FEBS Letters 514 (2002) 34–36 FEBS 25819

Minireview

Cysteinyl-tRNA formation and prolyl-tRNA synthetase

Clarisse Jacquin-Becker, Ivan Ahel, Alexandre Ambrogelly, Benfang Ruan, Dieter Söll*, Constantinos Stathopoulos

Department of Molecular Biophysics and Biochemistry, Yale University, P.O. Box 208114, 266 Whitney Avenue, New Haven, CT 06520-8114, USA

Received 25 November 2001; revised 10 December 2001; accepted 11 December 2001

First published online 31 January 2002

Edited by Lev Kisselev

Abstract Aminoacyl-tRNA (AA-tRNA) formation is a key step in protein biosynthesis. This reaction is catalyzed with remarkable accuracy by the AA-tRNA synthetases, a family of 20 evolutionarily conserved enzymes. The lack of cysteinyl-tRNA (Cys-tRNA) synthetase in some archaea gave rise to the discovery of the archaeal prolyl-tRNA (Pro-tRNA) synthetase, an enzyme capable of synthesizing Pro-tRNA and Cys-tRNA. Here we review our current knowledge of this fascinating process.

Key words: Aminoacyl-tRNA synthetase; Prolyl-tRNA synthetase; Amino acid activation; tRNA; Translation

1. Introduction

Aminoacyl-tRNA (AA-tRNA) synthesis is an essential biochemical process in every cell that provides the appropriate amino acid substrates for ribosomal translation. Aminoacylation is characterized by the high specificity of the covalent attachment of a particular amino acid to its cognate tRNA. The

$$\mathrm{AA} + \mathrm{ATP} + \mathrm{AARS} \rightleftharpoons \mathrm{AARS{\bullet}AA\text{-}AMP} + \mathrm{PP_i} \quad (1)$$

$$\mathrm{AARS{\bullet}AA\text{-}AMP} + \mathrm{tRNA} \rightleftharpoons \mathrm{AA\text{-}tRNA} + \mathrm{AMP} \quad (2)$$

two-step reaction begins with ATP-dependent amino acid activation (Eq. 1). In the next step the activated amino acid is transferred to the cognate tRNA (Eq. 2). The exquisite specificity in matching the amino acid with the tRNA's anticodon by the AA-tRNA synthetases (AARSs) is a key feature in the fidelity of expressing the genetic information. There are 20 members of this family of enzymes, as first proposed in Crick's adaptor hypothesis [1]. These enzymes are divided into two unrelated structural classes based on the topology of their ATP binding site (reviewed in [2]). These unique structural features make AARS genes also easily detectable by sequence similarity searches of known genomes from many organisms. In this way it became apparent that there are a number of deviations from the notion that all organisms contain 20 canonical AARSs (summarized in [3,4]). Here we review one of these unusual routes of AA-tRNA synthesis, the formation of cysteinyl-tRNA (Cys-tRNA) in archaea lacking *cysS*, the gene encoding cysteinyl-tRNA synthetase (CysRS).

2. Archaeal prolyl-tRNA synthetase (ProRS) can form Cys-tRNA

The fact that Cys-tRNA formation sometimes requires non-canonical enzymes is apparent from the knowledge that the canonical *cysS* gene is lacking in the complete genomes of three thermophilic or hyperthermophilic methanogenic archaea, *Methanocaldococcus* (*Methanococcus*) *jannaschii* [5], *Methanothermobacter thermoautotrophicus* [6] and *Methanopyrus kandleri* [7]. However, preliminary biochemical studies suggested that a Cys-tRNA forming activity was indeed present in cell-free extracts of these organisms. In addition, the formation of Cys-tRNA by modification of Ser-tRNA (analogous to selenocysteinyl-tRNA formation [8]) could not be demonstrated by in vitro experiments with pure *Methanothermobacter marburgensis* SerRS [9]. A biochemical purification procedure (based on charging of *M. jannaschii* tRNA with cysteine) then led to the isolation from *M. jannaschii* cell extracts of a protein whose N-terminal amino acid sequence matched that of ProRS [10,11]. Subsequent cloning and overexpression of the archaeal *proS* gene in *Escherichia coli* followed by purification of the recombinant protein showed that *M. jannaschii* ProRS could indeed form Cys-tRNA (with about 5-fold [12] or 36-fold [11] lower efficiency than synthesis of prolyl-tRNA (Pro-tRNA)). Cys-tRNA formation could be inhibited by proline analogs, again attesting the fact that ProRS catalyzed the cysteinylation reaction. To demonstrate the nature of the charged tRNA, periodate inactivation of the uncharged tRNAs present in a preparation of total *M. jannaschii* Cys-tRNA (acylated by ProRS) was used [10,11]. After deacylating the sample, the resulting tRNA could be charged by the archaeal ProRS only with cysteine. Moreover, the Cys-tRNA forming activity of ProRS was demonstrated in vivo in an *E. coli* strain carrying a *cysS*ts mutation. The presence of the *M. jannaschii proS* gene could rescue – albeit poorly – the growth of the mutant strain at the non-permissive temperature. Thus, ProRS may be the enzyme that provides Cys-tRNA in vivo in organisms that lack *cysS*.

Phylogenetic comparative analyses of ProRSs from many organisms clearly show that they cluster in two subgroups

*Corresponding author. Fax: (1)-203-432 6202.
E-mail address: soll@trna.chem.yale.edu (D. Söll).

Abbreviations: AARS, aminoacyl-tRNA synthetase; Cys-tRNA, cysteinyl-tRNA; Pro-tRNA, prolyl-tRNA; CysRS, cysteinyl-tRNA synthetase; ProRS, prolyl-tRNA synthetase; AA-tRNA, aminoacyl-tRNA

0014-5793/02/$22.00
PII: S0014-5793(02)02331-1

[13,14]; one bacterial-like and one archaeal-like (including all eukaryotes and some bacteria). The major difference between the two subgroups is an insertion domain (~180 amino acids) between motifs II and III in bacterial-like ProRSs. The archaeal genre enzyme has a variable length C-terminal extension instead. The wider distribution of the dual-specificity enzymes was verified with the biochemical detection of an archaeal type ProRS in the deep-rooted eukaryote *Giardia lamblia* [15] and in the *Thermus/Deinococcus* group (e.g. in *Thermus thermophilus* [16]). In both cases lateral gene transfer played an important role in the acquisition of the archaeal type ProRS [13,17] and both enzymes have been shown to recognize cysteine as a substrate [15,16]. There is even a structural argument for possible cysteine binding by ProRS. Modeling of a recent crystal structure of *T. thermophilus* ProRS suggests that a discrete cysteine binding pocket might be created from the proline binding site before the tRNA's acceptor helix attaches; this attachment may subsequently prevent a loop movement required for proline binding [18]. If this could be experimentally verified it may explain the property that many ProRS enzymes can activate cysteine (see below).

3. Mechanistic and biochemical aspects of cysteine activation by ProRS

The major challenge now is to decipher the mechanism by which the dual-specificity ProRS discriminates between its two 'cognate' amino acids proline and cysteine. Preliminary biochemical amino acid analog inhibition data suggesting overlapping amino acid binding sites on *M. jannaschii* ProRS were supported by the analysis of a number of mutant ProRS enzymes. A conserved region close to motif I appears to include amino acid residues important for binding of either proline or cysteine. Pro100Ala reduces dramatically CysRS activity while Glu103Ala abolishes ProRS activity [12]. This shows that certain structural elements of the amino acid binding pocket are in close proximity in the primary sequence. The crystal structure of *T. thermophilus* ProRS showed that this conserved region (PTXE) is indeed part of the active site [18]. The data further suggest that the formation of prolyl-adenylate, the activated amino acid intermediate (see Eq. 1), is a prerequisite for a fully ordered active site and that the 3′ end of the tRNA can only be properly bound in the enzyme after adenylate formation. A characteristic loop may seal the active site only for proline and the enzyme avoids that way the use of a pre-transfer editing mechanism for misactivated smaller amino acids such as alanine [18]. However, recent biochemical data suggest that *M. jannaschii* ProRS possesses both pre- and post-transfer editing mechanisms for alanine [18].

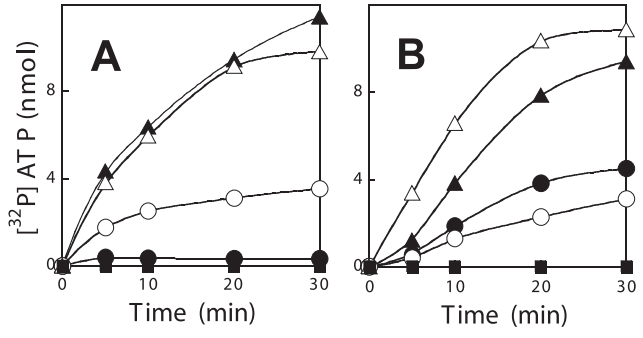

Fig. 1. Cysteine activation by *M. jannaschii* ProRS under different reaction conditions as measured by ATP–[^{32}P]PP$_i$ exchange [12]. A: Reaction in 50 mM HEPES pH 7.0, 50 mM KCl, 15 mM $MgCl_2$, 5 mM dithiothreitol, 2 mM KF and either 2 mM cysteine in the absence (●), or presence (▲) of unfractionated *M. jannaschii* tRNA, or 50 mM cysteine in the absence (○) or presence (△) of unfractionated *M. jannaschii* tRNA. B: Reaction [19,31] in 150 mM Tris–HCl pH 8.0, 10 mM $MgCl_2$, 10 mM 2-mercaptoethanol, 0.2 mg/ml bovine serum albumin, 2 mM KF and either 2 mM cysteine in the absence (●), or presence (▲) of unfractionated *M. jannaschii* tRNA, or 50 mM cysteine in the absence (○) or presence (△) of unfractionated *M. jannaschii* tRNA.

The role of tRNA in amino acid activation by *M. jannaschii* ProRS has been a subject of debate, as it may explain amino acid selection by the enzyme. Amino acid activation is measured by ATP–PP$_i$ exchange, the reverse reaction of Eq. 1 (see above). Like in all ProRS enzymes proline activation progresses in the absence of tRNA. However, the results of cysteine activation experiments appear to be influenced by reaction conditions and substrate concentration. It was reported that cysteine activation can be observed in the absence [11] or only in the presence [10,12] of total *M. jannaschii* tRNA in the reaction mixture. Therefore we compared cysteine activation under the different conditions (of buffer, pH or amino acid concentration) reported. Using the conditions of Stathopoulos et al. [12] (Fig. 1A) tRNA-independent activation of cysteine is hardly detectable above the background (reaction in the absence of either cysteine or enzyme), but it proceeds well in the presence of tRNA. However, under the conditions of Beuning and Musier-Forsyth [19] we observed (Fig. 1B) a significant amount of tRNA-independent cysteine activation, that could be stimulated 2–3-fold by addition of total *M. jannaschii* tRNA. Thus, it is clear that some of the reported discrepancies may be due to experimental conditions. As we did not observe much difference in similar experiments with proline (data not shown), it appears that cysteine activation by *M. jannaschii* ProRS is much more dependent on reaction conditions than is proline activation. The enzyme's sensitivity to its environment may be of advantage in vivo where the balance of the cell's free proline and cysteine determines the efficiency of Pro-tRNA and Cys-tRNA production.

Obligatory tRNA-dependent cysteine activation would be an appealing feature of the archaeal ProRS as it provides a plausible mechanism of amino acid discrimination [12]. On the other hand, if ProRS activates cysteine in the absence of tRNA, then a post-transfer editing mechanism would be essential for error correction, as cysteine could be transferred to any tRNA (including tRNA^{Pro}) that fits into the tRNA binding site. Such a mischarging event (the formation of Cys-tRNA^{Pro} using an in vitro transcript) was reported, but no post-transfer editing mechanism was shown to be involved [19].

Although *E. coli* and *Thermus aquaticus* ProRS have been purified and partially characterized many years ago, the spectrum of amino acid substrates was never determined [20,21]. It appears that cysteine activation may be an intrinsic characteristic of all ProRS enzymes. It was recently shown that *E. coli* ProRS, a bacterial type ProRS, can activate cysteine; either to 23% [19] or to $<0.1\%$ (under the conditions of Fig. 1A; I. Ahel, unpublished) of the k_{cat}/K_M value of proline. Many other ProRS enzymes also activate cysteine (I. Ahel and C. Stathopoulos, unpublished). If this leads to mischarged Cys-tRNA, it would provide a reason for the proposal that *E. coli*

ProRS possesses a post-transfer editing mechanism [22], the structural elements of which were attributed to the insertion domain (~180 amino acids) between motifs II and III, which is a characteristic feature of the bacterial type ProRS enzymes.

4. CysRS is not essential in *M. maripaludis*

Biochemical and genetic studies had shown that *M. maripaludis* contains a canonical CysRS in addition to a dual-specificity ProRS [10,23]. Although the K_M value of ProRS for cysteine is lower than that of the homologous CysRS, the overall Cys-tRNA formation rate for both enzymes is comparable [24]. If the dual-specificity enzyme were sufficient for Cys-tRNA formation, then the *cysS* gene may not be essential for the organism's growth. When the *M. maripaludis cysS* gene was disrupted the mutant cells were indeed viable [24], demonstrating that the canonical CysRS is not required for Cys-RNA formation in these cells. Thus, it is likely that ProRS (or an unknown protein) is the minimum requirement for cysteine incorporation into proteins in *M. maripaludis*.

This raises the general question whether deletion of other AARS genes may also be tolerated by certain organisms. Pertaining to this, it was reported in a transposon mutagenesis study of *Mycoplasma genitalium* that IleRS and TyrRS may be dispensable [25]. However, closer examination of the positions of transposon insertion made it plausible that these truncated enzymes could be functional and thus required for growth[26]. Furthermore, an examination of the phenotypes of yeast genome knockout strains shows that all 20 canonical AARSs involved in cytoplasmic protein synthesis are essential [27]. Thus, the likelihood of widespread occurrence of dual-specificity AARSs is not high. However, future extension of genome analysis to many organisms may uncover more such enzymes.

5. Outlook

The existence of dual-specificity AARSs for either charging or mischarging of tRNA is understandably very restricted. Therefore it is surprising to see that nature developed a number of different ways of activating cysteine. Four unrelated routes have been found to synthesize Cys-tRNA: the canonical CysRS, the archaeal genre dual-specificity ProRS, an unclassified [28] evolutionarily restricted CysRS [29], and a non-editing valyl-tRNA synthetase mutant [30]. What these diverse routes contribute to in vivo protein synthesis remains currently an open question.

Our understanding of the archaeal ProRS enzyme is just beginning. Detailed biochemical, biophysical and genetic analyses of the enzyme, the tRNA species and potential interacting proteins are essential. Is there a metabolic reason why only some organisms lack the canonical CysRS? What may these organisms tell us about horizontal gene transfer (of *cysS*)? Is the small dual-specificity ProRS an evolutionarily old enzyme? Genetic, genomic and more extensive biochemical studies in the relevant archaea are essential to provide in-depth answers to the many open questions.

Acknowledgements: We thank Hubert Becker, Shipra Bunjun-Srihari and Michael Ibba for critical discussions. This work was supported by grants from the National Institute of General Medical Sciences, the National Aeronautics and Space Administration, and the Office of Energy Biosciences of the Department of Energy.

References

[1] Crick, F.H.C. (1958) Symp. Soc. Exp. Biol. 12, 138–163.
[2] Ibba, M. and Söll, D. (2000) Annu. Rev. Biochem. 69, 617–650.
[3] Ibba, M., Becker, H.D., Stathopoulos, C., Tumbula, D.L. and Söll, D. (2000) Trends Biochem. Sci. 25, 311–316.
[4] Ibba, M. and Söll, D. (2001) EMBO Rep. 2, 382–387.
[5] Bult, C.J. et al. (1996) Science 273, 1058–1073.
[6] Smith, D.R. et al. (1997) J. Bacteriol. 179, 7135–7155.
[7] Slesarev, A.I. et al. (2002) Proc. Natl. Acad. Sci. USA 99, in press.
[8] Böck, A., Forchhammer, K., Heider, J. and Baron, C. (1991) Trends Biochem. Sci. 16, 463–467.
[9] Kim, H.S., Vothknecht, U.C., Hedderich, R., Celic, I. and Söll, D. (1998) J. Bacteriol. 180, 6446–6449.
[10] Stathopoulos, C., Li, T., Longman, R., Vothknecht, U.C., Becker, H.D., Ibba, M. and Söll, D. (2000) Science 287, 479–482.
[11] Lipman, R.S.A., Sowers, K. and Hou, Y.-M. (2000) Biochemistry 39, 7792–7798.
[12] Stathopoulos, C., Jacquin-Becker, C., Becker, H.D., Li, T., Ambrogelly, A., Longman, R. and Söll, D. (2001) Biochemistry 40, 46–52.
[13] Woese, C.R., Olsen, G.J., Ibba, M. and Söll, D. (2000) Microbiol. Mol. Biol. Rev. 64, 202–236.
[14] Ribas De Pouplana, L., Brown, J.R. and Schimmel, P. (2001) J. Mol. Evol. 53, 261–268.
[15] Bunjun, S., Stathopoulos, C., Graham, D., Min, B., Kitabatake, M., Wang, A.L., Wang, C.C., Vivares, C.P., Weiss, L.M. and Söll, D. (2000) Proc. Natl. Acad. Sci. USA 97, 12997–13002.
[16] Feng, L., Stathopoulos, C., Ahel, I., Mitra, A., Tumbula-Hansen, D.L., Hartsch, T. and Söll, D. (2002) Extremophiles, in press.
[17] Olendzenski, L., Liu, L., Zhaxybayeva, O., Murphey, M., Shin, D.G. and Gogarten, J.P. (2000) J. Mol. Evol. 51, 587–599.
[18] Yaremchuk, A., Tukalo, M., Grotli, M. and Cusack, S. (2001) J. Mol. Biol. 309, 989–1002.
[19] Beuning, J.P. and Musier-Forsyth, K. (2001) J. Biol. Chem. 276, 30779–30785.
[20] Lee, M.L. and Muench, K.H. (1969) J. Biol. Chem. 244, 223–230.
[21] Rivera, J.A., Wang, Q.S. and Wong, J.T. (1984) Can. J. Biochem. Cell. Biol. 62, 507–515.
[22] Beuning, J.P. and Musier-Forsyth, K. (2000) Proc. Natl. Acad. Sci. USA 97, 8916–8920.
[23] Li, T., Graham, D., Stathopoulos, C., Haney, P., Kim, H.-S., Vothknecht, U.C., Kitabatake, M., Hong, K.-W., Eggertsson, G., Curnow, A.W., Lin, W., Celic, I., Whitman, W. and Söll, D. (1999) FEBS Lett. 462, 302–306.
[24] Stathopoulos, C., Kim, W., Li, T., Anderson, I., Deutsch, B., Palioura, S., Whitman, W. and Söll, D. (2001) Proc. Natl. Acad. Sci. USA 98, 14292–14297.
[25] Hutchison, C.A., Peterson, S.N., Gill, S.R., Cline, R.T., White, O., Fraser, C.M., Smith, H.O. and Venter, J.C. (1999) Science 286, 2165–2169.
[26] Ruan, B.F. et al. (2001) Acta Biochim. Pol. 48, 313–321.
[27] Winzler, E.A. et al. (1999) Science 285, 901–906.
[28] Iyer, L.M., Aravind, L., Bork, P., Hofmann, K., Mushegian, A.R., Zhulin, I.B. and Koonin, E.V. (2001) Genome Biol. 2, 0051.1–0051.11.
[29] Fabrega, C., Farrow, M.A., Mukhopadhyay, B., de Crecy-Lagard, V., Ortiz, A.R. and Schimmel, P. (2001) Nature 411, 110–114.
[30] Döring, V., Mootz, H.D., Nangle, L.A., Hendrickson, T.L., de Crecy-Lagard, V., Schimmel, P. and Marliere, P. (2001) Science 292, 501–504.
[31] Heacock, D., Forsyth, C.J., Shiba, K. and Musier-Forsyth, K. (1996) Bioorg. Chem. 24, 273–289.