CMLS, Cell. Mol. Life Sci. 61 (2004) 2437–2445
1420-682X/04/202437-09
DOI 10.1007/s00018-004-4194-9

CMLS Cellular and Molecular Life Sciences

Review

Cys-tRNA^{Cys} formation and cysteine biosynthesis in *methanogenic archaea*: two faces of the same problem?

A. Ambrogelly [a, *], S. Kamtekar [a], A. Sauerwald [a], B. Ruan [a], D. Tumbula-Hansen [a, †], D. Kennedy [a], I. Ahel [a] and D. Söll [a, b, *]

[a] Department of Molecular Biophysics and Biochemistry, Yale University, P.O. Box 208114, 266 Whitney Avenue, New Haven, Connecticut 06520-8114 (USA), Fax: +1 203 432 6202, e-mail: soll@trna.chem.yale.edu
[b] Chemistry, Yale University, New Haven, Connecticut 06520-8114 (USA)

Received 4 May 2004; received after revision 27 May 2004; accepted 28 May 2004

Abstract. Aminoacyl-tRNA (transfer RNA) synthetases are essential components of the cellular translation machinery as they provide the ribosome with aminoacyl-tRNAs. Aminoacyl-tRNA synthesis is generally well understood. However, the mechanism of Cys-tRNA^{Cys} formation in three methanogenic archaea (*Methanocaldococcus jannaschii*, *Methanothermobacter thermautotrophicus* and *Methanopyrus kandleri*) is still unknown, since no recognizable gene for a canonical cysteinyl-tRNA synthetase could be identified in the genome sequences of these organisms. Here we review the different routes recently proposed for Cys-tRNA^{Cys} formation and discuss its possible link with cysteine biosynthesis in these methanogenic archaea.

Key words. Cys-tRNA^{Cys}; cysteine biosynthesis; cysteinyl-tRNA synthetase; archaea; methanogens.

Introduction

Faithful translation of the genetic information into proteins relies on the precise matching of the messenger RNA (mRNA) codons with their corresponding amino acids in accordance with the rules of the genetic code. To achieve this, transfer RNA (tRNA) through its anticodon specifically pairs with the mRNA codon, and when aminoacylated at its 3′ end, delivers the corresponding amino acid to the ribosome where the peptide bonds are formed. The overall fidelity of gene expression depends on the correct association of amino acids with their corresponding tRNA by the enzymes collectively known as the aminoacyl-tRNA synthetases (AARSs). As proteins are composed of 20 standard (canonical) amino acids, 20 members of this enzyme family exist – at least – in eukaryotes, one for each amino acid [1]. The AARSs generate aminoacyl-tRNAs (AA-tRNAs) in a two-step mechanism, in which the amino acid is first activated in presence of ATP to form an enzyme-bound aminoacyl-adenylate. The activated amino acid is subsequently transferred to one of the ribose hydroxyl groups at the 3′ terminus of the cognate tRNA molecule with the release of AMP. The AARSs are divided into two structurally unrelated groups, class I and class II, based on mutually exclusive sequence motifs that reflect distinct active site topologies and different mechanistic features such as ATP conformation, mode of tRNA binding and regiospecificity of the transfer reaction [2, 3]. Although discrepancies in the mechanism of formation of Lys-tRNA [4], Asn-tRNA and Gln-tRNA (reviewed in [5]) are well known, the absence of a canonical cysteinyl-tRNA synthetase (CysRS) and the mechanism by which Cys-tRNA is made has not yet been satisfactorily resolved in three

* Corresponding authors.
† Present address: Medical University of South Carolina, Department of Biochemistry and Molecular Biology, Charleston, South Carolina, 29412 (USA)

methanogenic archaea, *Methanocaldococcus jannaschi*, *Methanothermobacter thermautotrophicus* and *Methanopyrus kandleri*. Here we will review the recently proposed solutions to this remaining puzzle.

Cys-tRNA^{Cys} formation in nature

With the exception of *M. jannaschi* [6], *M. thermautotrophicus* [7] and *M. kandleri* [8], in all the organisms for which the genome sequences are known, Cys-tRNA^{Cys} formation is carried out by CysRS, a small canonical class I synthetase. The crystal structure of the *Escherichia coli* CysRS sheds light on specific cysteine recognition by the enzyme [9]. The structure revealed the presence of a zinc ion in the active site, coordinated via amino acid residues Cys28, Cys209, His234 and Glu238; the metal ion is also capable of coordinating with the sulfhydryl group of the cysteine substrate. The specific recognition of cysteine via metal ion coordination excludes alternative amino acids of similar molecular volume that are unable to coordinate the zinc ion, avoiding the need for any editing mechanism [9, 10].

As inferred from its phylogeny (fig. 1), CysRS does not correspond with the 16S ribosomal RNA (rRNA) phylogeny because archaeal CysRSs resolve poorly from their bacterial and eukaryal counterparts [11]. Such a discrepancy between the CysRS phylogeny and the usually accepted microbial taxonomy led to the conclusion that *cysS*, the gene encoding CysRS, was involved in considerable interdomain horizontal gene transfer, mainly affecting the archaeal domain [11]. While the phylogenetic repartition of the bacterial CysRS is in relatively good accordance with the usual bacterial phylogeny, the scattered phylogenetic repartition of the archaeal CysRS suggests several independent transfer events of *cysS* from bacteria into archaea. The existence of a canonical CysRS of bacterial origin in all but three archaea remains puzzling. In the absence of any data on tRNA cysteinylation in these three organisms, the molecular mechanisms underlying the vast dispersal of the bacterial *cysS* into the archaeal domain and the displacement of an archaeal specific cysteinylation route remain obscure. However, genetic experiments in *Methanococcus maripaludis*, a close relative of *M. jannaschii*, furthered our understanding of a possible *cysS* gene transfer scenario. In a *M. maripaludis* genetic *cysS* knockout strain, the fully functional CysRS was shown to be dispensable [12], suggesting the existence of a redundant tRNA cysteinylation route, possibly homologous to the one present in *M. jannaschii*. The CysRS phylogenetic repartition, together with the genetic analysis done in *M. maripaludis*, suggests that the archaeal acquisition of *cysS* is a relatively recent and still ongoing evolutionary event. The idea of such an incomplete evolutionary process is consistent with the proposals that cysteine was one of the late amino acids added to the genetic code [13, 14] and that the frequency of cysteine codons may have more than doubled since the last universal common ancestor [15, 16]. This pressure for higher levels of cysteine incorporation into proteins over a relatively short period might have led to the recruitment by archaeal organisms of a more efficient cysteinylation system from bacteria, explaining the modern phylogeny of the CysRSs.

Cysteine biosynthesis in nature

Several biosynthetic routes to cysteine have been identified in bacteria, eukarya and a few archaea. In bacteria, cysteine is synthesized starting from serine and then transformed into cysteine via a two-step pathway (fig. 2A) [17]. In the first step, the hydroxyl group of serine is activated in the form of *O*-acetylserine by a serine transacetylase (SAT, encoded by *cysE)* in the presence of acetyl-coenzyme A (CoA). In a second step, the methyl ester is displaced in the presence of a sulfur donor (sulfide or thiosulfate) by a pyridoxal-5′-phosphate (PLP)-dependent *O*-acetylserine sulfhydrylase (OASS-A or OASS-B, encoded by *cysK* and *cysM*, respectively) to yield cysteine and acetate as a by-product of the reaction. In contrast to plants, which use a pathway similar to bacteria (reviewed in [18]), in mammals, cysteine is synthesized via a transsulfuration pathway that utilizes methionine for the sulfhydryl function and serine for the α-amino acid stereocenter and side chain carbon atoms [19] (fig. 2B). Methionine is first converted to homocysteine through the intermediates *S*-adenosyl-methionine and *S*-adenosyl-homocysteine. Cystathionine β-synthase (CBS), a PLP-dependent enzyme, then combines homocysteine and serine to form cystathionine, which yields cysteine and α-ketobutyrate upon the action of another PLP-dependent enzyme, cystathionine γ-ligase (CGL).

Saccharomyces cerevisiae was shown to use a hybrid bacterial- and mammalian-like route for cysteine biosynthesis (reviewed in [20]). This organism can fix sulfur by synthesizing homocysteine starting from homoserine via *O*-acetylhomoserine in a pathway formally homologous to the one used for cysteine biosynthesis in bacteria (fig. 2C). *S. cerevisiae* possesses the homologs of CBS and CGL that allow transformation of homocysteine into cysteine in a pathway similar to the one present in mammalian cells.

Cysteine biosynthesis in archaea is still only beginning to be understood. *Methanosarcina barkeri* and *Methanosarcina thermophila* were shown to use the bacterial pathway [21, 22], because functional homologs of SAT and OASS could be identified in these organisms (table 1 [21, 22]). The presence of both homologs in *Methanosarcina acetivorans* and *Halobacterium* sp. NRC-1 (table 1) sug-

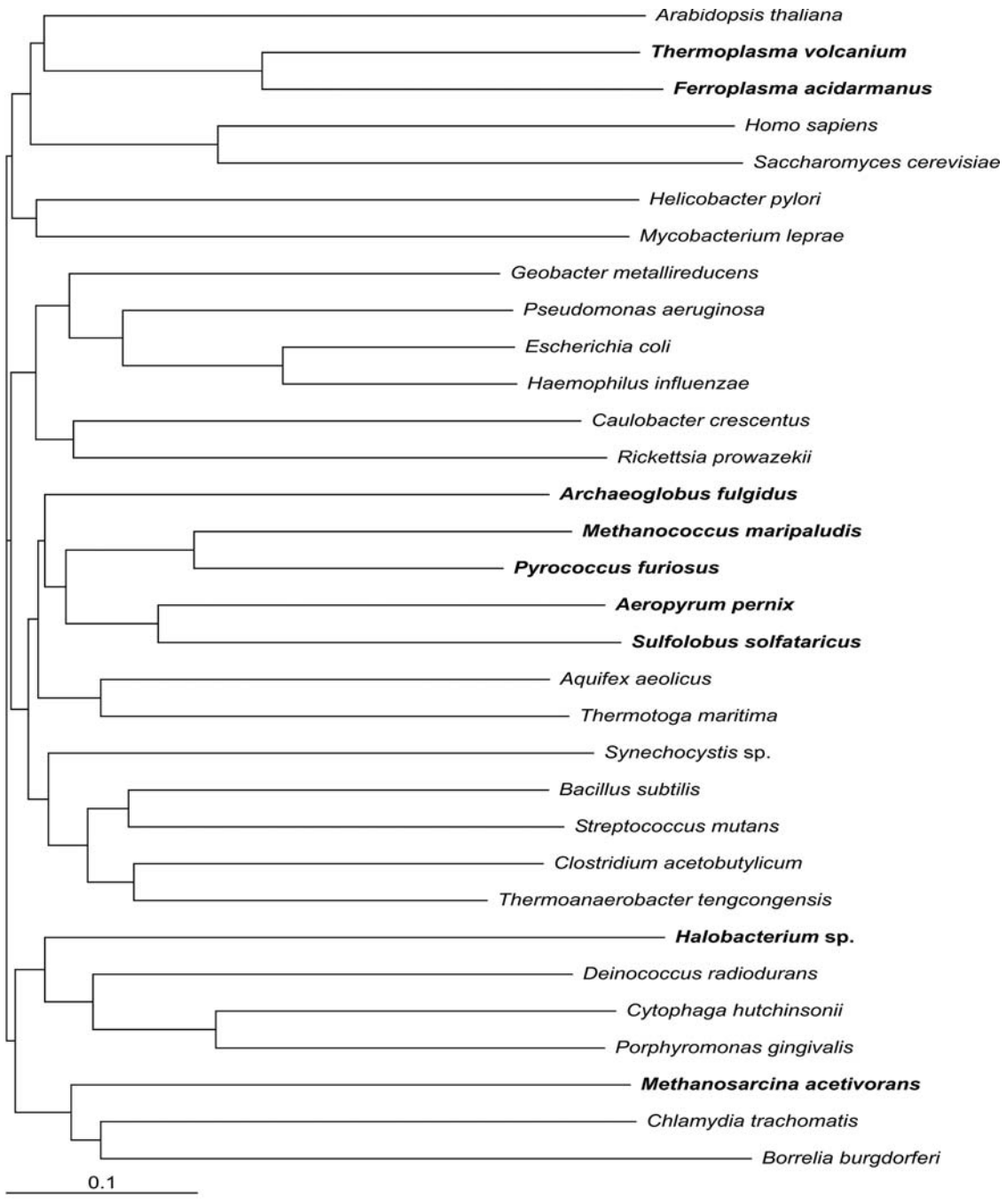

Figure 1. Consensus phylogenetic tree of CysRS proteins, inferred by the neighbor-joining method. CysRS sequences were obtained from available genomic databases and aligned using ClustalX [11]. Archaeal CysRSs are shown in bold. The bar represents 10 amino acid replacements per 100 positions.

gests a similar situation for these archaea. For all other archaea, cysteine biosynthesis remains obscure. Sequence homology searches throughout the completely sequenced genomes do not allow the nature of the cysteine biosynthesis pathway to be predicted. SAT is absent in most of the archaea (table 1), and Blast searches using *E. coli* OASS and mouse CBS as queries in most cases yield the same gene or set of genes (table 1). Functional analysis of one of the OASS and CBS homologs in *A. pernix* revealed that the protein had OASS activity as well as CBS activity in vitro [23]. It also showed that the protein preferred the more thermostable *O*-phosphoserine serine over *O*-acetylserine as substrate for the OASS activity [24]. Interestingly, *O*-phosphoserine is also the direct precursor for serine biosynthesis. It is possible that these archaea might use the same precursor for cysteine and serine

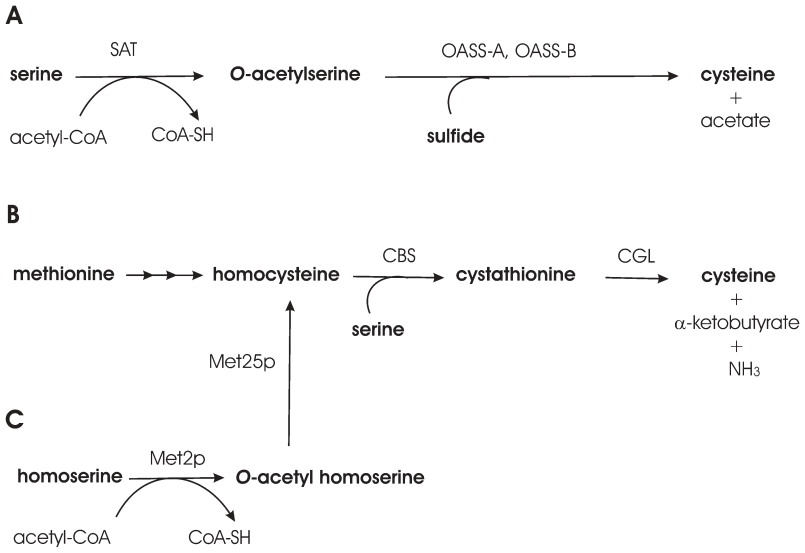

Figure 2. Biosynthesis of cysteine in bacteria (*A*), mammals (*B*) and *S. cerevisiae* (*C*). Abbreviations are CoA, coenzyme A; SAT, serine *O*-acetyltransferase; OASS, *O*-acetylserine sulfhydrylase; Met2p, homoserine transacetylase from *S. cerevisiae*; Met25p, homocysteine synthase from *S. cerevisiae*; CBS, cystathionine β-synthase; CGL, cystathionine γ-lyase.

Table 1. Presence of CysRS and cysteine biosynthesis proteins homologs in archaea.

Organism		Bacterial pathway	Animal pathway	
	CysRS	SAT	OASS CBS	CGL
Crenarcheaota				
Aeropyrum pernix	NP_148045	–	NP_148041* NP_147802	NP_147803
Sulfolobus solfataricus	NP_343652	–	NP_341900	NP_343729
Sulfolobus tokodaii	NP_378245	–	NP_377338	NP_376392
Pyrobaculum aerophilum	NP_558873	–	NP_559045	NP_559999
Euryarcheota				
Archaeoglobus fulgidus	NP_069247	–	–	–
Halobacterium sp NRC-1	NP_280014	NP_280304	NP_280167 NP_279635	NP_279780
Methanothermobacter thermautotrophicus	–	–	–	–
Methanocaldococcus jannaschii	–	–	–	–
Methanopyrus kandleri	–	–	–	–
Methanosarcina acetivorans	NP_615709	NP_617620	NP_617619	NP_617435
Methanosarcina barkeri	ZP_00077080	ZP_00077518*	ZP_00077517*	ZP_00757651
Methanosarcina mazei	NP_633935	NP_634963	–	NP_635109
Methanosarcina thermophila	Nd	AAG01805*	AAG01804*	Nd
Pyrococcus abyssi	NP_127080	–	NP_126065	NP_126585
Pyrococcus furiosus	NP_578753	–	NP_579587	NP_578995
Pyrococcus horikoshii	NP_142595	–	–	NP_142999
Ferroplasma acidarmanus	ZP_00000437	–	ZP_00000432 ZP_00001530	ZP_00001365
Thermoplasma acidophilum	NP_394604	–	NP_394010	NP_393559
Thermoplasma volcanium	NP_111763	–	NP_111108	NP_110693
Nanoarchaeota				
Nanoarchaeum equitans	NP_963349	–	–	–

Presence (NCBI accession number) or absence (–) of orthologs were determined using BLASTp with *Mus musculus* CBS and CGL as query or with *E. coli* SAT and OASS as query. * Enzymes experimentally isolated and characterized [21–23]. Nd, not determined because the genome sequence is not available. The shaded area corresponds to the three archaea missing a canonical class I CysRS.

biosynthesis, and therefore not require any SAT-like activity. On the other hand, in some archaea, labeling experiments so far are consistent with the presence of the mammalian-like transsulfuration pathway (fig. 2B). Zhou and White [25] used [^{34}S]-labeling to demonstrate that labeled methionine could provide the sulfur for both cysteine and methionine in *Haloferax marismortui* and *Sulfolobus acidocaldarius*, representing both major archaeal kingdoms. Here again sequence homology searches do not permit clarification of the situation. Homologs of the relevant enzymes, CBS and CGL, are widespread in both archaeal kingdoms (table 1). However, the CGL homologs are annotated without distinctions as CGL or as cystathionine γ-synthase (CGS), a closely related protein involved in cystathionine synthesis via coupling of cysteine to *O*-succinylhomoserine. In the absence of any functional analysis of these CGL homologs, it is therefore impossible to say whether they are involved in cysteine biosynthesis in archaea.

Finally, homologs of the bacterial SAT, OASS-A and OASS-B have not been identified in the archaeal genomes of *Archaeoglobus fulgidus*, *M. thermautotrophicus*, *M. jannaschii*, *M. kandleri* and *Pyrococcus horikoshii* (table 1). In these archaea, the only open reading frame to which SAT and OASS display significant similarity is the evolutionarily related PLP-dependent amino acid synthase, threonine synthase [26]. No homologs of CBS or CGL that could indicate the existence of a mammalian-like pathway are found in *M. jannaschii*, *M. kandleri* or *M. thermautotrophicus* (table 1) either. None of these archaea require cysteine for growth [27–29] and so must be synthesizing this essential amino acid by some route. Recently, possible intermediates in cysteine biosynthesis were identified in cell extracts of *M. jannaschii* using coupled gas chromatography and mass spectrometry [30]. Measurement of these intermediates was dependent upon the added substrates, and so while an enzymatic potential was determined, the natural presence of such intermediates is still not apparent. In this study, cysteine formation appeared to be dependent on the presence of homocysteine plus another substrate (*O*-phosphoserine, serine or *O*-acetylserine) or on cystathionine alone. This dependence on homocysteine or cystathionine suggests the mammalian pathway (fig. 2B). However, the production of these two intermediates in the *M. jannaschii* extracts indicated variations with the mammalian pathway. Cystathionine was not produced from homocysteine plus serine (fig. 2B) but from homocysteine plus *O*-phosphoserine. In this ability to use *O*-phosphoserine, these thermophilic archaea may be using an intermediate of serine biosynthesis as was suggested for the bacterial-like pathway in *A. pernix* [24]. Similarly, homocysteine production was measured upon addition of *O*-phosphohomoserine or *O*-acetylhomoserine, and was dependent on an endogenous, unidentified sulfur source that was not sulfide [30]. Whatever the source of the homocysteine and cystathionine, these archaea are still missing the homologs of CBS and CGL of the mammalian pathway (table 1) [6–8, 31]. The lack of a CBS homolog may be due in part to the observation that cystathionine formation was dependent on *O*-phosphoserine rather than serine [30]. Finally, perhaps this study demonstrates that methanogens use a novel enzyme to perform the cystathionine cleavage step. If cysteine is indeed produced in *M. jannaschii* by this proposed route, at this point in time, the identities of the enzymes catalyzing these steps are still unknown.

Investigation of Cys-tRNA^{Cys} formation in three methanogenic archaea

Codon frequencies reveal the presence of a normal level of cysteine in proteins of the methanogenic archaea. Functionally essential cysteine residues are conserved in a number of *M. jannaschii* proteins [32]. An analysis of the charging level of tRNA in freshly prepared *M. jannaschii* cell extracts indicates that tRNA^{Cys} is aminoacylated in vivo [C. Polycarpo, unpublished results]. This implies in vivo formation of Cys-tRNA^{Cys} and use of this AA-tRNA as substrate for ribosomal protein synthesis. While this, of course, was not unexpected, it raises the question how do these archaea form Cys-tRNA^{Cys}? Based on our knowledge of AA-tRNA formation, two routes for Cys-tRNA^{Cys} formation may be envisioned. (i) Direct acylation of tRNA to form Cys-tRNA^{Cys} is possible via a CysRS not recognized by bioinformatics means. Direct synthesis requires the presence of free cysteine in these archaea. Therefore, in addition to a novel CysRS, a cysteine biosynthetic pathway would also need to be identified. (ii) An indirect route, formally analogous to that of selenocysteine synthesis, is also a possibility. For this to happen, a cysteine precursor (e.g. serine) is first misacylated onto tRNA^{Cys} and subsequently modified to form cysteine. Such an indirect route could also provide cysteine for protein synthesis and possibly other metabolic pathways.

An indirect pathway?

Nature provides a number of cases for which the synthesis of a given amino acid relies on a tRNA-dependent two-step mechanism (reviewed in [5]). This biosynthetic route invokes the aminoacylation of a specific tRNA with a chemical precursor and subsequent modification of this precursor to yield the new amino acid still attached to the tRNA. A typical example of such a tRNA-dependent two-step mechanism is the biosynthesis of selenocysteine [33, 34] and formylmethione [35, 36]. Cysteine and methion-

ine are first attached to a specific selenocysteine tRNA (tRNA^{Sec}) and a specific initiator formylmethionine tRNA (tRNA^{fMet}) by action of seryl-tRNA synthetase (SerRS) and methionyl-tRNA synthetase, respectively. Upon action of selenocysteine synthase the Ser-tRNA^{Sec} is converted into a selenocysteinyl-tRNA^{Sec} available for protein synthesis [33, 34]. Similarly, formylation of the α-amino group of Met-tRNA^{fMet} leads to formylMet-tRNA^{fMet}, necessary for translation initiation in bacteria and organelles [35, 36]. In the light of the fact that the formation of several AA-tRNAs involves an indirect route, we have searched for a similar scenario in which serine is first aminoacylated onto tRNA^{Cys} by a mischarging SerRS to form Ser-tRNA^{Cys}. The hydroxyl moiety of the aminoacylated serine would be then replaced by a sulfhydryl group to yield Cys-tRNA^{Cys}. The presence of a phylogenetically divergent group of SerRSs in a number of methanogenic archaea, including *M. jannaschii*, *M. thermautotrophicus* and *M. kandleri*, seemed to correlate with this working hypothesis [37]. However, in vitro attempts to mischarge serine onto native or transcript tRNA^{Cys} species using purified, recombinant SerRSs from *M. thermautotrophicus*, *M. jannaschii* or *M. kandleri* never yielded any Ser-tRNA^{Cys} ([37]; D. Korencic, unpublished results), whereas the same enzymes correctly charged their cognate tRNA^{Ser} species. Thus, synthesis of Cys-tRNA^{Cys} via the indirect formation of Ser-tRNA^{Cys} seems unlikely. However, the above studies were performed with pure SerRS preparations and did not investigate the requirement of an additional protein cofactor allowing Ser-tRNA^{Cys} formation, nor the existence of an indirect pathway involving mischarging of any cysteine analogs besides serine.

Or a direct pathway?

Direct aminoacylation of tRNA^{Cys} with cysteine by a non-canonical AARS would be a possible alternative to an indirect route. An attractive precedent is the replacement of the canonical class II LysRS by a class I LysRS in several archaea and bacteria [4]. However, the sequencing of the *M. jannaschii* genome did not reveal any undesignated AARS-like genes potentially responsible for cysteinylation of tRNA^{Cys}.

One remarkable possibility was that an already specified AARS could provide the direct cysteine-charging activity. Classical biochemical purification of an enzymatic activity responsible for the direct cysteinylation of *M. jannaschii* unfractionated tRNA led to the isolation of prolyl-tRNA synthetase (ProRS) as the enzyme responsible for this activity [38, 39]. Recombinant *M. jannaschii* ProRS expressed in *E. coli* confirmed this unexpected finding. Determination of kinetic parameters showed that the overall activation and transfer onto tRNA of proline by *M. jannaschii* ProRS was ~5–100 times, depending on the experimental conditions, more efficient than that of cysteine [40–42].

In addition, a temperature-sensitive *E. coli cysS* (encoding CysRS) allele in was apparently rescued for growth at the non-permissive temperature by *proS* genes (encoding ProRS) from *M. jannaschii*, *M. thermautotrophicus* or *M. maripaludis*, suggesting that the archaeal ProRS was indeed able to functionally replace the *E. coli* CysRS [38]. However, analysis of the aminoacylation specificity of this highly unusual *M. jannaschii* ProRS later showed that, in vitro, cysteine was in fact transferred onto mature *M. jannaschii* tRNA^{Pro} and not onto tRNA^{Cys} [43]. In fact, ProRSs from organisms spanning the three domains of life were all able to mischarge cysteine unto tRNA^{Pro} to form Cys-tRNA^{Pro} albeit in variable proportions when compared to the formation of the cognate product Pro-tRNA^{Pro} [44].

Misacylation of cysteine onto tRNA^{Pro} did not correlate with the presence or absence of an editing domain, shown to be able to hydrolyze the incorrectly formed Ala-tRNA^{Pro} in bacterial type ProRSs [45]. Unlike alanine, cysteine is able to escape the quality control mechanism of these enzymes, as none of the ProRSs were able to hydrolyze Cys-tRNA^{Pro} [44]. Also, formation of the prolyl-adenylate by *Thermus thermophilus* ProRS was shown to induce major active site conformational rearrangements likely leading to the formation of a productive complex with tRNA^{Pro} and eventually to Pro-tRNA^{Pro} [46]. A specific prolyl-AMP-induced ordering of the active site was therefore proposed to be a possible mechanism selected by archaeal-like ProRSs to ensure higher aminoacylation specificity [46]. However, co-crystallization of *M. thermautotrophicus* ProRS with the aminoacyl-adenylate analogs cysteinyl- and prolyl-sulfamoyl-adenylate (CysAMS and ProAMS) indicated that this might not be the case [47]. Indeed, the analogs of the activated proline and cysteine were shown to induce identical conformational changes around the active site (fig. 3), leading to the same proposed productive conformation as the one observed for the *T. thermophilus* enzyme. The *M. thermautotrophicus* ProRS crystal structure showed that ProAMS and CysAMS occupy a very similar position in the active site, and that once bound to the active site cavity, the aminoacyl moieties of the adenylates occupies the same molecular volume [47]. Thus, in the absence of a specific amino acid discrimination mechanism, such as the one present in the CysRS active site, ProRS has difficulty discriminating, at least in vitro, between proline and cysteine both in the editing and activation sites.

Revisiting the in vivo complementation approaches suggested that the archaeal ProRSs were in fact not producing Cys-tRNA^{Cys}. Complete characterization of the *E. coli* temperature-sensitive *cysS* strain led to the identifi-

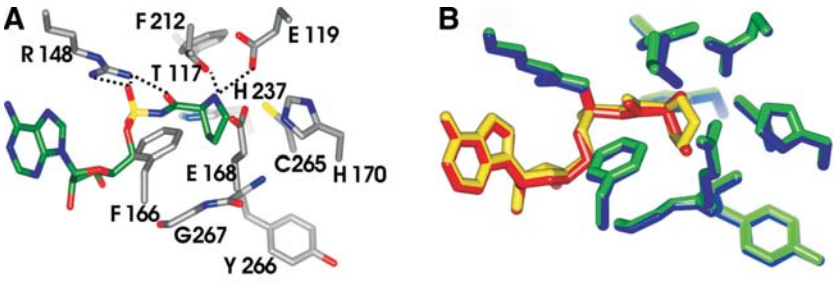

Figure 3. (*A*) Prolyl-sulfamoyl-adenylate bound to the active site of *M. thermautotrophicus* ProRS. All residues within 4 Å of the amino acid moiety of the sulfamoyl-adenylate are shown. Hydrogen bounds are indicated as black dotted lines. (*B*) Superimposition of the active site structure of *M. thermautotrophicus* ProRS with prolyl-sulfamoyl-adenylate (ligand in yellow and protein in green) and cysteinyl-sulfamoyl-adenylate (ligand in red and protein in dark blue). Figures are from [47].

cation of the precise mutation in the *cysS* gene that was responsible for the thermosensitivity of the strain [48]. The mutation, V27E, was found in the active site at a position near a residue responsible for cysteine recognition via chelation to a zinc ion [9, 48]. The mutation provoked a fourfold increase in K_M for cysteine, whereas the affinity for ATP and $tRNA^{Cys}$ remained similar to the one determined for the wild-type enzyme [48]. Consistent with this observation, an increase in the cysteine concentration in the growth medium restored the growth of the *E. coli* $cysS^{ts}$ strain at the non-permissive temperature. This result may explain the weak complementation observed previously using the archaeal-type *proS* genes [38, 49]. Therefore, the construction of a *cysS* deletion strain of *E. coli* provided a more stringent complementation test. The *proS* genes from *M. maripaludis* or *M. jannaschii* could not restore growth of this deletion strain, in contrast to the *cysS* genes from *E. coli* or *M. maripaludis* [48], confirming that *proS* is not functionally equivalent to *cysS*.

In conclusion, despite a remarkable ability to activate cysteine, ProRS is unable to transfer cysteine onto $tRNA^{Cys}$ either in vitro or in vivo in *E. coli*. ProRS remains a highly unusual enzyme as it is the only synthetase known to date that recognizes equally well a cognate and non-cognate canonical amino acid in vitro. One must assume that such mischarging does not occur in vivo or at least at a level unbearable to the cell. Also, the incidence of cysteine mischarging by ProRS would be minimized if cysteine was not as readily available to the enzyme. This situation would be the case in the presence of (i) a favorable intracellular ratio of proline to cysteine, (ii) a CysRS that could out-compete ProRS for cysteine or (iii) a tRNA-dependent cysteine biosynthetic route. Another option would be the existence of a specificity cofactor that would either prevent cysteine mischarging or conversely favor the production of correctly charged Cys-$tRNA^{Cys}$. Finally, if high mischarging were to occur in *M. jannaschii* cells, a physiological role for Cys-$tRNA^{Pro}$ might exist.

Another report [50] introduced an unknown *M. jannaschii* ORF MJ1477 as the missing CysRS. However, further bioinformatic [51] and genetic [48] studies did not support this assignment.

What protein(s) forms Cys-$tRNA^{Cys}$ in *M. jannaschii*?

Despite all the efforts to answer this question, the proteins that form cysteine and Cys-$tRNA^{Cys}$ in *M. jannaschii*, *M. thermautotrophicus* and *M. kandleri* are still unknown. It remains to be seen whether Cys-$tRNA^{Cys}$ is synthesized via an indirect pathway while simultaneously providing a biosynthetic pathway for cysteine or is formed via direct acylation of $tRNA^{Cys}$ with free cysteine by an unidentified protein. Involvement of the methanogenic SerRS, ProRS or MJ1477, in the context of the cell, cannot be definitively ruled out but seems doubtful in the absence of supporting biochemical evidence. Recent advances in understanding the molecular processes underlying fixation of sulfur in both bacteria and archaea could provide some insights into cysteine and Cys-$tRNA^{Cys}$ formation in *M. jannaschii*. Discovery of an enzyme in *M. jannaschii* able to transfer sulfide onto phosphoenolpyruvate, an analog of dehydroalanine, could be of interest [52]. Indeed, dehydroalanine is the enzymatic intermediate in the reaction catalyzed by the second enzyme, OASS, of the bacterial cysteine biosynthetic route. The presence of an enzyme formally able to catalyze a similar reaction in *M. jannaschii* could hint at new cysteine or/and cysteinyl-tRNA biosynthetic pathways. However, sequence homology searches in the genome of *M. jannaschii*, *M. thermautotrophicus* and *M. kandleri* using BLASTP [53] have not revealed candidate paralogs. More intriguing is the presence of a homoserine transacetylase in *M. thermautotrophicus*. As homoserine transacetylase is the first enzyme of the cysteine biosynthesis pathway in *S. cerevisiae* (fig. 2C), a similar mechanism could possibly exist in *M. thermautotrophicus* [54]. Here again the remaining enzymes of the pathway are missing in *M. thermautotrophicus*, and also no orthologs of the homoserine transacetylase can be recognized in the genomes of *M.*

jannaschii or *M. kandleri*. A broader look at the amino acid metabolism in *M. jannaschii, M. thermoautrophicus* and *M. kandleri* reveals the absence not only of CysRS and a cysteine biosynthetic route but also enzymes of the methionine biosynthetic pathway (encoded in bacteria by *metA*, *metB*, *metC*, *metE* and *metH*). Because a MetRS can be identified, the aminoacyl-tRNA synthesis and amino acid biosynthesis involving methionine are uncoupled. The genes for threonine biosynthesis (*thrB* and *thrC*) are, however, present. Threonine biosynthesis and methionine biosyntheses share a metabolic intermediate, homoserine, also a potential intermediate of cysteine biosynthesis (fig. 2C). And as proposed in a recent study [30], homoserine, upon phosphorylation in the first step of the threonine biosynthetic route, could equally serve as starting point for methionine and cysteine biosynthesis and possibly for Cys-tRNA^{Cys} formation.

In conclusion, identification of a CysRS in organisms living in such drastic environments (high pressure, temperature, salinity, and the absence of light and oxygen) has proven to be more complicated than initially envisioned and is bound to produce more surprises. Despite earlier optimism, the formation of Cys-tRNA^{Cys} still remains the last puzzle in aminoacyl-tRNA synthesis [55].

Acknowledgments. We would like to thank D. E. Graham, C. Polycarpo and J. C. Salazar for critical reading of the manuscript and C. Polycarpo and D. Korencic for communication of unpublished results. We thank Tom Steitz for his encouragement in this work. The work in the author's laboratory was supported by grants from the National Institute of General Medical Sciences (D.S. and T.S.) and the Department of Energy (D.S.).

1 Ibba M. and Söll D. (2000) Aminoacyl-tRNA synthesis. Annu. Rev. Biochem. **69:** 617–650
2 Eriani G., Delarue M., Poch O., Gangloff J. and Moras D. (1990) Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. Nature **347:** 203–206
3 Cusack S., Berthet-Colominas C., Härtlein M., Nassar N. and Leberman R. (1990) A second class of synthetase structure revealed by X-ray analysis of *Escherichia coli* seryl-tRNA synthetase at 2.5 A. Nature **347:** 249–255
4 Ibba M., Morgan S., Curnow A. W., Pridmore D. R., Vothknecht U. C., Gardner W. et al. (1997) A euryarchaeal lysyl-tRNA synthetase: resemblance to class I synthetases. Science **278:** 1119–1122
5 Ibba M. and Söll D. (2004) Aminoacyl-tRNAs: setting the limits of the genetic code. Genes Dev. **18:** 731–738
6 Bult C. J., White O., Olsen G. J., Zhou L., Fleischmann R. D., Sutton G. G. et al. (1996) Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. Science **273:** 1058–1073
7 Smith D. R., Doucette-Stamm L. A., Deloughery C., Lee H., Dubois J., Aldredge T. et al. (1997) Complete genome sequence of *Methanobacterium thermoautotrophicum* ΔH: functional analysis and comparative genomics. J. Bacteriol. **179:** 7135–7155
8 Slesarev A. I., Mezhevaya K. V., Makarova K. S., Polushin N. N., Shcherbinina O. V., Shakhova V. V. et al. (2002) The complete genome of hyperthermophile *Methanopyrus kandleri* AV19 and monophyly of archaeal methanogens. Proc. Natl. Acad. Sci. USA **99:** 4644–4649
9 Newberry K. J., Hou Y. M. and Perona J. J. (2002) Structural origins of amino acid selection without editing by cysteinyl-tRNA synthetase. EMBO J. **21:** 2778–2787
10 Fersht A. R. and Dingwall C. (1979) Cysteinyl-tRNA synthetase from *Escherichia coli* does not need an editing mechanism to reject serine and alanine: high binding energy of small groups in specific molecular interactions. Biochemistry **18:** 1245–1249
11 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
12 Stathopoulos C., Kim W., Li T., Anderson I., Deutsch B., Palioura S. et al. (2001) Cysteinyl-tRNA synthetase is not essential for viability of the archaeon *Methanococcus maripaludis*. Proc. Natl. Acad. Sci. USA **98:** 14292–14297
13 Brooks D. J. and Fresco J. R. (2002) Increased frequency of cysteine, tyrosine and phenylalanine residues since the last universal ancestor. Mol. Cell. Proteomics **1:** 125–131
14 Brooks D. J., Fresco J. R., Lesk A. M. and Singh M. (2002) Evolution of amino acid frequencies in proteins over deep time: inferred order of introduction of amino acids into the genetic code. Mol. Biol. Evol. **19:** 1645–1655
15 Wong J. T. (1975) A co-evolution theory of the genetic code. Proc. Natl. Acad. Sci. USA **72:** 1909–1912
16 Trifonov E. N. (2000) Consensus temporal order of amino acids and evolution of the triplet code. Gene **261:** 139–151
17 Kredich N. M. (1996) Biosynthesis of cysteine. In: *Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology, pp. 514–527, Neidhardt F. C., Curtiss R., Ingraham J. L., Lin E. C. C., Low K. B., Magasanik B. et al. (eds), American Society for Microbiology, Washington, DC
18 Noji M. and Saito K. (2002) Molecular and biochemical analysis of serine acetyltransferase and cysteine synthase towards sulfur metabolic engineering in plants. Amino Acids **22:** 231–243
19 Griffith O. W. (1987) Mammalian sulfur amino acid metabolism: an overview. Methods Enzymol. **143:** 366–376
20 Thomas D. and Surdin-Kerjan Y. (1997) Metabolism of sulfur amino acids in *Saccharomyces cerevisiae*. Microbiol. Mol. Biol. Rev. **61:** 503–532
21 Kitabatake M., So M. W., Tumbula D. L. and Söll D. (2000) Cysteine biosynthesis pathway in the archaeon *Methanosarcina barkeri* encoded by acquired bacterial genes? J. Bacteriol. **182:** 143–145
22 Borup B. and Ferry J. G. (2000) Cysteine biosynthesis in the Archaea: *Methanosarcina thermophila* utilizes O-acetylserine sulfhydrylase. FEMS Microbiol. Lett. **189:** 205–210
23 Mino K. and Ishikawa K. (2003) Characterization of a novel thermostable O-acetylserine sulfhydrylase from *Aeropyrum pernix* K1. J. Bacteriol. **185:** 2277–2284
24 Mino K. and Ishikawa K. (2003) A novel *O*-phospho-L-serine sulfhydrylation reaction catalyzed by *O*-acetylserine sulfhydrylase from *Aeropyrum pernix* K1. FEBS Lett. **551:** 133–138
25 Zhou D. and White R. H. (1991) Transsulfuration in archaebacteria. J. Bacteriol. **173:** 3250–3251
26 Alexander F. W., Sandmeier E., Mehta P. K. and Christen P. (1994) Evolutionary relationships among pyridoxal-5′-phosphate-dependent enzymes: regio-specific α, β and γ families. Eur. J. Biochem. **219:** 953–960
27 Whitman W. B., Bowen T. L. and Boone D. R. (1991) The methanogenic bacteria. In: The Prokaryotes, pp. 719–767, Balows A., Trüper H. G., Dworkin M., Harder W. and Schleifer K. H. (eds), Springer, New York
28 Stetter K. O. (1991) The genus *Archaeoglobus*. In: The Prokaryotes, pp. 707–711, Balows A., Trüper H. G., Dworkin M., Harder W. and Schleifer K. H. (eds), Springer, New York

29 Kurr M., Huber R., König H., Jannasch H. W., Fricke H., Trincone A. et al. (1991) *Methanopyrus kandleri*, gen. and sp. nov. respresents a novel group of hyperthermophilic methanogens, growing at 100 °C. Arch. Microbiol. **156:** 239–247

30 White R. H. (2003) The biosynthesis of cysteine and homocysteine in *Methanococcus jannaschii*. Biochim. Biophys. Acta **1624:** 46–53

31 Klenk H. P., Clayton R. A., Tomb J. F., White O., Nelson K. E., Ketchum K. A. et al. (1997) The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. Nature **390:** 364–370

32 Lee D. Y., Ahn B. Y. and Kim K. S. (2000) A thioredoxin from the hyperthermophilic archaeon *Methanococcus jannaschii* has a glutaredoxin-like fold but thioredoxin-like activities. Biochemistry **39:** 6652–6659

33 Böck A., Thanbichler M., Rother M. and Resch A. (2003) Selenocysteine. In: Aminoacyl-tRNA Synthetases, Ibba M., Francklyn C. and Cusack S. (eds), Landes Bioscience Publishing, in press

34 Lescure A., Fagegaltier D., Carbon P. and Krol A. (2002) Protein factors mediating selenoprotein synthesis. Curr. Protein Pept. Sci. **3:** 143–151

35 Schmitt E., Panvert M., Blanquet S. and Mechulam Y. (1998) Crystal structure of methionyl-tRNAfMet transformylase complexed with the initiator formyl-methionyl-tRNAfMet. EMBO J. **17:** 6819–6826

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

37 Kim H. S., Vothknecht U. C., Hedderich R., Celic I. and Söll D. (1998) Sequence divergence of seryl-tRNA synthetases in archaea. J. Bacteriol. **180:** 6446–6449

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

39 Lipman R. S., Sowers K. R. and Hou Y. M. (2000) Synthesis of cysteinyl-tRNA^{Cys} by a genome that lacks the normal cysteine-tRNA synthetase. Biochemistry **39:** 7792–7798

40 Lipman R. S., Beuning P. J., Musier-Forsyth K. and Hou Y. M. (2002) Amino acid activation of a dual-specificity tRNA synthetase is independent of tRNA. J. Mol. Biol. **316:** 421–427

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

42 Jacquin-Becker C., Ahel I., Ambrogelly A., Ruan B., Söll D. and Stathopoulos C. (2002) Cysteinyl-tRNA formation and prolyl-tRNA synthetase. FEBS Lett. **514:** 34–36

43 Ambrogelly A., Ahel I., Polycarpo C., Bunjun-Srihari S., Krett B., Jacquin-Becker C. et al. (2002) *Methanocaldococcus jannaschii* prolyl-tRNA synthetase charges tRNA^{Pro} with cysteine. J. Biol. Chem. **277:** 34749–34754

44 Ahel I., Stathopoulos C., Ambrogelly A., Sauerwald A., Toogood H., Hartsch T. et al. (2002) Cysteine activation is an inherent *in vitro* property of prolyl-tRNA synthetases. J. Biol. Chem. **277:** 34743–34748

45 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

46 Yaremchuk A., Tukalo M., Grotli M. and Cusack S. (2001) A succession of substrate induced conformational changes ensures the amino acid specificity of *Thermus thermophilus* prolyl-tRNA synthetase: comparison with histidyl-tRNA synthetase. J. Mol. Biol. **309:** 989–1002

47 Kamtekar S., Kennedy W. D., Wang J., Stathopoulos C., Söll D. and Steitz T. A. (2003) The structural basis of cysteine aminoacylation of tRNA^{Pro} by prolyl-tRNA synthetases. Proc. Natl. Acad. Sci. USA **100:** 1673–1678

48 Ruan B., Nakano H., Tanaka M., Mills J. A., DeVito J. A., Min B. et al. (2004) Cysteinyl-tRNA^{Cys} formation in *Methanocaldococcus jannaschii*: the mechanism is still unknown. J. Bacteriol. **186:** 8–14

49 Bunjun S., Stathopoulos C., Graham D., Min B., Kitabatake M., Wang A. L. et al. (2000) A dual-specificity aminoacyl-tRNA synthetase in the deep-rooted eukaryote *Giardia lamblia.* Proc. Natl. Acad. Sci. USA **97:** 12997–13002

50 Fàbrega C., Farrow M. A., Mukhopadhyay B., de Crécy-Lagard V., Ortiz A. R. and Schimmel P. (2001) An aminoacyl tRNA synthetase whose sequence fits into neither of the two known classes. Nature **411:** 110–114

51 Iyer L. M., Aravind L., Bork P., Hofmann K., Mushegian A. R., Zhulin I. B. et al. (2001) *Quod erat demonstrandum?* The mystery of experimental validation of apparently erroneous computational analyses of protein sequences. Genome Biol. **2:** RESEARCH0051

52 Graham D. E., Xu H. and White R. H. (2002) Identification of coenzyme M biosynthetic phosphosulfolactate synthase: a new family of sulfonate-biosynthesizing enzymes. J. Biol. Chem. **277:** 13421–13429

53 Altschul S. F., Gish W., Miller W., Myers E. W. and Lipman D. J. (1990) Basic local alignment search tool. J. Mol. Biol. **215:** 403–410

54 Borup B. and Ferry J. G. (2000) *O*-Acetylserine sulfhydrylase from *Methanosarcina thermophila*. J. Bacteriol. **182:** 45–50

55 Li T., Graham D. E., Stathopoulos C., Haney P. J., Kim H. S., Vothknecht U. et al. (1999) Cysteinyl-tRNA formation: the last puzzle of aminoacyl-tRNA synthesis. FEBS Lett. **462:** 302–306