoding site is occupied by the corresponding UAG.

factor *Thermus thermophilus* EF-Tu, raising the possibility that the mechanism of pyrrolysine insertion is divergent to that employed for selenocysteine (Theobald-Dietrich et al. 2004). Although this finding questions the assertion that essentially identical routes insert selenocysteine and pyrrolysine, genome sequence analysis nevertheless indicates that two other idiosyncratic elements of selenoprotein synthesis are also employed here. The first are conserved stem-loop structures similar to known SECIS elements (termed PYLIS) immediately 3' of the in-frame UAG codons in the *mtmB1* and *mtmB2* genes of *M. barkeri*, *Methanosarcina acetivorans*, and *Methanosarcina mazei* (Namy et al. 2004), and the *mttB* gene in *D. hafniense* (Fig. 2). The second key feature is the presence of genes encoding truncated SelB-like proteins in all three known *Methanosarcineae* genomes (MA4654, Galagan et al. 2002; MM1309, Deppenmeier et al. 2002; GenBank accession no. ZP_00078665) and in *Methanococcoides burtonii*, even though none of these genomes apparently encode selenocysteine. The methanosarcinaeal *selB* paralogs are smaller than the bacterial SelB proteins, so the presence of an additional interacting protein is probably required to enable SelB to perform both PYLIS recognition and elongation factor function, as also proposed for the shorter methanococcal SelB proteins (Rother et al. 2001). Although the precise details may differ, these observations suggest that much of the general selenoprotein synthesis strategy has effectively been usurped in the *Methanosarcineae*, and perhaps also in *D. hafniense*, for pyrrolysine insertion (Fig. 1). In this context, the comparative rarity of pyrrolysine cotranslation may be explained by its having evolved from an older, more common, selenocysteine pathway. If this
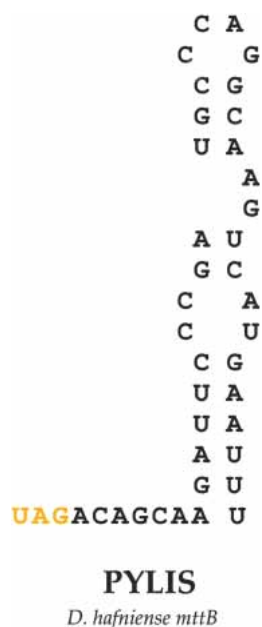
were indeed the case, it would also suggest that other noncanonical amino acids could be added to the genetic code in a similar fashion by dual use of existing codons.

## Translating glutamine and asparagine

The cotranslational insertions of formylmethionine, selenocysteine, and pyrrolysine require the indirect synthesis of noncanonical aa-tRNAs that are subsequently accepted by the ribosome. In a very large number of bacterial and archaeal organisms, the cotranslational insertion of the canonical amino acids glutamine and asparagine is also preceded by pretranslational amino acid modification. The noncognate aa-tRNA species, Asp-tRNA$^{Asn}$ and Glu-tRNA$^{Gln}$, are formed by a nondiscriminating-type aaRS (see below) as precursors for subsequent conversion to Asn-tRNA$^{Asn}$ and Gln-tRNA$^{Gln}$ by Asp/Glu-tRNA amidotransferase (Ibba et al. 2000). The noncognate species do not present a threat to translational fidelity, as they are effectively discriminated against by EF-Tu (Stanzel et al. 1994; Becker and Kern 1998; Asahara and Uhlenbeck 2002). Genome sequence and biochemical analyses provided a major surprise in revealing the lack of conservation in Gln-tRNA formation. This is the only known example where nature employs a different biosynthetic route in each of the three kingdoms that comprise the living world. The eukaryotic cytoplasm utilizes direct acylation of tRNA$^{Gln}$ by glutaminyl-tRNA synthetase (GlnRS), whereas most prokaryotes use the transamidation route. However, bacteria and archaea differ in their choice of enzyme; bacteria use a heterotrimeric Asp/Glu-tRNA amidotransferase, and the archaea employ a related heterodimeric enzyme (Tumbula et al. 2000). The existence of two alternate and redundant routes of Asn-tRNA and Gln-tRNA synthesis effectively allowed organisms the choice of pathway. Genome analysis and biochemical data reveal in bacteria four different scenarios, all of which are based on the presence of different operational pathways for the formation of Gln-tRNA and/or Asn-tRNA (Fig. 3). (1) Organisms with AsnRS and GlnRS: Some γ-proteobacteria (e.g., *E. coli*) acquired GlnRS from eukaryotes by horizontal gene transfer (Lamour et al. 1994) and therefore produce Asn-tRNA and Gln-tRNA by direct acylation. (2) Organisms lacking AsnRS and GlnRS using Asp/Glu-AdT: A larger set of bacteria (e.g., *Chlamydia trachomatis*) lacks these two synthetases. Instead, the bacterial Asp/Glu-tRNA amidotransferase carries out both functions (Raczniak et al. 2001). Obviously, this requires that both glutamyl-tRNA synthetase (GluRS) and aspartyl-tRNA synthetase (AspRS) be nondiscriminating (ND), that is, that they be able to respectively recognize the noncognate tRNA$^{Gln}$ and tRNA$^{Asn}$ just as effectively as the cognate tRNA$^{Glu}$ and tRNA$^{Asp}$. (3) Bacteria with GlnRS but lacking AsnRS: Some organisms (e.g., *Pseudomonas aeroginosa*; Akochy et al. 2004) lack an ND GluRS and employ GlnRS for Gln-tRNA formation while Asp/Glu-AdT supplies Asn-tRNA. (4) Bacteria with AsnRS but lacking GlnRS: Here is the opposite situation in which AsnRS forms Asn-tRNA and Glu/



```
          C   A
        C       G
          C   G
          G   C
          U   A
                A
                G
        A   U
        G   C
      C       A
      C       U
          C   G
          U   A
          U   A
        A   U
        G   U
UAGACAGCAA      U
```

**PYLIS**

*D. hafniense mttB*

**Figure 2.** The putative PYLIS element of the *D. hafniense mttB* gene (GenBank accession no. NZ_AAAW02000173). RNAstructure 3.71 was used for folding (http://rna.chem.rochester.edu/RNAstructure.html).
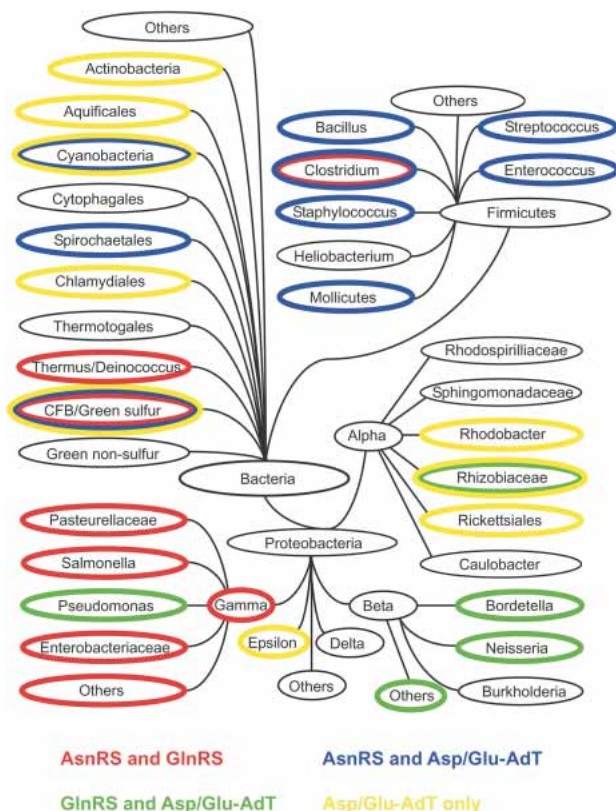
**Figure 3.** Distribution of direct and indirect pathways for Asn-tRNA$^{Asn}$ and Gln-tRNA$^{Gln}$ synthesis in bacterial lineages. The occurrence of genes encoding AsnRS, GlnRS, and Asp/Glu-Adt in genome sequences from representative bacterial groups is indicated. The organisms contain the following: (red) AsnRS and GlnRS; (green) GlnRS and Asp/Glu-AdT; (blue) AsnRS and Asp/Glu-AdT; (yellow) Asp/Glu-AdT (for further explanation see text). The bacterial phylogeny is based upon 16S rRNA, adapted from http://www.ncbi.nlm.nih.gov/sutils/genom_tree.cgi.

Asp-AdT provides Gln-tRNA (e.g., *Bacillus subtilis*). Although the reasons for this unexpected diversity remain unclear, clues to some of the underlying determinants can be gleaned from a number of recent studies.

The presence of the notionally redundant transamidation and direct acylation pathways in their active forms has also been observed, for example in *Deinococcus radiodurans* and *Thermus thermophilus* (Curnow et al. 1998; Becker and Kern 1998). These organisms contain all possible enzymes for Asn-tRNA and Gln-tRNA formation, AsnRS, GlnRS and Asp/Glu-AdT. Biochemical and genetic studies demonstrated that the tRNA-dependent Asp → Asn conversion is the only route to asparagine biosynthesis available in these organisms. Genome sequence analysis revealed that unless another still unknown enzyme for Asp → Asn conversion exists, Asp/Glu-AdT is essential for asparagine synthesis in many organisms (Min et al. 2002). These data lend support to the notion that the transamidation pathways are intimately linked to amino acid metabolism, as has also been proposed for other aspects of translation (Francklyn

2003; see also below). This interesting relationship between an enzyme involved in protein synthesis and in amino acid biosynthesis was further highlighted by the discovery that *Pyrococcus abyssi* contains a truncated AsnRS gene which is an asparagine synthetase functional in converting Asp → Asn (Roy et al. 2003). The close structural relationship between this enzyme, AspRS, and the tRNA-independent AsnA-type asparagine synthetase suggests that they may have diverged from the archaeal genre AspRS, as all of these enzymes use ATP and give rise to an aspartyl adenylate intermediate (Gatti and Tzagoloff 1991; Roy et al. 2003). As alluded to above, the transamidation routes leading to amide aa-tRNA synthesis are dependent on the inherently inaccurate ND versions of GluRS and AspRS. Although it is moot to speculate whether the discriminating (D) or the ND form is the evolutionarily more ancient, some recent studies have shown that the change from D-AspRS to ND-AspRS rests on very few amino acid changes (Charron et al. 2003) or even just a single amino acid replacement (Feng et al. 2003). Thus the distributions of the Asp/Asn translation and amino acid biosynthesis machineries may be more a reflection of the physiological demands of contemporary organisms rather than a manifestation of their deeper evolutionary origin. In contrast, GlnRS, generally held to be the most recently evolved aaRS, seems to have no obvious residual relationship to amino acid metabolism. Instead the simultaneous existence of both D-GluRS and ND-GluRS, as still observed in a few bacteria, would seem to facilitate the replacement of the indirect route of Gln-tRNA$^{Gln}$ synthesis by GlnRS (Salazar et al. 2003; Skouloubris et al. 2003), either by bacterial GluRS evolution or horizontal gene transfer from eukarya (Lamour et al. 1994).

## Quality control and aa-tRNA

The translation of a number of codons, as described above, relies on the synthesis of certain noncognate aa-tRNAs (deliberately noncognate species; Table 1). More often, however, the unintended synthesis of noncognate aa-tRNAs presents a threat to the accuracy of translation. Consequently a variety of quality control strategies are employed by the cell to ensure that typically only about one in every $10^4$ codons is mistranslated, even though significantly higher rates at particular codons can be tolerated (Ruusala et al. 1982; Kurland 1992; Min et al. 2003; Núñez et al. 2004). Aside from codon-anticodon pairing (Ogle et al. 2003), the mechanisms of translational quality control are broadly of three types; specificity of substrate selection by aa-tRNA synthetases, proofreading, and exclusion from the ribosome. Exhaustive studies over the last four decades have described in intricate detail the molecular recognition strategies that allow synthetases to select from the cellular pool particular canonical amino acids and tRNAs to generate correctly matched aa-tRNAs (for review, see Ibba and Söll 2000). This includes a variety of different editing mechanisms designed to ensure that misactivated amino acids or mischarged aa-tRNA will be hydrolyzed (Nureki

et al. 1998; Dock-Bregeon et al. 2000; Ferri-Fioni et al. 2001; Lincecum et al. 2003). In particular, some aaRSs contain regions of the protein (editing domains) with the ability to carry out this specific hydrolytic reaction (Lin et al. 1996). More recently it has become clear that amino acid specificity is further enhanced by the existence of a natural reservoir of diverse synthetase alleles. This pool of synthetases, with slight differences in specificity towards molecules outside the canonical set of amino acids, serves to both exclude amino acid analogs from translation (Jester et al. 2003) and provide resistance against inhibitory amino acid mimics (Brown et al. 2003).

## Synthetase-like proteins

Comparative analyses of the highly conserved aaRS genes revealed that numerous bacterial genomes contain open reading frames that encode synthetase-like proteins, many of which are smaller than the corresponding functional synthetases and are not active in aa-tRNA formation (for review, see Schimmel and Ribas De Pouplana 2000). The motivation to determine a function for these genes was greatly enhanced by the discovery that a histidyl-tRNA synthetase-like protein, the paralog hisZ, had a required role in histidine biosynthesis in some organisms (Sissler et al. 1999). A different role was assigned to the paralogs of alanyl-tRNA synthetase (AlaRS), prolyl-tRNA synthetase (ProRS), and threonyl-tRNA synthetase (ThrRS). Although the editing domains of some synthetases (e.g., AlaRS) are always located within the multidomain protein and perform the editing reaction in *cis* (Beebe et al. 2003), even the isolated domain displays reduced in vitro activity in hydrolyzing misacylated tRNA (Lin et al. 1996). The absence of an editing domain in many archaeal ThrRSs and in some bacterial ProRSs suggested that their paralogs carry out *trans*-editing of mischarged tRNAs. Examination of the naturally occurring paralogs AlaX, ProX, ThrX (Ahel et al. 2003), and the ProRS paralog Ybak of *Haemophilus influenzae* (Wong et al. 2003) showed that this is the case, as these proteins efficiently and specifically hydrolyze misacylated tRNA substrates. These data suggest that the editing domains in aaRSs may have originated from similar autonomous editing modules. Interestingly, a widely occurring GluRS paralog (*yadB* in *E. coli*) that lacks the enzyme's anticodon binding domain is not involved in tRNA editing, but rather attaches glutamate specifically to tRNA$^{Asp}$ (Dubois et. al. 2004; Salazar et al. 2004). However, glutamate is not esterified to the normal aminoacylation position at the tRNA's 3′-terminal adenosine, but is attached to the hypermodified nucleoside queuosine located in the first position of the tRNA$^{Asp}$ anticodon (Salazar et al. 2004). Thus, this aa-tRNA synthetase fragment contributes to nucleotide modification of tRNA, further broadening the range of functions thus far attributed to synthetase paralogs.

## Beyond aa-tRNA synthesis

The routine detection of errors in translation indicates that, despite the best efforts of synthetases, some incor-

rectly charged tRNAs are able to clear the first hurdle of quality control. The next step in protein synthesis requires the association of aa-tRNAs with translation factors, and this is used to provide an additional checkpoint. Although EF-Tu discrimination is not absolute (Min et al. 2003; Núñez et al. 2004), it can distinguish both precursor noncognate tRNAs (see above) and other incorrectly charged tRNA species (LaRiviere et al. 2001; Asahara and Uhlenbeck 2002), whereas initiation factor 2 specifically binds initiator aa-tRNAs (Mayer et al. 2003). Finally, it has been suggested that the ribosome itself might also provide some degree of selectivity towards cognate aa-tRNAs (Wolfson et al. 2001). Perhaps not surprisingly, these various quality control mechanisms are directed at natural noncognate aa-tRNAs and are relatively insensitive to the introduction of synthetic amino acids. This has allowed the design of both bacterial and eukaryotic in vitro systems for the site-specific cotranslational insertion of synthetic amino acids in response to in-frame stop codons (Wang et al. 2001; Chin et al. 2003; Köhrer et al. 2003; Mehl et al. 2003). The common feature of all these systems is the existence of aaRSs with appropriately re-engineered substrate specificities (Kobayashi et al. 2003). No further changes to the cellular quality control machinery are necessary to accommodate these numerous expansions of the genetic code, providing a graphic illustration of how aa-tRNA synthesis is the single most important factor in setting the limits within which the genetic code is interpreted.

## Acknowledgments

## References

Ahel, I., Korencic, D., Ibba, M., and Söll, D. 2003. Trans-editing of mischarged tRNAs. *Proc. Natl. Acad. Sci.* **100:** 15422–15427.

Akochy, P.M., Bernard, D., Roy, P.H., and Lapointe, J. 2004. Direct glutaminyl-tRNA biosynthesis and indirect asparaginyl-tRNA biosynthesis in *Pseudomonas aeruginosa* PAO1. *J. Bacteriol.* **186:** 767–776.

Asahara, H. and Uhlenbeck, O.C. 2002. The tRNA specificity of *Thermus thermophilus* EF-Tu. *Proc. Natl. Acad. Sci.* **99:** 3499–3504.

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

Beebe, K., Ribas De Pouplana, L., and Schimmel, P. 2003. Elucidation of tRNA-dependent editing by a class II tRNA synthetase and significance for cell viability. *EMBO J.* **22:** 668–675.

Böck, A., Thanbichler, M., Rother, M., and Resch, A. 2004. Selenocysteine. In *Aminoacyl-tRNA synthetases* (eds. M. Ibba, C.S. Francklyn, and S. Cusack). Landes Bioscience (in press).

Brown, J.R., Gentry, D., Becker, J.A., Ingraham, K., Holmes, D.J., and Stanhope, M.J. 2003. Horizontal transfer of drug-resis-

tant aminoacyl-transfer-RNA synthetases of anthrax and Gram-positive pathogens. *EMBO Rep.* **4:** 1–7.

Bumsted, R.M., Dahl, J.L., Söll, D., and Strominger, J.L. 1968. Biosynthesis of the peptidoglycan of bacterial cell walls. X. Further study of the glycyl transfer ribonucleic acids active in peptidoglycan synthesis in *Staphylococcus aureus*. *J. Biol. Chem.* **243:** 779–782.

Castellano, S., Novoselov, S.V., Kryukov, G.V., Lescure, A., Blanco, E., Krol, A., Gladyshev, V.N., and Guigo, R. 2004. Reconsidering the evolution of eukaryotic selenoproteins: A novel nonmammalian family with scattered phylogenetic distribution. *EMBO Rep.* **5:** 71–77.

Charron, C., Roy, H., Blaise, M., Giegé, R., and Kern, D. 2003. Nondiscriminating and discriminating aspartyl-tRNA synthetases differ in the anticodon-binding domain. *EMBO J.* **22:** 1632–1643.

Chin, J.W., Cropp, T.A., Anderson, J.C., Mukherji, M., Zhang, Z., and Schultz, P.G. 2003. An expanded eukaryotic genetic code. *Science* **301:** 964–967.

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

Deppenmeier, U., Johann, A., Hartsch, T., Merkl, R., Schmitz, R.A., Martinez-Arias, R., Henne, A., Wiezer, A., Bäumer, S., Jacobi, C., et al. 2002. The genome of *Methanosarcina mazei:* Evidence for lateral gene transfer between Bacteria and Archaea. *J. Mol. Microbiol. Biotechnol.* **4:** 453–461.

Dock-Bregeon, A., Sankaranarayanan, R., Romby, P., Caillet, J., Springer, M., Rees, B., Francklyn, C.S., Ehresmann, C., and Moras, D. 2000. Transfer RNA-mediated editing in threonyl-tRNA synthetase. The class II solution to the double discrimination problem. *Cell* **103:** 877–884.

Driscoll, D.M. and Copeland, P.R. 2003. Mechanism and regulation of selenoprotein synthesis. *Annu. Rev. Nutr.* **23:** 17–40.

Dubois, D.Y., Blaise, M., Becker, H.D., Campanacci, V., Keith, G., Giegé, R., Cambillau, C., Lapointe, J., and Kern, D. 2004. An aminoacyl-tRNA synthetase-like protein encoded by the *Escherichia coli yadB* gene glutamylates specifically tRNA$^{Asp}$. *Proc. Natl. Acad. Sci.* (in press).

Feng, L., Tumbula-Hansen, D., Toogood, H., and Söll, D. 2003. Expanding tRNA recognition of a tRNA synthetase by a single amino acid change. *Proc. Natl. Acad. Sci.* **100:** 5676–5681.

Ferri-Fioni, M.L., Schmitt, E., Soutourina, J., Plateau, P., Mechulam, Y., and Blanquet, S. 2001. Structure of crystalline D-Tyr-tRNA$^{Tyr}$ deacylase: A representative of a new class of tRNA-dependent hydrolase. *J. Biol. Chem.* **276:** 47285–47290.

Francklyn, C. 2003. tRNA synthetase paralogs: Evolutionary links in the transition from tRNA-dependent amino acid biosynthesis to de novo biosynthesis. *Proc. Natl. Acad. Sci.* **100:** 9650–9652.

Galagan, J.E., Nusbaum, C., Roy, A., Endrizzi, M.G., Macdonald, P., FitzHugh, W., Calvo, S., Engels, R., Smirnov, S., Atnoor, D., et al. 2002. The genome of *M. acetivorans* reveals extensive metabolic and physiological diversity. *Genome Res.* **12:** 532–542.

Gatti, D.L. and Tzagoloff, A. 1991. Structure and evolution of a group of related aminoacyl-tRNA synthetases. *J. Mol. Biol.* **218:** 557–568.

Giglione, C. and Meinnel, T. 2001. Organellar peptide deformylases: Universality of the N-terminal methionine cleavage mechanism. *Trends Plant Sci.* **6:** 566–572.

Gromer, S., Johansson, L., Bauer, H., Arscott, L.D., Rauch, S., Ballou, D.P., Williams Jr., C.H., Schirmer, R.H., and Arner, E.S. 2003. Active sites of thioredoxin reductases: Why sele-

noproteins? *Proc. Natl. Acad. Sci.* **100:** 12618–12623.

Hao, B., Gong, W., Ferguson, T.K., James, C.M., Krzycki, J.A., and Chan, M.K. 2002. A novel UAG encoded residue in the structure of a methanogen methyltransferase. *Science* **296:** 1462–1466.

Hatfield, D.L. and Gladyshev, V.N. 2002. How selenium has altered our understanding of the genetic code. *Mol. Cell Biol.* **22:** 3565–3576.

Ibba, M. and Söll, D. 2000. Aminoacyl-tRNA synthesis. *Annu. Rev. Biochem.* **69:** 617–650.

Ibba, M., Becker, H.D., Stathopoulos, C., Tumbula, D.L., and Söll, D. 2000. The adaptor hypothesis revisited. *Trends Biochem. Sci.* **25:** 311–316.

James, C.M., Ferguson, T.K., Leykam, J.F., and Krzycki, J.A. 2001. The amber codon in the gene encoding the monomethylamine methyltransferase isolated from *Methanosarcina barkeri* is translated as a sense codon. *J. Biol. Chem.* **276:** 34252–34258.

Jester, B., Levengood, J., Roy, H., Ibba, M., and Devine, K. 2003. Nonorthologous replacement of lysyl-tRNA synthetase prevents addition of lysine analogs to the genetic code. *Proc. Natl. Acad. Sci.* **100:** 14351–14356.

Kobayashi, T., Nureki, O., Ishitani, R., Yaremchuk, A., Tukalo, M., Cusack, S., Sakamoto, K., and Yokoyama, S. 2003. Structural basis for orthogonal tRNA specificities of tyrosyl-tRNA synthetases for genetic code expansion. *Nat. Struct. Biol.* **10:** 425–432.

Köhrer, C., Yoo, J.H., Bennett, M., Schaack, J., and RajBhandary, U.L. 2003. A possible approach to site-specific insertion of two different unnatural amino acids into proteins in mammalian cells via nonsense suppression. *Chem. Biol.* **10:** 1095–1102.

Kryukov, G.V., Castellano, S., Novoselov, S.V., Lobanov, A.V., Zehtab, O., Guigo, R., and Gladyshev, V.N. 2003. Characterization of mammalian selenoproteomes. *Science* **300:** 1439–1443.

Kurland, C.G. 1992. Translational accuracy and the fitness of bacteria. *Annu. Rev. Genet.* **26:** 29–50.

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

LaRiviere, F.J., Wolfson, A.D., and Uhlenbeck, O.C. 2001. Uniform binding of aminoacyl-tRNAs to elongation factor Tu by thermodynamic compensation. *Science* **294:** 165–168.

Lennarz, W.J., Nesbitt III, J.A., and Reiss, J. 1966. The participation of sRNA in the enzymatic synthesis of O-L-lysyl phosphatidylgylcerol in *Staphylococcus aureus. Proc. Natl. Acad. Sci.* **55:** 934–941.

Lin, L., Hale, S.P., and Schimmel, P. 1996. Aminoacylation error correction. *Nature* **384:** 33–34.

Lincecum, T.L., Tukalo, M., Yaremchuk, A., Mursinna, R.S., Williams, A.M., Sproat, B.S., Van Den, E.W., Link, A., Van Calenbergh, S., Grotli, M., et al. 2003. Structural and mechanistic basis of pre- and posttransfer editing by leucyl-tRNA synthetase. *Mol. Cell* **11:** 951–963.

Mayer, C., Köhrer, C., Kenny, E., Prusko, C., and RajBhandary, U.L. 2003. Anticodon sequence mutants of *Escherichia coli* initiator tRNA: Effects of overproduction of aminoacyl-tRNA synthetases, methionyl-tRNA formyltransferase, and initiation factor 2 on activity in initiation. *Biochemistry* **42:** 4787–4799.

Mehl, R.A., Anderson, J.C., Santoro, S.W., Wang, L., Martin, A.B., King, D.S., Horn, D.M., and Schultz, P.G. 2003. Generation of a bacterium with a 21 amino acid genetic code. *J. Am. Chem. Soc.* **125:** 935–939.

Min, B., Pelaschier, J.T., Graham, D.E., Tumbula-Hansen, D., and Söll, D. 2002. Transfer RNA-dependent amino acid biosynthesis: An essential route to asparagine formation. *Proc. Natl. Acad. Sci.* **99:** 2678–2683.

Min, B., Kitabatake, M., Polycarpo, C., Pelaschier, J., Raczniak, G., Ruan, B., Kobayashi, H., Namgoong, S., and Söll, D. 2003. Protein synthesis in *Escherichia coli* with mischarged tRNA. *J. Bacteriol.* **185:** 3524–3526.

Namy, O., Rousset, J.P., Napthine, S., and Brierley, I. 2004. Reprogrammed genetic decoding in cellular gene expression. *Mol. Cell* **13:** 157–168.

Novoselov, S.V., Rao, M., Onoshko, N.V., Zhi, H., Kryukov, G.V., Xiang, Y., Weeks, D.P., Hatfield, D.L., and Gladyshev, V.N. 2002. Selenoproteins and selenocysteine insertion system in the model plant cell system, *Chlamydomonas reinhardtii. EMBO J.* **21:** 3681–3693.

Núñez, H., Lemfil, L., Min, B., Söll, D., and Orellana, O. 2004. In vivo formation of glutamyl-tRNAGln in *Escherichia coli* by heterologous glutamyl-tRNA synthetases. *FEBS Lett.* **557:** 133–135.

Nureki, O., Vassylyev, D.G., Tateno, M., Shimada, A., Nakama, T., Fukai, S., Konno, M., Hendrickson, T.L., Schimmel, P., and Yokoyama, S. 1998. Enzyme structure with two catalytic sites for double-sieve selection of substrate. *Science* **280:** 578–582.

O'Donoghue, P. and Luthey-Schulten, Z. 2003. On the evolution of structure in aminoacyl-tRNA synthetases. *Microbiol. Mol. Biol. Rev.* **67:** 550–573.

Ogle, J.M., Carter, A.P., and Ramakrishnan, V. 2003. Insights into the decoding mechanism from recent ribosome structures. *Trends Biochem. Sci.* **28:** 259–266.

Oku, Y., Kurokawa, K., Ichihashi, N., and Sekimizu, K. 2004. Characterization of the *Staphylococcus aureus* mprF gene, involved in lysinylation of phosphatidylglycerol. *Microbiology* **150:** 45–51.

Petit, J.F., Strominger, J.L., and Söll, D. 1968. Biosynthesis of the peptidoglycan of bacterial cell walls. VII. Incorporation of serine and glycine into interpeptide bridges in *Staphylococcus epidermidis. J. Biol. Chem.* **243:** 757–767.

Polycarpo, C., Ambrogelly, A., Ruan, B., Tumbula-Hansen, D., Ataide, S.F., Ishitani, R., Yokoyama, S., Nureki, O., Ibba, M., and Söll, D. 2003. Activation of the pyrrolysine suppressor tRNA requires formation of a ternary complex with class I and class II lysyl-tRNA synthetases. *Mol. Cell* **12:** 287–294.

Raczniak, G., Becker, H., Min, B., and Söll, D. 2001. A single amidotransferase forms asparaginyl-tRNA and glutaminyl-tRNA in *Chlamydia trachomatis. J. Biol. Chem.* **276:** 45862–45867.

Rao, M., Carlson, B.A., Novoselov, S.V., Weeks, D.P., Gladyshev, V.N., and Hatfield, D.L. 2003. *Chlamydomonas reinhardtii* selenocysteine tRNA[Ser]Sec. *RNA* **9:** 923–930.

Rother, M., Resch, A., Wilting, R., and Böck, A. 2001. Selenoprotein synthesis in archaea. *Biofactors* **14:** 75–83.

Rother, M., Mathes, I., Lottspeich, F., and Böck, A. 2003. Inactivation of the selB gene in *Methanococcus maripaludis*: Effect on synthesis of selenoproteins and their sulfur-containing homologs. *J. Bacteriol.* **185:** 107–114.

Roy, H., Becker, H.D., Reinbolt, J., and Kern, D. 2003. When contemporary aminoacyl-tRNA synthetases invent their cognate amino acid metabolism. *Proc. Natl. Acad. Sci.* **100:** 9837–9842.

Ruusala, T., Ehrenberg, M., and Kurland, C.G. 1982. Is there proofreading during polypeptide synthesis? *EMBO J.* **1:** 741–745.

Salazar, J.C., Ahel, I., Orellana, O., Tumbula-Hansen, D., Krieger, R., Daniels, L., and Söll, D. 2003. Coevolution of an aminoacyl-tRNA synthetase with its tRNA substrates. *Proc. Natl. Acad. Sci.* **100:** 13863–13868.

Salazar, J.C., Ambrogelly, A., Crain, P.F., McCloskey, J.A., and Söll, D. 2004. A truncated aminoacyl-tRNA synthetase modifies RNA. *Proc. Natl. Acad. Sci.* (in press).

Schauer, S., Chaturvedi, S., Randau, L., Moser, J., Kitabatake, M., Lorenz, S., Verkamp, E., Schubert, W.D., Nakayashiki, T., Murai, M., et al. 2002. *Escherichia coli* glutamyl-tRNA reductase. Trapping the thioester intermediate. *J. Biol. Chem.* **277:** 48657–48663.

Schimmel, P. and Ribas De Pouplana, L. 2000. Footprints of aminoacyl-tRNA synthetases are everywhere. *Trends Biochem. Sci.* **25:** 207–209.

Sissler, M., Delorme, C., Bond, J., Ehrlich, S.D., Renault, P., and Francklyn, C. 1999. An aminoacyl-tRNA synthetase paralog with a catalytic role in histidine biosynthesis. *Proc. Natl. Acad. Sci.* **96:** 8985–8990.

Skouloubris, S., De Pouplana, L.R., De Reuse, H., and Hendrickson, T.L. 2003. A noncognate aminoacyl-tRNA synthetase that may resolve a missing link in protein evolution. *Proc. Natl. Acad. Sci.* **100:** 11297–11302.

Srinivasan, G., James, C.M., and Krzycki, J.A. 2002. Pyrrolysine encoded by UAG in Archaea: Charging of a UAG-decodong specialized tRNA. *Science* **296:** 1459–1462.

Stanzel, M., Schön, A., and Sprinzl, M. 1994. Discrimination against misacylated tRNA by chloroplast elongation factor Tu. *Eur. J. Biochem.* **219:** 435–439.

Stortchevoi, A., Varshney, U., and RajBhandary, U.L. 2003. Common location of determinants in initiator transfer RNAs for initiator-elongator discrimination in bacteria and in eukaryotes. *J. Biol. Chem.* **278:** 17672–17679.

Tan, T.H., Bochud-Allemann, N., Horn, E.K., and Schneider, A. 2002. Eukaryotic-type elongator tRNAMet of *Trypanosoma brucei* becomes formylated after import into mitochondria. *Proc. Natl. Acad. Sci.* **99:** 1152–1157.

Thanbichler, M. and Böck, A. 2002. The function of SECIS RNA in translational control of gene expression in *Escherichia coli. EMBO J.* **21:** 6925–6934.

Theobald-Dietrich, A., Frugier, M., Giege, R., and Rudinger-Thirion, J. 2004. Atypical archaeal tRNA pyrrolysine transcript behaves towards EF-Tu as a typical elongator tRNA. *Nucleic Acids Res.* **32:** 1091–1096.

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

Vial, L., Gomez, P., Panvert, M., Schmitt, E., Blanquet, S., and Mechulam, Y. 2003. Mitochondrial methionyl-tRNA[fMet] formyltransferase from *Saccharomyces cerevisiae*: Gene disruption and tRNA substrate specificity. *Biochemistry* **42:** 932–939.

Wang, L., Brock, A., Herberich, B., and Schultz, P.G. 2001. Expanding the genetic code of *Escherichia coli. Science* **292:** 498–500.

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

Wolfson, A.D., LaRiviere, F.J., Pleiss, J.A., Dale, T., Asahara, H., and Uhlenbeck, O.C. 2001. tRNA conformity. *Cold Spring Harb. Symp. Quant. Biol.* **66:** 185–193.

Wong, F.C., Beuning, P.J., Silvers, C., and Musier-Forsyth, K. 2003. An isolated class II aminoacyl-tRNA synthetase insertion domain is functional in amino acid editing. *J. Biol. Chem.* **278:** 52857–52864.

Zavacki, A.M., Mansell, J.B., Chung, M., Klimovitsky, B., Harney, J.W., and Berry, M.J. 2003. Coupled tRNA[Sec]-dependent assembly of the selenocysteine decoding apparatus. *Mol. Cell* **11:** 773–781.

**Erratum**

**Genes & Development 18:** 731–738 (2004)

**Aminoacyl-tRNAs: setting the limits of the genetic code**
Michael Ibba and Dieter Söll

In the above-mentioned paper, an incorrect e-mail address was given for Michael Ibba. The correct e-mail address is ibba.1@osu.edu.

# Aminoacyl-tRNAs: setting the limits of the genetic code

Michael Ibba and Dieter Söll

| | |
|---|---|
| **Related Content** | **Errata for vol. 18, p. 731**<br>Genes Dev. May , 2004 18: 1106 |
| **References** | This article cites 73 articles, 41 of which can be accessed free at:<br>**http://genesdev.cshlp.org/content/18/7/731.full.html#ref-list-1**<br><br>Articles cited in:<br>**http://genesdev.cshlp.org/content/18/7/731.full.html#related-urls** |
| **License** | |
| **Email Alerting Service** | Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or  **click here.** |