Molecular recognition of tRNA^{Pro} by *Escherichia coli* proline tRNA synthetase *in vitro*

Hongjian Liu, Richard Peterson, Jason Kessler and Karin Musier-Forsyth*

Department of Chemistry, University of Minnesota, 207 Pleasant Street S.E., Minneapolis, MN 55455, USA

Received September 1, 1994; Revised and Accepted November 15, 1994

ABSTRACT

In this study, we identify a subset of nucleotides that specify aminoacylation of tRNA^{Pro} by *Escherichia coli* proline tRNA synthetase *in vitro*. Twenty-two tRNA^{Pro} variants were prepared by *in vitro* transcription and their efficiency of aminoacylation with proline (k_{cat}/K_M) was measured. From this analysis, we conclude that recognition elements for tRNA^{Pro} aminoacylation by ProRS are located in at least three domains of the tRNA molecule. The largest decreases in the kinetic parameters for aminoacylation resulted from single substitutions at position G72 of the acceptor stem and position G36 of the anticodon. Anticodon nucleotide G35 and position A73 in the acceptor stem were also identified as major recognition elements. Moreover, bases that are believed to be important for maintaining the tertiary structure of the tRNA (G15 and C48) appear to be important for efficient recognition of tRNA^{Pro} by ProRS *in vitro*.

INTRODUCTION

Aminoacyl-tRNA synthetases catalyze a two step reaction that results in esterification of an amino acid onto the 3'-adenosine of their cognate tRNA substrates. These enzymes recognize specific tRNAs and discriminate among non-specific tRNA isoacceptors. There are twenty synthetases, one specific for each amino acid, and they have been partitioned into two classes of ten each based on regions of sequence similarity and structural data (1). Studies of tRNA recognition by aminoacyl-tRNA synthetases in recent years have focused on determining which nucleotides and structural features are important for positive recognition by the cognate synthetase and which negative elements are used to prevent tRNA recognition by noncognate synthetases (2, 3). Both positive and negative recognition elements are used to define the complete 'identity set' for a particular tRNA (4). The correlation between identity sets and the two synthetase classes is not entirely clear (5). These elements may be localized primarily in one domain of the tRNA molecule, or may be scattered throughout two or more domains. The two regions of the tRNA that have the highest density of identity nucleotides are the acceptor helix and the anticodon located at opposite ends of the L-shaped structure. Indeed, in the co-crystal structure of class II aspartyl-tRNA

synthetase complexed to tRNA^{Asp}, the enzyme clearly interacts with these two distant tRNA domains (6). Contacts which span the entire length of the molecule are also seen in the co-crystal structure of the class I glutamyl-tRNA synthetase-tRNA complex (7). In contrast, class II seryl-tRNA synthetase interacts with the acceptor-TΨC stem and loop domains as well as with the long variable arm of tRNA^{Ser}, whereas the anticodon is not contacted at all (8).

Escherichia coli proline-tRNA synthetase (ProRS) is a dimer of identical subunits with a total molecular mass of 127 kD (1). All three of the conserved sequence motifs that define class II synthetases have been identified in the primary sequence of this enzyme (1). An *in vivo* study showed that ProRS uses nucleotides in both the acceptor stem and the anticodon loop to recognize its cognate tRNA (9). Moreover, these authors concluded that ProRS requires a distinctive structure at the end of the acceptor helix for efficient aminoacylation. Recent experiments suggest that maintenance of structural elements in the core of tRNA^{Pro} is important for recognition by ProRS *in vitro* (10). In the present analysis, we use site-directed mutagenesis of *in vitro* transcribed tRNAs to probe the role of nucleotides in three different domains of tRNA^{Pro}. This information will contribute to efforts to obtain the complete recognition set of all *E. coli* tRNAs (4).

MATERIALS AND METHODS

Protein purification and assays

E. coli SY327 carrying two plasmids, pGT1-2 and pGT1-2# was a gift from Mike Syvaney. pGT1-2 contains the T7 RNA polymerase gene under the heat inducible λP_L promoter. pGT1-2# contains the ProRS gene under a T7 RNA polymerase promoter. The procedure used to prepare ProRS using this double plasmid expression system has been recently described (10).

ProRS protein concentrations were based on active site titrations using the adenylate burst assay (11). Aminoacylation assays of full-length tRNA variants by ProRS were conducted at 25°C as described (10).

T7 RNA polymerase was purified according to Grodberg and Dunn (12) from *E. coli* strain BL-21/pAR 1219 which was a gift of F. William Studier.

* To whom correspondence should be addressed

RESULTS

All three tRNA^{Pro} molecules from *E.coli* contain a unique C1:G72 base pair at the end of the acceptor stem helix (Figure 1). Because T7 RNA polymerase prefers a G at position 1 for efficient *in vitro* transcription, it is difficult to obtain high yields of a C1-containing transcript. Using a DNA template that does not encode C1 and priming with a CpG dinucleotide, we were able to obtain sufficient quantities of the full-length tRNA molecule to compare its aminoacylation efficiency with that of a transcript lacking the first base (Δ C1). Since the second base in tRNA^{Pro} is a G, high yields of Δ C1 transcripts could be obtained. In our hands, the latter was actually a better substrate for ProRS than the C1-transcript. Under our assay conditions, the K_M for the Δ C1 substrate is 4.2 μ M and the k_{cat} is 0.27 s⁻¹ ($k_{cat}/K_M = 0.064$ s⁻¹· μ M⁻¹) (10). The k_{cat}/K_M of the full-length C1-containing transcript was reduced approximately 3.5-fold relative to the Δ C1 transcript. This result and the relative ease in preparing large quantities of Δ C1 'wild-type' and mutant transcripts, prompted us

to prepare all of the mutants at positions other than C1 in the context of a $\Delta C1$ tRNA^{Pro}.

Figure 1 shows the structure of the UGG isoacceptor of *E. coli* tRNA^{Pro} used in this analysis (19). This figure also indicates the 15 conserved (open circles) and 8 semi-conserved (boxed) nucleotides found in most non-mitochondrial elongator tRNAs (19). The semi-conserved positions are constant purines or pyrimidines. Many of these 23 positions are important for the correct folding of tRNAs into their characteristic L-shape structure. The nucleotides indicated by shaded circles in this Figure are bases other than the universally-conserved or semi-conserved nucleotides that are common to all eight tRNA^{Pro} isoacceptors from *E. coli*, phages T4 and T5, and *Salmonella typhimurium* (20, 21). It should be noted that the D-loop nucleotide U17a is absent in the phage T5 tRNA^{Pro} sequence. However, this tRNA contains an additional C following the conserved guanosines at position 18 and 19 of the D-loop (not shown).

Acceptor stem mutants

Figure 2 shows the 22 mutants prepared in this study. Each arrow points to a single or multiple change that was made. The number in parenthesis is the change (x-fold) in k_{cat}/K_M relative to the $\Delta C1$ tRNA^{Pro} transcript. These values are also reported in Table 1. Due to the relatively low rates of aminoacylation of many of the mutants, individual values for k_{cat} and K_M could not be reliably determined for all of the transcripts and are, therefore, not reported. Single acceptor stem changes were made at the 4 positions A73, G72, C70, and C1. Three of these positions (A73, G72, and C1) are conserved in all *E.coli* tRNA^{Pro} isoacceptors

Figure 2. Summary of tRNA^{Pro} variant transcripts prepared and tested in this study. Single and multi-base changes are indicated by the arrows. Only those bases that were changed are shown explicitly. Symbols around these bases have the same meaning as in Figure 1. Other positions are indicated by a black filled-in circle. All changes made at positions other than 1 were made in the context of the $\Delta C1$ tRNA^{Pro} framework. The number in parenthesis is the change (x-fold) in k_{cat}/K_M relative to the 'wild-type' $\Delta C1$ tRNA^{Pro} transcript.

(Figure 1). However, only mutations at A73 and G72 resulted in significant decreases in k_{cat}/K_M (Table 1). All three base substitutions were made at the so-called 'discriminator base', position A73, and were somewhat higher (31-fold to 185-fold) when G72 was changed. As mentioned above, C1 could be removed or changed to a G with only minor effects on aminoacylation. A G1:C72 double mutant was also prepared. The decrease in aminoacylation seen with this mutant (77-fold) is approximately equal to the product of the decreases measured for the single changes made at these sites (31×2). Base pair 3:70, which is the location of important recognition elements in other tRNAs (15, 22–25), was also changed. The G3:C70 base pair was changed to either an A3:U70 base pair or a G3:U70 wobble base pair. These changes resulted in minor decreases of 3- and 6-fold, respectively (Table 1).

Anticodon-loop mutants

Single changes were made at all three anticodon positions (Figure 2 and Table 1). As expected, the U34→C change at the first position of the anticodon, which varies in all three tRNA^{Pro} isoacceptors, had relatively minor effects on aminoacylation

(4-fold). The middle anticodon base, however, is conserved among all isoacceptors, and the G35→C change had a larger effect (14-fold). Of all of the anticodon mutants tested in this analysis, the largest decrease in aminoacylation was seen with the G36→C change at the third position of the anticodon (164-fold). The effect of the single G36→C mutation was even larger than the 154-fold reduction in aminoacylation efficiency seen as a result of a triple UGG→CUA change at the anticodon positions. Single G36→A and G36→U changes had much less severe effects of 40- and 9-fold, respectively. An additional anticodon-loop nucleotide, the semi-conserved purine G37, was changed to cytosine. Even though this nucleotide is conserved among all tRNA^{Pro} isoacceptors and is next to a critical anticodon base, this change had no effect on aminoacylation.

Table 1. Specificity constants for aminoacylation with proline of tRNA^{Pro} mutant transcripts

Mutant type	tRNA ^{Pro} variant	k_{cat}/K_M	Loss of specificity
		(relative)	(x-fold)
tRNA ^{Pro}	$\Delta C1$	1	1
Acceptor stem	$\Delta C1 + A73 \rightarrow C73$	0.023	43
	$\Delta C1 + A73 \rightarrow G73$	0.0087	115
	$\Delta C1 + A73 \rightarrow U73$	0.034	29
	$\Delta C1 + G72 \rightarrow A72$	0.0054	185
	$\Delta C1 + G72 \rightarrow C72$	0.032	31
	$\Delta C1 + G72 \rightarrow U72$	0.0064	156
	C1→G1	0.48	2
	C1:G72→G1:C72	0.013	77
	$\Delta C1 + C70 \rightarrow U70$	0.17	6
	$\Delta C1 + G3:C70 \rightarrow A3:U70$	0.38	3
Anticodon loop	$\Delta C1 + U34 \rightarrow C34$	0.24	4
	$\Delta C1 + G35 \rightarrow C35$	0.070	14
	$\Delta C1 + G36 \rightarrow A36$	0.025	40
	$\Delta C1 + G36 \rightarrow C36$	0.0061	164
	$\Delta C1 + G36 \rightarrow U36$	0.11	9
	$\Delta C1 + G37 \rightarrow C37$	0.71	1
	$\Delta C1 + U34, G35, G36 \rightarrow$		
	C34, U35, A36	0.0065	154
	$\Delta C1 + G15 \rightarrow C15$	0.016	63
	$\Delta C1 + U17a \rightarrow C17a$	0.30	3
D-loop and variable loop	$\Delta C1 + C48 \rightarrow G48$	0.051	20
	$\Delta C1 + G15, C48 \rightarrow C15, G48$	0.051	20

The specificity constant (k_{cat}/K_M) for aminoacylation of tRNA^{Pro} variant transcripts was derived from a Lineweaver–Burke plot. Five tRNA concentrations (0.5 to 12 μM) were used and the ProRS concentration was 50 nM. The results are averages of 2–3 determinations with standard deviations of $\pm 25\%$.

D-loop and variable loop mutants

A final set of $\Delta C1$ tRNA^{Pro} variants was prepared containing changes in the two regions of the tRNA that are involved in extensive tertiary interactions, the D-loop and variable loop.

These domains are part of the core region of the tRNA and are important in maintaining the L-shape structure. To probe the significance of the conserved 'extra' nucleotide U17a in the D-loop, we prepared a U17a→C mutant. The minor reduction in aminoacylation seen with this mutant (3-fold) suggests that it is not an element involved in positive tRNA^{Pro} recognition by ProRS.

Recent experiments using chemically-synthesized tRNA fragments corresponding to the 5'-18 nucleotides of tRNA^{Pro}, demonstrated that deletion of G15 resulted in a dramatic decrease in aminoacylation efficiency. This result suggested that maintenance of the semi-conserved trans base pair involving nucleotides 15 and 48 is important for efficient recognition of tRNA^{Pro} by ProRS (10). To test this hypothesis, we constructed three mutants that altered the proposed hydrogen bonding interactions of these bases (Figure 2 and Table 1). A G15→C change that creates a C15,C48 mismatch had a large effect on aminoacylation (63-fold). The single C48→G change had a less severe, but still significant effect (20-fold). The C15:G48 double mutant was also tested and resulted in a 20-fold reduction in aminoacylation efficiency relative to the wild-type ΔC1 tRNA transcript. This reduction is much less than the decrease expected if these sites were acting independently (63×20), and supports the importance of this interaction.

DISCUSSION

A previous computer analysis of 67 tRNA sequences revealed that the nucleotides distinguishing tRNA^{Pro} from other tRNAs include the C1:G72 base pair in the acceptor stem, U17a in the D-loop, and G35, G36, and G37 in the anticodon loop (9). This analysis predicted three regions of the molecule may be important for tRNA^{Pro} identity. An *in vivo* experimental study by these same researchers using the amber-suppressor tRNA system, concluded that A73, C1:G72 and one or more of the anticodon nucleotides contribute to the tRNA^{Pro} identity set. Furthermore, seven changes (U17a, C1:G72, C16, G45, C51, and A59) were sufficient to partially switch the identity of tRNA^{Phe} to tRNA^{Val} *in vivo*. An independent *in vitro* analysis showed that a tRNA^{Val} mutant (A73) with the UGG anticodon of proline and with a G1:G72 base pair was charged with proline by partially purified *E.coli* aminoacyl-tRNA synthetase (26). Discriminator base changes to G, C, or U, however, prevented aminoacylation of the tRNA^{Val} mutant.

In the present study, an *in vitro* analysis of tRNA^{Pro} recognition by purified ProRS was carried out. In Figure 3, the nucleotides tested are explicitly shown. The arrows point to those nucleotides identified as positive recognition elements based on the aminoacylation results shown in Figure 2 and Table 1. In this analysis, a 10-fold decrease in aminoacylation efficiency is considered to be 'significant'. Mutations at six positions, A73, G72, C48, G36, G35, and G15 had significant effects on kinetic parameters (Figure 3, arrows). Although C1, U17a, and G37 are among those nucleotides that a computer analysis predicted may be used to distinguish tRNA^{Pro} from other tRNAs (9), the present *in vitro* analysis shows that they are not critical ProRS recognition elements.

The fourth base from the 3'-end, the 'discriminator base', is important for the identity of many tRNAs (4). ProRS appears to recognize this position as well, since all three base substitutions at position 73 had significant effects on aminoacylation (Table 1).

Figure 3. A summary of the recognition nucleotides in tRNA^{Pro} identified in this study. Symbols around the bases have the same meaning as in Figures 1 and 2. Recognition nucleotides are indicated by the arrows. Site-directed changes at these positions resulted in decreases of > 10 -fold in the specificity constant (k_{cat}/K_M) for aminoacylation with proline. Smaller arrows point to bases that are part of a conserved tertiary interaction found to be important for recognition by ProRS.

Our experiments also agree well with the computer prediction and *in vivo* data that suggest G72 is an identity element for tRNA^{Pro} (9). Furthermore, this study demonstrates the importance of anticodon loop nucleotides G35 and G36 in tRNA^{Pro} recognition by purified ProRS *in vitro*. Interestingly, the effects of different substitutions at position 36 varied over a wide range (9- to 164-fold). Similar results have been obtained with anticodon mutations in other tRNAs. For example, changes at position 35 of yeast tRNA^{Asp} resulted in 19- to 530-fold decreases in aspartylation (16). In *E.coli* tRNA^{Gln}, large decreases of k_{cat}/K_M ranging from 2×10^3 -fold to 3×10^5 -fold were seen upon substitution of anticodon positions C34→A, U35→A or C, and G36→A (15). In comparison, a G36→U substitution had a relatively modest 67-fold effect (15). A recent study of *E.coli* tRNA^{Ile} found that while no bases can efficiently substitute for the original ones at positions 34 and 35, decreases in k_{cat}/K_M ranging from 6- to 220-fold were measured upon substitution of U36 (27).

Single-stranded D- and TΨC-loop residues constitute the domain of tRNAs known as the 'variable pocket'. Insertions or deletions in the D-loop could change the shape of this pocket and could form part of the identity set for a particular tRNA (28). The results of a previous *in vitro* study using annealed tRNA^{Pro} fragments suggest that while the 'extra' nucleotide, U17a, in the D-loop and the size of the D-loop *per se* are not critical for aminoacylation with proline, certain D-loop tertiary interactions

are important (10). The mutagenesis results presented here agree with this analysis and with the conclusion that maintenance of the conserved tertiary interaction between G15 and C48 in the core of tRNA^{Pro} is important for recognition by ProRS *in vitro* (Figure 3, small arrows).

The first base of tRNA^{Pro}, C1, is part of a unique C1:G72 base pair. Surprisingly, it is dispensable for aminoacylation. In the three-dimensional crystal structure of class I *E.coli* glutamyl-tRNA synthetase complexed with tRNA^{Gln}, the first base pair (U1:A72) is disrupted (7). Furthermore, a Δ U1 tRNA^{Gln} transcript is an efficient substrate for glutamyl-tRNA synthetase (15). The crystal structure of the ProRS-tRNA^{Pro} complex has not yet been solved. Nevertheless, the fact that a Δ C1 tRNA^{Pro} transcript is a slightly better substrate (by approximately 3.5-fold) than a wild-type C1-containing transcript supports the previous suggestion that the first base pair of tRNA^{Pro} is disrupted upon complex formation with its cognate class II synthetase (9).

It has been previously observed that the larger the number of identity nucleotides in a tRNA molecule, the less the importance of their individual contributions (5). Of the positions identified as recognition elements for tRNA^{Pro} (Figure 3), only modest decreases in aminoacylation efficiency were seen ranging from 14-fold to 185-fold. Furthermore, this and previous work suggest that elements that stabilize the tertiary structure of tRNA^{Pro} also contribute significantly to recognition by ProRS (9, 10). From these studies, we can conclude that *E.coli* tRNA^{Pro} belongs to the subset of tRNAs with a large number of relatively weak determinants located in at least three domains of the tRNA molecule.

ACKNOWLEDGEMENTS

We would like to thank Dr Mike Syvaney for providing us with the over-expression plasmid for ProRS and Dr F. William Studier for the overexpression plasmid for T7 RNA polymerase. This work was funded by grant number GM49928 from the National Institutes of Health. Acknowledgment is also made to the donors of The Petroleum Research Fund, administered by the American Chemical Society and to the American Cancer Society (Grant IN-13-33-15) for partial support of this research.

REFERENCES

- 1 Eriani, G., Delarue, M., Poch, O., Gangloff, J. and Moras, D. (1990) *Nature*, 347, 203–206.
- 2 McClain, W.H. (1993) *J. Mol. Biol.*, 234, 257–280.
- 3 Saks, M.E., Sampson, J.R. and Abelson, J.N. (1994) *Science*, 263, 191–197.
- 4 Schulman, L.H. (1991) *Prog. Nucleic Acid Res. and Mol. Biol.*, 41, 23–87.
- 5 Giegé, R., Puglisi, J.D. and Florentz, C. (1993) *Prog. Nucleic Acid Res. and Mol. Biol.*, 45, 129–205.
- 6 Ruff, M., Krishnaswamy, S., Boeglin, M., Poterszman, A., Mitschler, A., Podjarny, A., Rees, B., Thierry, J.C. and Moras, D. (1991) *Science*, 252, 1682–1689.
- 7 Rould, M.A., Perona, J.J., Söll, D. and Steitz, T.A. (1989) *Science*, 246, 1135–1142.
- 8 Biou, V., Yaremchuk, A., Tukalo, M. and Cusack, S. (1994) *Science*, 263, 1404–1410.
- 9 McClain, W.H., Schneider, J. and Gabriel, K. (1994) *Nucleic Acids Res.*, 22, 522–529.
- 10 Liu, H. and Musier-Forsyth, K. (1994) *Biochemistry*, 33, 12708–12714.
- 11 Fersht, A.R., Ashford, J.S., Bruton, C.J., Jakes, R., Koch, G.L.E. and Hartley, B.S. (1975) *Biochemistry*, 14(1), 1–4.
- 12 Grodberg, J.D. and Dunn, J.J. (1988) *J. Bacteriol.*, 170, 1245–1253.
- 13 Kunkel, T.A. (1985) *Proc. Natl. Acad. Sci. U.S.A.*, 82, 488.
- 14 Milligan, J.F. and Uhlenbeck, O.C. (1989) *Methods Enzymol.*, 180, 51–62.
- 15 Jahn, M., Rogers, M.J. and Söll, D. (1991) *Nature*, 352, 258–260.
- 16 Pütz, J., Puglisi, J.D., Florentz, C. and Giegé, R. (1991) *Science*, 252, 1696–1699.
- 17 Sampson, J.R. and Uhlenbeck, O.C. (1988) *Proc. Natl. Acad. Sci. U.S.A.*, 85, 1033–1037.
- 18 Yue, D., Kintanar, A. and Horowitz, J. (1994) *Biochemistry*, 33, 8905–8911.
- 19 Sprinzl, M., Hartmann, T., Weber, J., Blank, J. and Zeidler, R. (1989) *Nucleic Acids Res.*, 17, r1–r172.
- 20 Steinberg, S., Misch, A. and Sprinzl, M. (1993) *Nucleic Acids Res.*, 21, 3011–3015.
- 21 Kuchino, Y., Yabusaki, Y., Mori, F. and Nishimura, S. (1984) *Nucleic Acids Res.*, 12, 1559–1562.
- 22 Francklyn, C., Shi, J.-P. and Schimmel, P. (1992) *Science*, 255, 1121–1125.
- 23 Hou, Y.-M. and Schimmel, P. (1988) *Nature*, 333, 140–145.
- 24 Normanly, J., Ogden, R.C., Horvath, S.J. and Abelson, J. (1986) *Nature*, 321, 213–219.
- 25 McClain, W.H. and Foss, K. (1988) *Science*, 241, 1804–1807.
- 26 Shimizu, M., Asahara, H., Tamura, K., Hasegawa, T. and Himeno, H. (1992) *J. Mol. Evol.*, 35, 436–443.
- 27 Nureki, O., Niimi, T., Muramatsu, T., Kanno, H., Kohno, T., Florentz, C., Giegé, R. and Yokoyama, S. (1994) *J. Mol. Biol.*, 226, 710–724.
- 28 Normanly, J. and Abelson, J. (1989) *Ann. Rev. Biochem.*, 58, 1029–1049.