Archaeal Aminoacyl-tRNA Synthesis: Diversity Replaces Dogma

Debra Tumbula,* Ute C. Vothknecht,* Hyun-soo Kim,* Michael Ibba,¹ Bokkee Min,* Tong Li,* Joanne Pelaschier,* Constantinos Stathopoulos, Hubert Becker* and Dieter Söll²

*Department of Molecular Biophysics and Biochemistry and ¹Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, Connecticut 06520-8114 and ²Center for Biomolecular Recognition, IMBG, The Panum Institute, DK-2200 Copenhagen N, Denmark

Manuscript received March 25, 1999
Accepted for publication May 5, 1999

ABSTRACT

Accurate aminoacyl-tRNA synthesis is essential for faithful translation of the genetic code and consequently has been intensively studied for over three decades. Until recently, the study of aminoacyl-tRNA synthesis in archaea had received little attention. However, as in so many areas of molecular biology, the advent of archaeal genome sequencing has now drawn researchers to this field. Investigations with archaea have already led to the discovery of novel pathways and enzymes for the synthesis of numerous aminoacyl-tRNAs. The most surprising of these findings has been a transamidation pathway for the synthesis of asparaginyl-tRNA and a novel lysyl-tRNA synthetase. In addition, seryl- and phenylalanyl-tRNA synthetases that are only marginally related to known examples outside the archaea have been characterized, and the mechanism of cysteinyl-tRNA formation in Methanococcus jannaschii and Methanobacterium thermoautotrophicum is still unknown. These results have revealed completely unexpected levels of complexity and diversity, questioning the notion that aminoacyl-tRNA synthesis is one of the most conserved functions in gene expression. It has now become clear that the distribution of the various mechanisms of aminoacyl-tRNA synthesis in extant organisms has been determined by numerous gene transfer events, indicating that, while the process of protein biosynthesis is orthologous, its constituents are not.

The synthesis of aminoacyl-tRNAs is one of the core reactions of gene expression. By matching the anticodon trinucleotide of a tRNA with a particular amino acid, the aminoacyl-tRNAs allow the translation of a series of mRNA codons into a corresponding polypeptide sequence. These small RNAs, which have an amino acid attached to the adenosine at their 3′-end, are the sole substrates for ribosomal translation of mRNA. Consequently, the fidelity of translation is highly dependent on the accuracy of aminoacyl-tRNA synthesis. The majority of aminoacyl-tRNAs are directly synthesized by the aminoacyl-tRNA synthetases, a family of 20 enzymes (1 discrete enzyme per amino acid-tRNA pair) that attach the cognate amino acid to the tRNA. However, in a number of organisms and organelles, glutaminyl- or asparaginyl-tRNAs are formed in an indirect pathway (a tRNA-dependent amino acid transformation) that requires the modification of an incorrectly charged amino acid while bound to tRNA (see below). The aminoacyl-tRNA synthetases have been studied in considerable detail, resulting in a clear picture of the mechanisms that ensure that only a particular amino acid is attached to certain tRNAs. With only a few notable exceptions, the essential nature of the aminoacyl-tRNA synthetases had long been assumed to limit their variability in living organisms. This assumption was initially borne out when no fundamental difference was seen between the aminoacyl-tRNA synthetases of numerous organisms as diverse as Escherichia coli, yeast, and human. Not surprisingly, until recently most studies of the aminoacyl-tRNA synthetases had been mainly confined to these model organisms. In particular, very little information had accumulated on the mechanisms of aminoacyl-tRNA synthesis in archaea—and none of the data had suggested that studying this process in such notoriously intractable organisms would be a particularly worthwhile endeavor. Any such preconceptions were destroyed after the publication in 1996 of the first genomic sequence from an archaean, that of Methanococcus jannaschii (Bult et al. 1996).

The sequence of the M. jannaschii genome provided many surprises, among them the observation that the genome encoded only 16 known aminoacyl-tRNA synthetase homologs. The “missing” synthetases were the asparaginyl- (AsnRS), cysteinyl- (CysRS), glutaminyl- (GlnRS), and lysyl-tRNA synthetases (LysRS). In addition, the assignment of an open reading frame (ORF) encoding seryl-tRNA synthetases (SerRS) was based on rather low homology, while there appeared to be three reasonable candidates to encode the two subunits of phenylalanyl-tRNA synthetase (PheRS). Furthermore, key components of the selenocysteinyl-tRNA synthesis pathway could not be identified. Archaeal protein synthesis utilizes the 20 canonical amino acids (as well as
aspartyl-tRNA synthetase (AspRS) forms the misamino- (Figure 1), a nondiscriminating glutamyl-(GluRS) or an indirect way of making Gln-tRNA. This observation transamidation pathway. have thwarted a more universal recognition of this fact.

trachomatis nature of the enzymes involved in this pathway may established that Gln-tRNA$_{Gln}$ and Asn-tRNA$_{Asn}$ formation can archaea have unlinked genetic description of some of these enzymes and the Thermus thermophilus

gatCAB

genes for its three subunits are arranged in an operon (gatCAB). The Gat proteins have no similarity to the known glutamine amidotransferases (Curnow et al. 1997). The A subunit has sequence similarity to certain bacterial amidases, but the known characterized enzymes (Kobayashi et al. 1993) do not use glutamine or asparagine as substrates. However, GatA (coexpressed with GatC) does convert glutamine to glutamate (A. W. Curnow, unpublished results). On the basis of this information it was proposed that the A subunit functions as amide donor by hydrolyzing glutamine, while the B subunit may be a tRNA-binding protein. The role of the C subunit may be in the stabilization or correct folding of the A subunit (Curnow et al. 1997).

How many subunits are needed for activity? Genomic analysis suggests that the yeast mitochondrial Glu-AdT has only an A and a B subunit. However, the lack of a sequence with similarity to gatC does not rule out the presence of a GatC-like protein structure encoded by a dissimilar sequence. Yet a Thermus thermophilus Asp-AdT enzyme was highly active, but contained only the two subunits corresponding to gatA and gatB (Becker and Ker n 1998). Thus, GatA/ GatB heterodimers may also be active, a fact to be kept in mind when searching in genomes for ORFs responsible for the tRNA-dependent transamidation pathway.

A look at genomic sequences reveals that bacteria carry two forms of gat gene arrangements: some of them have operon structures like B. subtilis (e.g., Chlamydia trachomatis), while in others the genes for the subunits are dispersed throughout the chromosome (e.g., Deinococcus radiodurans). The latter gene arrangement is the one found in archaea; currently most known sequenced archaea have unlinked gatA, gatB, and gatC genes. Additionally, a gatB ortholog and a gatB paralog are in each of the genomic sequences of M. jannaschii, Methanococcus thermooautotrophicum $\Delta$H, and Archaeoglobus fulgidus (Bul t et al. 1996; Klenk et al. 1997; Smit h et al. 1997). In comparison, the only discernible gat gene in the genome of Pyrococcus horikoshii, which lacks GlnRS, is gatB (Kawarabay si et al. 1998); thus the pathways of Gln-tRNA formation in this organism remain to be established. However, an AsnRS gene is present in the Pyrococcus genome. From these genome examinations it is clear that all archaea lack GlnRS and instead use the transamidation pathway (Table 1). However, none of the archaearal transamidation enzymes have been puriﬁed and characterized, which is necessary to reveal the
**TABLE 1**

Biochemical evidence for aminoacyl-tRNA formation by transamidation

<table>
<thead>
<tr>
<th>Aminoacyl-tRNA</th>
<th>Organisms and references</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gln-tRNA</td>
<td>Archaea&lt;sup&gt;a,b&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td><em>Haloferax volcanii</em> (&lt;span&gt;Curnow&lt;/span&gt; et al. 1996)</td>
</tr>
<tr>
<td></td>
<td><em>Halobacterium cutirubrum</em> (White and Bayley 1972)</td>
</tr>
<tr>
<td></td>
<td><em>Methanococcus maripaludis</em>&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td><em>Methanothermbacterium thermoautotrophicum</em>&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Bacteria</td>
<td>Gram-positive&lt;sup&gt;1&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td><em>Bacillus subtilis</em> (Wilcox and Nirenberg 1968)</td>
</tr>
<tr>
<td></td>
<td><em>Bacillus megaterium</em> (Wilcox and Nirenberg 1968)</td>
</tr>
<tr>
<td></td>
<td><em>Lactobacillus acidophilus</em> (Wilcox and Nirenberg 1968)</td>
</tr>
<tr>
<td></td>
<td><em>Lactobacillus bulgaricus</em> (Schön et al. 1988b)</td>
</tr>
<tr>
<td></td>
<td><em>Streptococcus faecalis</em> (Wilcox 1969)</td>
</tr>
<tr>
<td></td>
<td><em>Streptomyces antibioticus</em> (Wilcox 1969)</td>
</tr>
<tr>
<td>Gram-negative</td>
<td><em>Rhizobium meliloti</em> (Gagnon et al. 1996)</td>
</tr>
<tr>
<td></td>
<td><em>Helicobacter pylori</em>&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Exceptions: <em>Escherichia coli</em> (Lazzarini and Mehe 1964), <em>Azotobacter vinelandii</em>, and <em>Proteus vulgaris</em> (Wilcox 1969)</td>
</tr>
<tr>
<td>Synechocystis sp. strain PCC6803&lt;sup&gt;g&lt;/sup&gt; (Schön et al. 1988a)</td>
<td></td>
</tr>
<tr>
<td>Eukaryal organelles</td>
<td>Chloroplasts&lt;sup&gt;1&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td><em>Barley</em> (Schön et al. 1988a)</td>
</tr>
<tr>
<td></td>
<td><em>Spinach</em> (Schön et al. 1988a)</td>
</tr>
<tr>
<td></td>
<td><em>Chlamydomonas reinhardtii</em> (Jahn et al. 1990)</td>
</tr>
<tr>
<td>Mitochondria</td>
<td>Yeast (Martin and Rabinowicz 1984)</td>
</tr>
<tr>
<td></td>
<td><em>Mouse</em> (Schön et al. 1988a)</td>
</tr>
<tr>
<td></td>
<td><em>Mung bean</em> (Schön et al. 1988a)</td>
</tr>
<tr>
<td></td>
<td>Exceptions: <em>Leishmania tarentolae</em> and <em>Trypanosoma brucei</em> (Nabholz et al. 1997)</td>
</tr>
<tr>
<td>Asn-tRNA</td>
<td>Archaea&lt;sup&gt;a,b&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td><em>Haloferax volcanii</em> (&lt;span&gt;Curnow&lt;/span&gt; et al. 1996)</td>
</tr>
<tr>
<td></td>
<td><em>Methanococcus maripaludis</em>&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td><em>Methanothermbacterium thermoautotrophicum</em>&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Bacteria&lt;sup&gt;1&lt;/sup&gt;</td>
<td><em>Deinococcus radiodurans</em> (&lt;span&gt;Curnow&lt;/span&gt; et al. 1998)</td>
</tr>
<tr>
<td></td>
<td><em>Thermus thermophilus</em> (Becker and Kern 1998)</td>
</tr>
<tr>
<td></td>
<td><em>Helicobacter pylori</em>&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup> Within this group, exceptions to this transamidation route have yet to be demonstrated biochemically.

<sup>b</sup> Except for Pyrobaculum aerophilum and Pyrococcus horikoshii, archaeal genomic sequences lack GlnRS and AsnRS genes and contain homologs of *gatC, gatA,* and *gatB* (Bult et al. 1996; Klenk et al. 1997; Smith et al. 1997).

<sup>c</sup> A. W. Curnow, M. Ibba and D. Söll (unpublished results).


<sup>f</sup> The genomic sequence of *E. coli* has a GlnRS gene and lacks homologs of the *gat* genes (Blattner et al. 1997).

<sup>g</sup> The Synechocystis genome sequence lacks a GlnRS gene and contains homologs of *gatC, gatA,* and *gatB* (Kaneko et al. 1996).

<sup>h</sup> Most bacterial genomic sequences contain an AsnRS gene.

exact details of how these organisms make Gln-tRNA and Asn-tRNA.

Two transamidases of differing specificities (Glu-AdT and Asp-AdT) are suggested in some archaea by sequence data. According to this the two enzymes would share the C and A subunits, but would differ in the B subunit; an ortholog (to *B. subtilis*) for Glu-AdT and a paralog for Asp-AdT. Both transamidation pathways have been demonstrated by tRNA-dependent amino acid transformations of labeled substrates in cell extracts.
of H. volcanii, Methanococcus maripaludis, and M. thermoautotrophicum Marburg (Curnow et al. 1996; H. D. Becker, A. W. Curnow, M. Ibba and D. Tumbula, unpublished results). However, this clear distinction has been muddied recently by the demonstration that, in addition to Gln-tRNA synthesis, the gatCAB-encoded Glu-AdT enzymes from B. subtilis and D. radiodurans can also carry out the transamidation of Asp-tRNAAsn to Asn-tRNAAsn (Curnow et al. 1998). Thus, there is no necessity for the pyrococcal enzyme also displays the unusual tRNA in the aspartylation reaction (related to the same aminoacyl-tRNA? The initial assumption was “yes,” and, in line with the incidence of the direct pathways, direct glutaminylatin and asparaginylation of tRNA from H. volcanii and H. cutirubrum could not be demonstrated (Griffiths and Bayley 1969; Gupta and Curnow 1984; Curnow et al. 1996). However, contrary to initial expectations, bacteria have now been found that contain both direct and indirect pathways of Gln-tRNA and Asn-tRNA formation (Becker and Kern 1998; Curnow et al. 1998). Without further biochemical and metabolic investigations no conclusions can be reached on the respective roles of these pathways. It is clear that studies in the archaea will clarify the issue of number and nature of tRNA-dependent amidotransferases.

Acetic acid, a discriminatory synthetase, is the required presence of nondiscriminating GluRS or AspRS enzymes. Fortunately, the first crystal structure of an archaeal aminoacyl-tRNA synthetase was that of AspRS from Pyrococcus kodakaraensis KOD1 (Schmitt et al. 1998). Knowledge of the structures of the native enzyme and its complexes with ATP and aspartic acid suggested a basis for the degenerate specificity of a naturally occurring misacylating enzyme. A loop in yeast AspRS, a discriminating synthetase, which specifically interacts with the tRNA anticodon (positions 36 and 37) is shortened in the corresponding region of the pyrococcal AspRS. Modeling the pyrococcal structure with tRNA suggests that the loop and anticodon do not closely interact in the archaeal AspRS. In analogy to its closely related ortholog in T. thermophilus the pyrococcal AspRS is a nondiscriminating enzyme (Becker and Kern 1998). The pyrococcal enzyme also displays the unusual characteristic that ATP can be substituted by GTP or UTP in the aspartylation reaction (Fujiwara et al. 1996). No information is yet available for the archaea, or indeed any other, nondiscriminating GluRS.

**Lysyl-tRNA synthetase:** The absence of AsnRS and GlnRS from archaeal genomes was easily explained by their replacement with tRNA-dependent transamidation pathways (see above). The lack of an obvious ortholog for LysRS was initially more problematic (Ibba et al. 1997a). The possibility that LysRS was absent, with Lys-tRNA instead made via mischarging, seemed unlikely because there exist no obvious pathways to easily transform another amino acid into lysine (particularly at the 3′-end of a tRNA). These ideas, which suggest that LysRS must be present in some form, were confirmed by the observation that cell-free extracts of both H. volcanii and M. maripaludis are able to directly attach lysine to tRNA (M. Ibba and A. W. Curnow, unpublished results). This assay was then used as the basis for the purification of M. maripaludis LysRS, the N-terminal sequencing of which allowed the cloning of the corresponding gene (lysS). Heterologous expression of this gene in E. coli gave rise to a protein with canonical LysRS activity. The sequence of the M. maripaludis lysS gene showed no similarity to known LysRS-encoding genes, but was highly homologous to unassembled open reading frames in all of the archaeal genomes that lacked an identifiable LysRS (Ibba et al. 1997b). Interestingly, this group of archaeal lysS genes all encoded proteins containing sequence motifs characteristic of class I aminoacyl-tRNA synthetases, in contrast to all other then-known LysRS proteins that were members of class II. The division of the aminoacyl-tRNA synthetases into two classes is based upon distinct and unrelated active site topologies, with membership of a particular class being conserved throughout the living kingdom (Erian i et al. 1990; the so-called “class rule”). The violation of the class rule (of which this remains the only known example to date) by the existence of both class I and class II LysRS proteins has subsequently been confirmed by various biochemical approaches (Ibba et al. 1999).

The initial finding that many, but not all, archaea contain a novel LysRS led to the suggestion that this protein might be confined to the archaeal kingdom. However, subsequent studies arising from bacterial genomics have shown that the class I LysRS is also present in several bacteria to the exclusion of the more common class II protein (Ibba et al. 1997c). Comparisons of the predicted amino acid sequences for the 15 known class I LysRS proteins show that the bacterial examples do not cluster together but instead form subgroups most closely related to particular groups of archaeal proteins (Figure 2). The class I LysRS is found in both Euryarchaeota (two species of Methanococcaceae and Pyrococcus, M. thermoautotrophicum, Thermoplasma acidophilum, Methanosarcina mazei, and A. fulgidus) and in Crenarchaeota (a species of Cenarchaeum and Aeropyrum pernix). The known bacterial class I lysS genes are found in disparate taxa, the spirochetes (Borrelia burgdorferi and Treponema pallidum), a G + C rich gram-positive bacterium (Streptomyces coelicolor) and the α-proteobacteria (Rickettsiia prowazekii and Rhodobacter capsulatus). The euryarchaeal
versions of the enzyme are a coherent phylogenetic grouping in agreement with the 16S rRNA phylogeny (Olson et al. 1994; Barns et al. 1996) from which the crenarchaeal examples are distant. In contrast, the bacterial LysRS proteins do not form a single grouping. The relationships between the various class I LysRS proteins suggest that the bacterial examples have been acquired by at least two horizontal gene transfer events from archaea, with the existing class II LysRS subsequently displaced (Ibba et al. 1999). Thus, the LysRS in α-proteobacteria may have arisen from a crenarchaeal LysRS and that in the spirochetes and Streptomyces from the pyrococcal LysRS.

As mentioned above, not all archaea possess a class I-type LysRS. For example, the crenarchaeotes Pyrobaculum aerophilum (S. Fitz-Gibbon, unpublished data) and Sulfolobus solfataricus (Sensen et al. 1996; C. W. Sensen, personal communication) both contain only class II LysRS proteins. Closer examination of these crenarchaeal sequences shows that they are more similar to each other than to any other LysRS and that they have reasonably high homology to bacterial proteins (e.g., both show over 43% identity at the amino acid level to Thermotoga maritima LysRS). This last observation in particular suggests that these crenarchaeal class II LysRS might have been acquired from bacteria.

The apparent two-way transfer of the different LysRS encoding genes, both to and from the archaeal kingdom, has provided clear evidence of the dynamic nature of aminoacyl-tRNA synthetase evolution. This is in stark contrast to previous predictions that suggested that the obvious link between the genetic code and the aminoacyl-tRNA synthetases would limit the variation of this enzyme family. What instead has emerged from studies of the class I LysRS is clear evidence that the tRNAs themselves predate the aminoacyl-tRNA synthetases (Ibba et al. 1999; Schimmel and Ribas de Pouplana 1999) and by implication must be the true determinants of the genetic code. This is borne out by the observations that the class I and class II LysRS proteins can recognize the same tRNALys in vitro and in vivo (Ibba et al. 1999). The genomic sequences of P. horikoshii and Pyrococcus furiosus exhibit one gene each for the two PheRS subunits (Kawarabayasi et al. 1998; WIT1 ORFs RPF01697 and RPF01699). However, each of the genome se-

1 As these sequences are not yet in the GenBank database, the WIT ORF number is given; see http://wit.mcs.anl.gov/WIT2/.
quences of M. jannaschii, M. thermoautotrophicum \( \Delta H \), and A. fulgidus contains three ORFs with similarity to PheRS subunits (Bult et al. 1996; Klenc et al. 1997; Smith et al. 1997). One ORF (MTH0770) displayed high sequence similarity to bacterial and eukaryal PheRS \( \beta \)-subunits. In addition, there are two ORFs (e.g., MTH0742 and MTH1501) with similarity to \( \alpha \)-subunits from non-archaea. Their sizes (in kilodaltons) are 58.5 and 61, respectively. Which of these potential \( \alpha \)-subunits is found in the active PheRS?

We addressed this question by purifying PheRS from M. thermoautotrophicum Marburg. SDS-polyacrylamide gel electrophoresis of the purified enzyme revealed two proteins in about equimolar ratio (U.C. Vothknecht, unpublished results). The aminoterminal sequences were obtained for both proteins. The sequence of the 70-kD protein matched that of ORF MTH0770 (which has sequence similarity with known PheRS \( \beta \)-subunits), while the N-terminal sequence of the 60-kD protein corresponded to that of ORF MTH0742. Thus, this ORF encodes the PheRS \( \alpha \)-subunits used in this organism. It is likely that in M. jannaschii and A. fulgidus the PheRS \( \alpha \)-subunit is encoded by the ORFs with the highest homology to the M. thermoautotrophicum protein (i.e., MJO487 and AF1955).

What is the function of the protein encoded by ORF MTH1501 in M. thermoautotrophicum and the other euryarchaeaotes? It may be a pseudogene, although the highly conserved sequence suggests a functional protein. Alternatively, this ORF may encode a synthetase-like protein with unknown function. There are several examples of the involvement of synthetase-like proteins in reactions other than tRNA aminoacylation (see, e.g., Girich et al. 1997). However, not much is known to date about the function of these proteins.

**Seryl-tRNA synthetase:** SerRS is a class II synthetase characterized by the three signature motifs characteristic for all class II enzymes (Cusack et al. 1996). Surprisingly, three kinds of SerRS are found in the archaea (Doolittle and Handy 1998; Kim et al. 1998). The enzyme from the halophilic archaeon Haloarcula marismortui was the first archaeal SerRS to be purified and its gene sequence identified (Taupin et al. 1997). It possesses high sequence similarity to bacterial SerRS proteins and clusters within the bacterial branch in maximum likelihood analyses. A different clad comprising the SerRS enzymes from several other archaea (including P. horikoshii, P. furiosus, P. aerophilum, and A. fulgidus) shares high sequence similarity to its eukaryotic and bacterial counterparts (Klenc et al. 1997; WIT database; S. Fitz-Gibbon, unpublished data; Kawai et al. 1997). In contrast, the three methanogenic archaea M. jannaschii, M. maripaludis, and M. thermoautotrophicum contain SerRS-encoding genes (serS) that display only low sequence similarity to the known serS genes from all other organisms and encode a structurally uncommon enzyme with an altered motif II (Kim et al. 1998). Biochemical analysis of purified and heterologously expressed SerRS from M. thermoautotrophicum and M. maripaludis showed that these proteins in fact have canonical SerRS activity (Kim et al. 1998). While it has been demonstrated that the M. thermoautotrophicum and M. maripaludis SerRS enzymes aminoacylate E. coli tRNA, detailed biochemical, mutational, and structural studies are warranted to elucidate the structure-function relationships in this unusual SerRS enzyme.

**Cysteinyl-tRNA synthetase:** Biochemical, genetic, and genomic studies have revealed that many archaea contain a relatively recent bacterial version of CysRS and that lateral transfer of this gene between bacteria and archaea was widespread (Li et al. 1999). However, as discussed above, there are two organisms (M. jannaschii and M. thermoautotrophicum) with known complete genomic sequences that do not contain an ORF for a recognizable CysRS. Barring the possibility that the published genomes miss some genes, one is led to believe that these organisms must possess a currently unrecognizable enzyme responsible for Cyst-RNA formation. Because domain searches have not revealed unidentified ORFs with clear class I and class II signatures (except the unassigned PheRS-like ORF; see above), such an enzyme would not have similarity with the known synthetases.

On the other hand, an indirect pathway of Cyst-RNA\(^{5s}\) formation by thiolation of Ser-tRNA\(^{5s}\) analogous to the biosynthetic route for selenocysteinyltRNA was still a possibility (Baron and Bock 1995). For this reason it was intriguing to realize that the structurally uncommon SerRS (see above) is found in organisms whose genomic sequence lacks a recognizable gene encoding CysRS (M. jannaschii and M. thermoautotrophicum). However, in vitro experiments with pure M. thermoautotrophicum SerRS did not provide support for mischarging of tRNA\(^{5s}\) with serine (Kim et al. 1998). Nevertheless, M. maripaludis, a relative of M. jannaschii for which a genomic sequence is not yet available, contains a canonical CysRS with similarity to the pyrococcal enzyme (C. Stathopoulos, D. Graham, P. Haney and A. W. Curnow, unpublished results). However, this does not rule out the existence of another CysRS in this organism or the presence of a tRNA-dependent amino acid transformation pathway of Cyst-RNA formation in some archaea. Defining the pathway of Cyst-RNA synthesis in M. jannaschii and M. thermoautotrophicum is the last great challenge to our knowledge of aminoacyl-tRNA synthesis; it appears to be strictly an archaeal mystery.

**Selenocysteine formation:** Selenocysteine, a UGA-directed, cotranslationally inserted, essential amino acid, is also found in many archaea (see, e.g., Wilting et al. 1997). The mechanism of selenocysteine formation in archaea is currently unknown. While orthologs to the E. coli selD (selenophosphate synthase) and selC (selenocysteine tRNA) genes have been found, archaea appear to employ nonorthologous proteins in the function of
**Future prospects:** The study of aminoacyl-tRNA synthesis in archaea has questioned, and in some cases overturned, several dogmatic views. The question now arises, since nearly all of the puzzles initially presented by the genomic sequence of *M. jannaschii* have been solved, should we return to our trusted model organisms such as *E. coli*? The answer is most certainly not, for the simple reason that the investigation of archaea has made it abundantly clear that the concept of a model organism for the study of synthetases is now redundant. For example, while it was first shown in archaea that a transaminase pathway can synthesize Asn-tRNA and that a class I LysRS exists, both have subsequently proved to be common in bacteria. Similarly, the apparently universal absence of GlnRS from archaea, together with its scarcity in bacteria, has confirmed the prediction that this is essentially a eukaryotic enzyme (Siat et al. 1998). These examples clearly illustrate how the study of aminoacyl-tRNA synthesis in archaea has profoundly altered our preconceptions about this essential process.

To date, the utilization of alternative pathways and enzymes to synthesize particular aminoacyl-tRNAs shows no predictable distribution in the living kingdom (Table 1), with some bacteria even employing two pathways to make the same Asn-tRNA (Becker and Kern 1998; Curnow et al. 1998). This lack of a consistent evolutionary pattern is seen among virtually all the aminoacyl-tRNA synthetase families (Diaz-Lacoz et al. 1998; C. Woese, G. Olsen, M. Ibba and D. Söll, unpublished results). To better understand these unexpected levels of both biochemical and evolutionary complexity in aminoacyl-tRNA synthesis, the sampling of a wider range of organisms is now essential: the expansion of experimental work on archaeal aminoacyl-tRNA synthesis forms an integral part of such studies. In light of the numerous unexpected results that have already come from working with archaea, it seems reasonable to assume that more surprises are in store for the aminoacyl-tRNA synthesis field.

We thank Rajdeep Das for help with PhRs purification. The work described in this article was supported by grants from the National Institute of General Medical Sciences and Bristol-Myers Squibb (D.S.). D.L.T. holds a postdoctoral fellowship from the National Institute of General Medical Sciences, and H.B. is a European Molecular Biology Organization postdoctoral fellow. M.I. gratefully acknowledges the support of the Alfred Benzon Foundation.

**LITERATURE CITED**


Cusack, S., A. Yamarchuk and M. Tukalo, 1996 The crystal structure of the ternary complex of *Thermus thermophilus* seryl-tRNA synthetase with tRNA<sup>38</sup> and a seryl-adenylate analogue reveals a conformational switch in the active site. EMBO J. 15: 2834–2842.


Communicating editor: W. B. Whitman