Evolutionary Coadaptation of the Motif 2−Acceptor Stem Interaction in the Class II Prolyl-tRNA Synthetase System†

Brian Burke, Fan Yang,‡ Fei Chen, Catherine Stehlin,§ Barden Chan,|| and Karin Musier-Forsyth*

Department of Chemistry, University of Minnesota, Minneapolis, Minnesota 55455

Received August 4, 2000; Revised Manuscript Received September 25, 2000

ABSTRACT: Known crystal structures of class II aminoacyl-tRNA synthetases complexed to their cognate tRNAs reveal that critical acceptor stem contacts are made by the variable loop connecting the β-strands of motif 2 located within the catalytic core of class II synthetases. To identify potential acceptor stem contacts made by *Escherichia coli* prolyl-tRNA synthetase (ProRS), an enzyme of unknown structure, we performed cysteine-scanning mutagenesis in the motif 2 loop. We identified an arginine residue (R144) that was essential for tRNA aminoacylation but played no role in amino acid activation. Cross-linking experiments confirmed that the end of the $tRNA^{Pro}$ acceptor stem is proximal to this motif 2 loop residue. Previous work had shown that the $tRNA^{Pro}$ acceptor stem elements A73 and G72 (both strictly conserved among bacteria) are important recognition elements for *E. coli* ProRS. We carried out atomic group "mutagenesis" studies at these two positions of *E. coli* $tRNA^{Pro}$ and determined that major groove functional groups at A73 and G72 are critical for recognition by ProRS. Human $tRNA^{Pro}$, which lacks these elements, is not aminoacylated by the bacterial enzyme. An analysis of chimeric $tRNA^{Pro}$ constructs showed that, in addition to A73 and G72, transplantation of the *E. coli* $tRNA^{Pro}$ D-domain was necessary and sufficient to convert the human tRNA into a substrate for the bacterial synthetase. In contrast to the bacterial system, base-specific acceptor stem recognition does not appear to be used by human ProRS. Alanine-scanning mutagenesis revealed that motif 2 loop residues are not critical for tRNA aminoacylation activity of the human enzyme. Taken together, our results illustrate how synthetases and tRNAs have coadapted to changes in protein−acceptor stem recognition through evolution.

Aminoacyl-tRNA synthetases catalyze the attachment of amino acids to the 2′- or 3′-hydroxyl of the terminal ribose of cognate tRNA substrates. The aminoacylated tRNAs are then delivered to the ribosome, where they interact with the mRNA codons, delivering amino acids to the site of protein synthesis. The fidelity of this process is dependent upon the ability of synthetases to accurately recognize only their cognate tRNAs in vivo and to discriminate against noncognate tRNA isoacceptors. Studies of synthetases from *Escherichia coli* have demonstrated that the anticodon plays an important role in tRNA recognition in 17 of the 20 aminoacylation systems (*1*). In most systems, important base or backbone interactions between the tRNA acceptor stem and residues in or near the active site of the synthetase are also critical for specific tRNA aminoacylation (*2*, *3*).

The aminoacyl-tRNA synthetases are divided into two classes based on common catalytic domains (*4*, *5*). The class I synthetases contain the conserved HIGH and KMSKS motifs and a Rossmann fold catalytic core (*6*). The catalytic domain of class II synthetases contains motifs 2 and 3 and is made up of an antiparallel β-sheet motif (*7*, *8*). The majority of class II synthetases are dimeric and motif 1 constitutes the dimer interface (*4*). The class II synthetases are further divided into subclasses. On the basis of the location of the anticodon binding domain, class II prolyl-tRNA synthetase (ProRS)[1] is classified as a class IIa enzyme (*5*). Phylogenetic analysis has shown that ProRSs can be further separated into two distinct groups based on the presence of either a large insertion between motifs 2 and 3 ("prokaryotic-like") or a C-terminal extension ("eukaryotic-like") (*9*, *10*). Previously, it was shown that prokaryotic-like *E. coli* ProRS recognizes A73 and G72 in the acceptor stem of *E. coli* $tRNA^{Pro}$ (*11*). In contrast, eukaryotic-like human ProRS does not appear to recognize specific bases in the acceptor stem of human $tRNA^{Pro}$ (*10*). In both the *E. coli* and human systems, however, base-specific recognition occurs in the anticodon at positions G35 and G36 (*10*, *11*). These biochemical data are in good agreement with an X-ray

† This work was funded by Grant GM49928 from the National Institutes of Health. The donors of the Petroleum Research Fund, administered by the American Chemical Society, are also acknowledged for partial support of this research. B.B. is partially supported by a National Institutes of Health predoctoral training grant (T32 GM08277).

* To whom correspondence should be addressed at the Department of Chemistry, University of Minnesota, 207 Pleasant St. S. E., Minneapolis, MN 55455. Telephone (612) 624-0286; fax (612) 626-7541; e-mail musier@chem.umn.edu.

‡ Present address: University of Colorado Health Science Center, Denver, CO 80262.

§ Present address: Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/INSERM, Université Louis Pasteur, 1 Rue Laurent Fries, 67404 Illkirch Cedex, France.

|| Present address: Renal Division, Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA 02215.

[1] Abbreviations: AspRS, aspartyl-tRNA synthetase; BMH, 1,6-bis-(maleimido)hexane; ProRS, prolyl-tRNA synthetase.

10.1021/bi001835p CCC: $19.00
Published on Web 11/16/2000

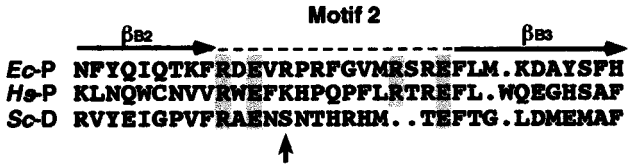

FIGURE 1: Sequence alignment comparing the motif 2 region of *E. coli* (*Ec*-P) and human (*Hs*-P) ProRS with that of yeast AspRS (*Sc*-D). Shaded boxes indicate residues that are conserved among all ProRSs. The small arrow points to residue S329 of yeast AspRS.

crystal structure showing the anticodon domain bound to *Thermus thermophilus* ProRS (*9*).

Cocrystal structures of class II synthetases complexed to cognate tRNAs have shown that specific contacts occur between the acceptor stem of the tRNA and the motif 2 loop of the synthetase (*8*, *12*−*15*). Given the similarity in the anticodon recognition between the two groups of ProRS, we imagined that the differences observed in acceptor stem recognition between prokaryotic-like and eukaryotic-like ProRS may be reflected in species-specific differences in functional interactions with the motif 2 loop. Moreover, on the basis of sequence alignments, we noticed that the motif 2 loop sequence is highly conserved within each group of ProRSs (*10*), but differs between the two groups (Figure 1). Although a cocrystal structure of *T. thermophilus* ProRS, a eukaryotic-like enzyme, has been solved complexed to $tRNA^{Pro}$, the acceptor stem is not bound in the active site and only the anticodon interactions are clearly resolved (*9*). In addition, the structure of a ProRS from the prokaryotic-like group has not been reported. Thus, to test the hypothesis that there are species-specific differences in the role of motif 2 in $tRNA^{Pro}$ recognition, we prepared mutations within the motif 2 loop sequence of a representative ProRS from each group. Aminoacylation assays using these mutant ProRS constructs revealed that motif 2 loop residues are critical for aminoacylation by *E. coli* ProRS but not for activity of the human enzyme.

In general, class II synthetases approach the tRNA acceptor stem from the major groove side and thus are more likely to make specific contacts within this groove (*8*). For the *E. coli* system, to establish whether the critical motif 2 loop interactions occur in the major groove, we carried out atomic group "mutagenesis" (*16*). In particular, base analogues with specific atomic group changes were incorporated into positions 72 and 73 of a semi-synthetic *E. coli* $tRNA^{Pro}$ construct (*17*). In a separate study, aminoacylation assays were carried out with chimeric $tRNA^{Pro}$ constructs to identify the barrier to cross-species aminoacylation of human $tRNA^{Pro}$ by *E. coli* ProRS. The combined results of this study and previous work have allowed us to propose a model for acceptor stem recognition in the prokaryotic-like ProRS system. We also show here that this RNA−protein interaction has changed through evolution and that these differences contribute to species-specific tRNA recognition in the ProRS system.

MATERIALS AND METHODS

Enzyme Purification and Site-Directed Mutagenesis of ProRS. Purification of histidine-tagged wild-type *E. coli* and human ProRS was accomplished as described previously with plasmids pCS-M1S and p*K*S-509, respectively (*18*, *19*). Site-directed mutagenesis of each of the five *E. coli* motif 2 loop residues was carried out in the background of a C443G mutant version of ProRS containing no Cys residues. The C443G change in motif 3 has only minor effects on aminoacylation activity (*19*), and use of this mutant allowed single Cys substitutions to be incorporated into the motif 2 loop. Site-directed mutagenesis was performed either by the Kunkel method (*20*) or via overlap extension PCR (*21*) with DNA primers encoding the desired amino acid changes (V143C, R144C, P145C, R146C, and F147C in *E. coli* ProRS and F1083A, K1084A, H1085A, P1086A, and Q1087A in human ProRS). Following the mutagenesis procedures, the plasmid encoding *E. coli* ProRS was digested with *Bam*HI and *Pst*I to generate a 511 base pair fragment. The plasmid encoding the human ProRS was digested with *Sac*I and *Spe*I to generate a 718 base pair fragment. These fragments were subcloned back into the respective wild-type or C443G mutant plasmids that had been digested with the same restriction enzymes. The entire 511 or 718 base pair regions of the final constructs were sequenced to confirm the existence of only the desired mutation. The mutant plasmids were then transformed into *E. coli* strains SG13009 [pREP4] (Qiagen, *E. coli* ProRS) or BL21(DE3) pLys3 (Novagen, human ProRS) for overexpression and purification as described previously (*18*, *19*).

Activity Assays. ProRS concentrations were based on active-site titrations determined by the adenylate burst assay (*22*). Aminoacylation assays were carried out as described previously (*18*). The ATP−PP_i exchange assay was also carried out using published procedures (*18*, *23*, *24*). The kinetic constants were derived from Lineweaver−Burk plots.

RNA Preparation and Mutagenesis. All full-length tRNAs were prepared by in vitro transcription as described (*18*). To facilitate in vitro transcription, all *E. coli* $tRNA^{Pro}$ constructs lacked C1 as described previously (*25*). Generation of mutant tRNAs was accomplished either by cloning a set of six overlapping synthetic DNA oligonucleotides encoding the desired mutant tRNA gene downstream of a T7 RNA polymerase promoter (*26*) or via mutagenesis by overlap extension PCR (*21*). Following all mutagenesis and cloning procedures, the existence of only the desired mutations was verified by sequencing of the entire tRNA coding region. The semi-synthetic tRNAs containing base analogue substitutions were prepared as previously described (*17*). Briefly, modified nucleotide analogues (Glen Research and Chem-Genes Corp.) were incorporated into a synthetic 3′-16-mer oligonucleotide during automated chemical synthesis by the phosphoramidite method on a Gene Assembler Special (Pharmacia). These 3′-fragments were then heat-annealed to a 5′-$^3/_4$-length in vitro transcribed RNA fragment to form a semi-synthetic full-length *E. coli* ΔC1-$tRNA^{Pro}$. All RNAs were purified by denaturing 12% or 16% polyacrylamide gel electrophoresis as previously described (*17*).

Cross-Linking Procedure. A bismaleimide homobifunctional cross-linking agent, 1,6-bis(maleimido)hexane (BMH) (Pierce) was used to site-specifically cross-link the 5′-end of a 5′-phosphorothioate-containing $tRNA^{Pro}$ to *E. coli* ProRS mutants containing a single Cys residue in the motif 2 loop (C443G/V143C and C443G/R144C). The wild-type enzyme, which contains a single Cys in motif 3 (C443), and the C443G ProRS mutant, which has no Cys residues, were also tested as controls. The tRNAs used in the cross-linking

Table 1: Effect of Single Amino Acid Changes in the Motif 2 Loop of *E. coli* and Human ProRS on Aminoacylation of Their Cognate $tRNA^{Pro}$ Transcripts[a]

ProRS species	mutation	k_{cat}/K_M ($s^{-1}\cdot\mu M^{-1}$)	k_{cat}/K_M (relative)	*x*-fold change
E. coli[b]	*e*WT	2.02×10^{-2}	3.7	+3.7
	C443G	5.39×10^{-3}	1.0	
	C443G/V143C	1.79×10^{-3}	0.33	−3.0
	C443G/R144C	nd	nd	−1000
	C443G/P145C	2.56×10^{-4}	0.047	−21.0
	C443G/R146C	6.79×10^{-5}	0.012	−79.1
	C443G/F147C	3.85×10^{-4}	0.071	−14.0
human[c]	hWT	4.3×10^{-3}	1.0	
	F1083A	3.10×10^{-3}	0.72	−1.4
	K1084A	6.64×10^{-3}	1.5	+1.5
	H1085A	2.30×10^{-3}	0.53	−1.9
	P1086A	8.15×10^{-3}	1.9	+1.9
	Q1087A	2.10×10^{-3}	0.50	−2.0

[a] Concentrations of tRNA ranging from 0.5 to 20 μM were used, and the ProRS concentration was 50 or 125 nM. Each assay was carried out at least twice with the values differing by <19%. [b] *e*WT corresponds to wild-type *E. coli* ProRS, which was assayed on the same day as each mutant. Motif 2 mutations were made in the context of a C443G mutant form of *E. coli* ProRS and the relative k_{cat}/K_M is normalized to the value obtained for this construct, which was set at 1.0. [c] hWT refers to wild-type human ProRS and the relative k_{cat}/K_M of all human mutant constructs is normalized to the wild-type enzyme, which was assayed on the same day as each mutant. nd indicates not detected.

experiments were prepared by in vitro transcription with 1 mM each NTP and 5 mM 5′-αS-GMP (Amersham). This results in specific incorporation of a single phosphorothioate at the 5′-end of the transcript. A small-scale (50-μL) transcription was also performed to generate [^{32}P]-labeled tRNA. This reaction contained 4 mM each ATP, CTP, and UTP, 1 mM GTP, 5 mM 5′-αS-GMP, and 17 mCi/mL [α-^{32}P]GTP (Amersham). Cross-linking reactions were carried out in 20-μL reaction volumes containing 2 μM enzyme monomer and 4 μM tRNA preincubated in 20 mM sodium phosphate and 150 mM NaCl, pH 7.0, at room temperature for 5 min. Addition of BMH to 250 μM initiated the reaction. At different times (1−4 h), the reaction was quenched with 100 mM DTT. Protein loading buffer was added and the samples were boiled for 5 min preceding electrophoresis on denaturing polyacrylamide gels. Gels were visualized by Coomassie Blue staining, autoradiography, and phosphorimaging.

RESULTS

Motif 2 Loop Mutations. To test the hypothesis that differences in motif 2 loop interactions contribute to species-specific differences in $tRNA^{Pro}$ acceptor stem recognition, we carried out Cys- and Ala-scanning mutagenesis of residues that constitute the motif 2 loop of *E. coli* and human ProRS, respectively. In the case of the *E. coli* enzyme, we chose to incorporate Cys rather than the more standard Ala because this would facilitate site-specific attachment of a cross-linking probe. Therefore, mutagenesis was carried out in the context of a Cys-less C443G-ProRS background. In the case of the human enzyme, which contains a total of 12 Cys residues, this approach was not practical and so a more standard Ala scan was done. Table 1 shows the results of aminoacylation assays with the 10 motif 2 loop mutant constructs prepared in this work. We chose to target amino acids that were the least well conserved between the two groups of ProRS and that aligned with amino acids in the motif 2 loop of yeast aspartyl-tRNA synthetase (AspRS) shown to form hydrogen bonds with G73 of cognate $tRNA^{Asp}$ (*12*). Thus, we mutagenized the VRPRF sequence in *E. coli* ProRS and the corresponding FKHPQ sequence in the human system (Figure 1). However, in contrast to the yeast AspRS system, where significant effects on aminoacylation were observed upon mutagenesis of motif 2 loop residues (*27*, *28*), Ala-scanning mutagenesis indicates that none of the motif 2 loop residues tested in human ProRS are critical for aminoacylation activity (Table 1). The largest decreases observed, H1085A and Q1087A, were only ∼2-fold reduced relative to wild-type human ProRS, while two mutations, K1084A and P1086A, resulted in a slight increase in aminoacylation efficiency.

Cysteine-scanning mutagenesis of the motif 2 loop residues of *E. coli* ProRS yielded strikingly different results (Table 1). A small decrease in aminoacylation activity (3-fold) was observed for the V143C mutant, while moderate decreases were observed for P145C, R146C, and F147C (14−79-fold). However, mutating R144 to Cys completely abolished aminoacylation activity. Based on the level of detection of these assays, the lack of activity corresponds to a ≥1000-fold effect. In contrast to the dramatic effect on aminoacylation, the R144C mutation had very little effect on proline activation as determined by the ATP−PP_i exchange assay. The k_{cat}/K_M^{Pro} of the C443G/R144C double mutant was determined to be 6.0 s^{-1} mM^{-1}, which is very similar to the value of 6.3 s^{-1} mM^{-1} measured for the C443G variant (*19*). These results indicate that R144 is not critical for amino acid binding or activation but is required for amino acid transfer to *E. coli* $tRNA^{Pro}$.

E. coli ProRS Cross-Linking Studies. To further characterize the motif 2 loop interaction with the acceptor stem of *E. coli* $tRNA^{Pro}$, we designed a cross-linking experiment that took advantage of the single-Cys ProRS variants. In particular, the most active *E. coli* ProRS mutant (C443G/V143C, Table 1), which contains a single Cys adjacent to the critical R144 residue, was used in this work. This variant was incubated with 5′-phosphorothioate-containing *E. coli* $tRNA^{Pro}$ and the BMH cross-linking agent (Figure 2A). Following the cross-linking reaction, samples were analyzed on a denaturing polyacrylamide gel. The results are shown in Figure 2B. We observe a time-dependent increase in a slower migrating band, which represents the ProRS−tRNA cross-linked complex (Figure 2B, lanes 1−5). These results indicate that the motif 2 loop residue V143 is in close proximity to the 5′-end of the acceptor stem of *E. coli* $tRNA^{Pro}$. The control reactions carried out either in the absence of BMH (Figure 2B, lane 6) or with C443G-ProRS that lacks a Cys residue (Figure 2B, lane 9) show no cross-linked product. Wild-type ProRS, which has a single Cys in motif 3, also does not form a cross-link (Figure 2B, lane 8). In contrast to the V143C variant, the inactive R144C mutant also does not cross-link to *E. coli* $tRNA^{Pro}$ (data not shown). Due to its complete lack of detectable aminoacylation activity (Table 1), we are unable to establish whether this mutant is defective in k_{cat} or K_M. However, the cocrystal structure of *T. thermophilus* ProRS complexed to $tRNA^{Pro}$ shows that the anticodon interactions can occur even in the absence of acceptor stem binding (*9*). Thus, one possibility is that the R144C mutant binds to the tRNA with the acceptor stem

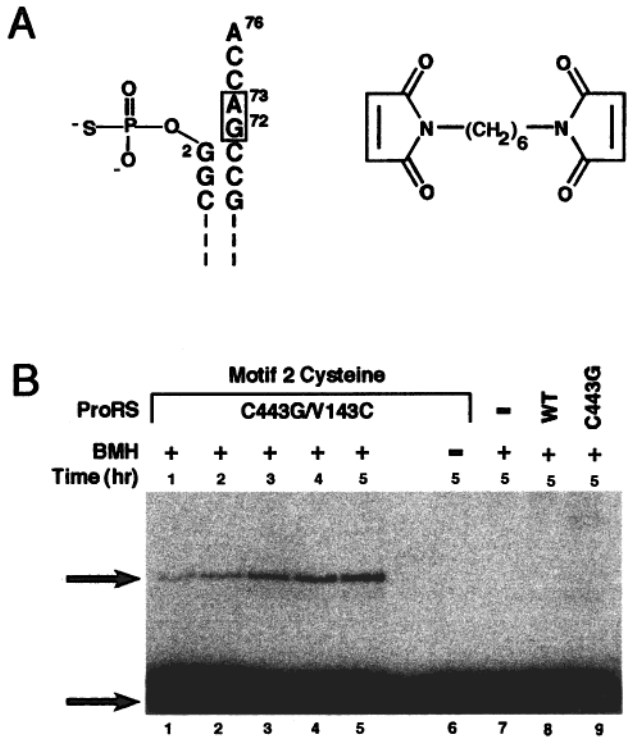

FIGURE 2: Design and results of cross-linking experiment using *E. coli* ProRS variants. (A) Top part of the acceptor stem of 5′-phosphorothioate-containing *E. coli* ΔC1-tRNA^{Pro} with the major recognition elements boxed (left) and structure of the BMH cross-linker used (right). (B) Phosphorimage of an SDS−10% polyacrylamide gel showing the results of a cross-linking experiment. [^{32}P]-Labeled tRNA^{Pro} was incubated with C443G/V143C *E. coli* ProRS in the presence (lanes 1−5) or absence (lane 6) of BMH for the times indicated above each lane. Experiments were also conducted in the absence of ProRS (lane 7), with wild-type ProRS (lane 8), and with C443G ProRS, which lacks a Cys residue (lane 9). The arrows indicate the positions of the free tRNA (bottom) and the cross-linked product (top).

bound in a nonproductive manner such that the CCA-3′ end of the tRNA does not reach the catalytic site.

Atomic Group Mutagenesis. Previous studies identified G72 and A73 as important recognition elements for *E. coli* ProRS (*11*, *29*). The cross-linking data support the view that the critical motif 2 loop residues are proximal to these sites at the top of the tRNA acceptor stem (Figure 2A). To gain further evidence for specific interactions at G72 and A73, we incorporated base analogues at these positions of semi-synthetic tRNA^{Pro} constructs. Figure 3 shows the purine analogues that were examined.

Previously, it was shown that an A73G mutation reduced aminoacylation activity by 115-fold (*11*). Our new data help to explain the strong preference for an A over a G at this "discriminator base" position (*30*) and provide additional insights into recognition at this site. Whereas the exocyclic amino group is not a major recognition element, as shown by the modest effect of the 2′-deoxypurine base substitution at this site (−3.6-fold), introduction of an amino group in the minor groove at position 2 has a strong negative effect (Figure 3A). For example, substitution with 2′-deoxy-2-aminopurine results in an 18-fold decrease in aminoacylation relative to substitution with 2′-deoxypurine, and substitution with 2′-deoxydiaminopurine results in a 48-fold decrease in activity relative to wild-type tRNA^{Pro} (Figure 3A). The 2′-deoxy-7-deaza-A substitution results in a 32-fold decrease

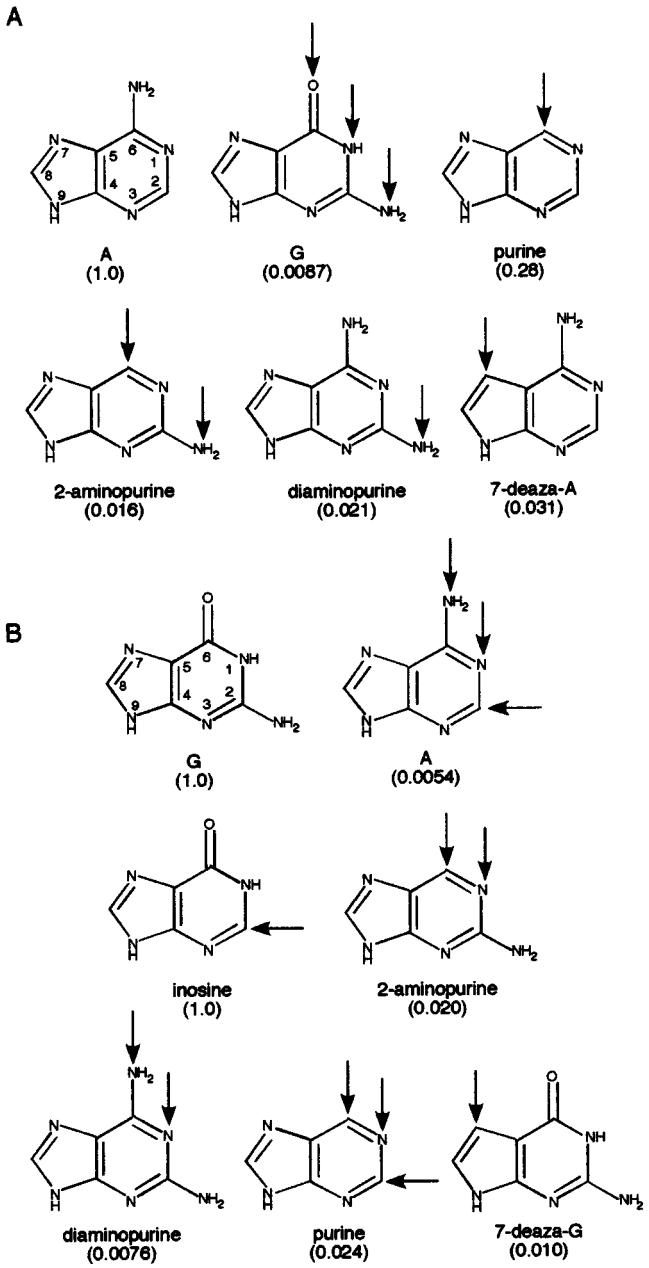

FIGURE 3: Structures of purine bases substituted at positions 73 (panel A) and 72 (panel B) of a semi-synthetic *E. coli* tRNA^{Pro} construct. The arrows indicate atomic group changes relative to that of the wild-type base. The numbers in parentheses are k_{cat}/K_M, given relative to that of wild-type semi-synthetic *E. coli* tRNA^{Pro}, which was set at 1.0. The 2′-hydroxyl group was previously shown to be dispensable at position 72 (*47*) and position 73 (this work and data not shown). We therefore incorporated the commercially available deoxynucleotide version of the base analogues purine, diaminopurine, 2-aminopurine, 7-deaza-A, and 7-deaza-G. In these assays, the tRNA concentration ranged from 1 to 20 μM, and the ProRS concentration was 25 or 50 nM. Each assay was carried out at least three times with an average standard deviation of ±33%.

in aminoacylation. Thus, in addition to a strong preference for A over G, the major groove N7 nitrogen contributes to positive recognition by ProRS.

A preference for G over A was previously observed at position 72 of *E. coli* tRNA^{Pro} (*11*). The minor groove amino group of G72 was shown to be dispensable for aminoacylation, as a G to inosine substitution was well tolerated (*31*). In contrast, deletion of major groove elements by substitution with 2′-deoxy-2-aminopurine resulted in a substantial de-

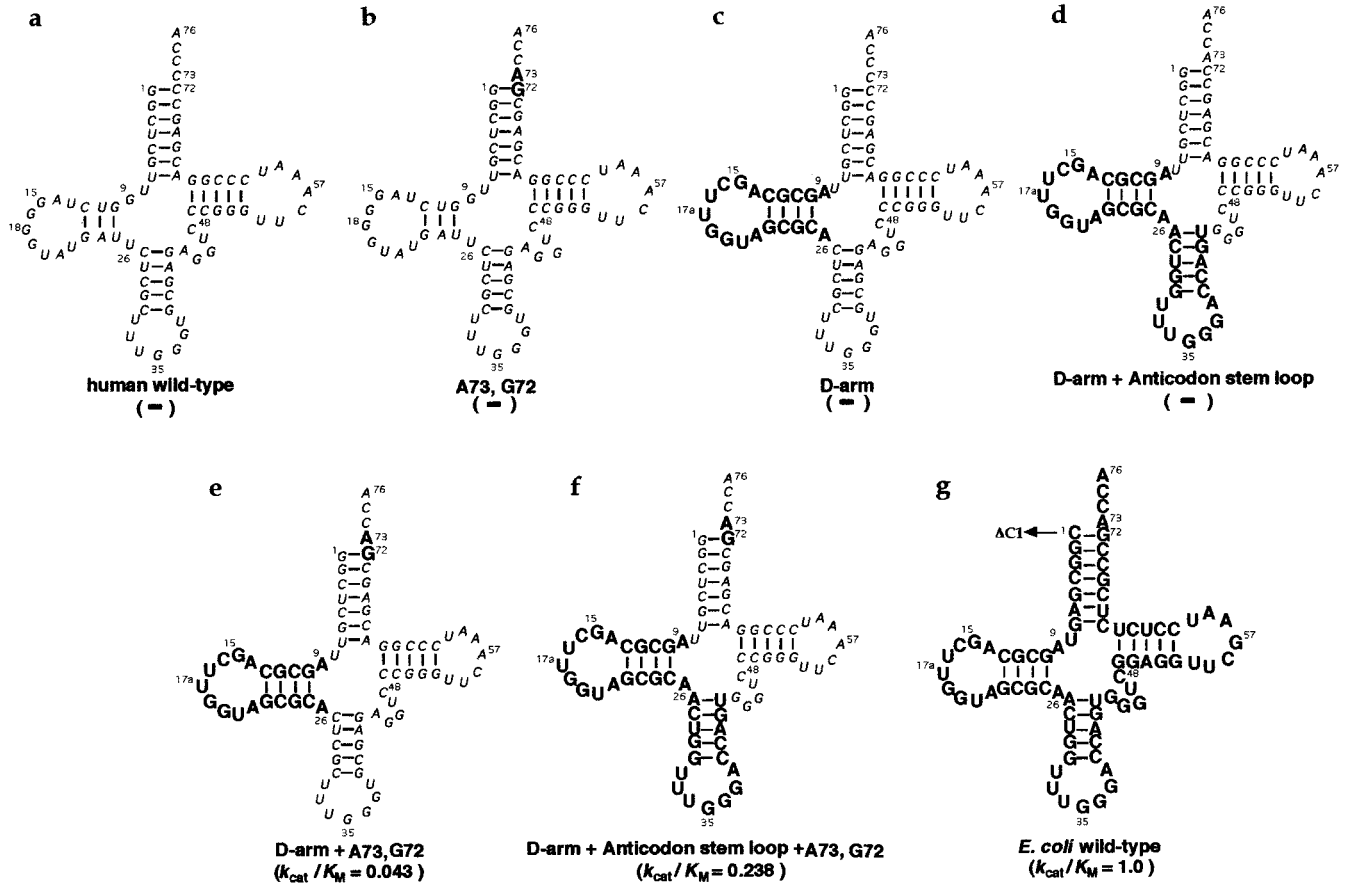

FIGURE 4: Wild-type human (a) and *E. coli* (g) $tRNA^{Pro}$ and chimeric (b−f) human and *E. coli* $tRNA^{Pro}$ variants tested as substrates for *E. coli* ProRS. Human sequence elements are in a smaller font, while *E. coli* sequence elements are in a larger boldface font. The arrow pointing to ΔC1 (g) indicates that in this transcript the first C was deleted to facilitate in vitro transcription as described previously (*25*). k_{cat}/K_M is given relative to that of the ΔC1 wild-type *E. coli* $tRNA^{Pro}$ transcript, which was set at 1.0. The notation (−) below the structure indicates that no aminoacylation was observed. In these assays, the tRNA concentrations ranged from 2 to 20 μM, and the ProRS concentration was 25 or 50 nM. Each assay was carried out at least three times with an average standard deviation of ±29%.

crease (50-fold) in aminoacylation efficiency (*31*). We have now substituted three additional purine analogues at this site, and our results are consistent with the importance of major groove elements. For example, substitution of the 6-keto group of G with an amino group (2′-deoxydiaminopurine) results in a 131-fold decrease in activity (Figure 3B), while deletion of the 6-keto group (2′-deoxypurine) results in a 42-fold decrease. Moreover, substitution of N7 with carbon (2′-deoxy-7-deaza-G) has an even larger effect at position 72 than it had at position 73, resulting in a 100-fold decrease in activity. These results strongly support the conclusion that base-specific major groove interactions at the top of the acceptor stem are critical for aminoacylation by *E. coli* ProRS.

Cross-Species Aminoacylation by E. coli ProRS. In contrast to the *E. coli* system, base-specific acceptor stem recognition does not appear to play a critical role in aminoacylation of $tRNA^{Pro}$ by human ProRS (*10*). Human $tRNA^{Pro}$ contains C73 and C72 and is not aminoacylated by *E. coli* ProRS, as expected. By transplanting *E. coli* $tRNA^{Pro}$ elements into the framework of human $tRNA^{Pro}$, we attempted to prepare a $tRNA^{Pro}$ chimeric construct that can be efficiently cross-acylated by the *E. coli* enzyme. Figure 4 shows the seven $tRNA^{Pro}$ constructs that were examined in this study. As mentioned above, the wild-type human $tRNA^{Pro}$ transcript is not aminoacylated by *E. coli* ProRS (Figure 4a). Transplantation of the two critical acceptor stem nucleotides A73 and G72 in the context of human $tRNA^{Pro}$ did not confer cross-species aminoacylation by *E. coli* ProRS (Figure 4b). The *E. coli* $tRNA^{Pro}$ D-domain contains an "extra" nucleotide at position 17a and four strong G•C base pairs in its stem (Figure 4g). In contrast, human $tRNA^{Pro}$ is missing a nucleotide at position 17 and contains a relatively weak three base pair D-stem (Figure 4a). Therefore, human $tRNA^{Pro}$ would be predicted to have a shorter D-anticodon helix than *E. coli* $tRNA^{Pro}$, and this may affect its capability to dock the acceptor stem correctly into the *E. coli* ProRS active site. To test this hypothesis, the D-domain of the human tRNA (nucleotides 9−26) was replaced with the corresponding sequence from *E. coli*. This D-arm swap alone did not confer cross-acylation (Figure 4c). However, when A73 and G72 were also transplanted into the D-arm construct, aminoacylation was observed (Figure 4e). Aminoacylation of this chimeric tRNA was reduced about 23-fold relative to that of wild-type *E. coli* $tRNA^{Pro}$. Additional stimulation of aminoacylation was observed upon the addition of the *E. coli* anticodon domain (Figure 4f). This chimera is now only ~4-fold less active than wild-type *E. coli* $tRNA^{Pro}$. It is worth noting that full-length C1-containing *E. coli* $tRNA^{Pro}$ is approximately 3-fold less active than the ΔC1-$tRNA^{Pro}$ transcript used in these studies (Figure 4g) (*25*). Thus, at least part of the decrease observed in the chimeric constructs shown in Figure 4e,f may be due to the presence of a nucleotide at position 1. Finally, transplantation

of the D- and anticodon domains alone is not sufficient to confer cross-acylation (Figure 4d), confirming the importance of the acceptor stem elements A73 and G72 for *E. coli* ProRS recognition.

DISCUSSION

Phylogenetic analyses previously established the existence of two very distinct groups of ProRS (*9, 10*). The prokaryotic-like group primarily includes sequences from bacteria and eukaryotic mitochondrial enzymes. The eukaryotic-like group contains sequences from eucarya and archaea, in addition to several sequences from bacteria. Using representative enzymes from each of these distinct groups (*E. coli* and human), we previously demonstrated species-specific differences in acceptor stem recognition of $tRNA^{Pro}$ (*10, 11*). In particular, the *E. coli* enzyme depends strongly on base-specific recognition of G72 and A73, two residues that are also strictly conserved in bacteria. In contrast, despite the fact that C72 and C73 are strictly conserved among cytoplasmic $tRNA^{Pro}$ sequences in eucarya, the human enzyme does not recognize acceptor stem nucleotides in a base-specific fashion. We hypothesized that the two groups of ProRS may correspond to two types of acceptor stem recognition and that there are likely to be sequence elements in the corresponding synthetases that have coadapted to reflect these changes through evolution (*10*). In the present study, we begin to address this latter question by focusing on the role of the variable loop sequence that connects the two β-strands of motif 2 in both *E. coli* and human ProRS. X-ray crystallography studies of class II synthetases complexed with tRNAs have shown that the motif 2 loop interacts closely with the acceptor stem domain (*8, 12*−*15, 32*). Thus, we imagined that species-specific differences in acceptor stem recognition might be reflected in differences in functional interactions with the motif 2 loop.

The sequence alignments of 11 prokaryotic-like and 9 eukaryotic-like ProRS sequences available when we began this work indicated that the motif 2 loop sequence was fairly well conserved within each group but differed significantly between the two groups (*10*). Figure 1 shows the alignment of the motif 2 loop of the *E. coli* and human enzymes studied here. The results of our site-directed mutagenesis studies (Table 1) indicate that the human motif 2 loop residues are not critical for tRNA aminoacylation. A new sequence alignment was recently carried out for the 20 known eukaryotic-like ProRS sequences currently available in the database (Kiyotaka Shiba, personal communication). Although there is not a single strictly conserved residue within the motif 2 loop, the Lys residue of the human FKHPQ motif is the most highly conserved amino acid, appearing in 17 of 20 sequences. The neighboring His residue is the next most conserved amino acid, appearing in 8 of 20 sequences. Mutation of these residues to Ala did not significantly affect aminoacylation efficiency by human ProRS (Table 1), and there is no evidence that this enzyme makes base-specific acceptor stem contacts (*10*). In the crystal structure of class II *T. thermophilus* phenylalanine-tRNA synthetase complexed to $tRNA^{Phe}$, motif 2 loop residues interact with the phosphate backbone of nucleotides 72−75 and base-specific acceptor stem contacts are not observed (*14*). Perhaps backbone recognition is more important than base-specific recognition in the case of human ProRS. Specific phosphorothioate substitutions will allow us to test the hypothesis. Although the basic residues K1084 and H1085 are good candidates for interacting with phosphate oxygens of the human $tRNA^{Pro}$ acceptor stem, our mutagenesis data suggest that main chain amino groups are more likely to be involved in backbone recognition.

In contrast to the results obtained with the human enzyme, mutagenesis of the *E. coli* ProRS motif 2 loop residues P145 and R146 resulted in large decreases in aminoacylation efficiency (21- and 79-fold, respectively) with less severe effects observed upon mutagenesis of V143 or F147 (3- and 14-fold, respectively). Our data indicate that R144 is absolutely essential for tRNA aminoacylation (Table 1) but is dispensable for prolyl-adenylate formation. A very similar result was obtained upon mutagenesis of S329 within the motif 2 loop of yeast AspRS (*15*). This residue aligns with R144 of *E. coli* ProRS (Figure 1, small arrow). The hydroxyl side chain of S329 interacts directly with the N1 of G73 of yeast $tRNA^{Asp}$ (*12*). The RP motif is present within the motif 2 loop sequence of all 18 prokaryotic-like cytoplasmic ProRSs sequenced to date (Kiyotaka Shiba, personal communication). The second Arg of the *E. coli* VRPRF motif is also highly conserved. It is present in 17 of 18 sequences, while the final Phe residue is present in 15 of 18 known sequences. On the other hand, the Val at position 143 is strictly conserved in only 5 of 18 sequences. Thus, the minor role of V143 and the relative importance of the RPR motif for aminoacylation by *E. coli* ProRS (Table 1) are in excellent agreement with predictions from sequence alignments.

On the basis of these data, we hypothesize that the RPR motif within the motif 2 loop sequence of prokaryotic-like ProRSs makes base-specific contacts with the acceptor stem of cognate $tRNA^{Pro}$. To gain further support for this hypothesis, we performed a cross-linking study wherein a 5′-phosphorothioate-containing *E. coli* $tRNA^{Pro}$ was specifically cross-linked to residue 143 of the motif 2 loop of *E. coli* ProRS. The cross-linking study confirmed that the critical R144 residue within the motif 2 loop was indeed proximal to the first (1:72) base pair of the tRNA acceptor stem (Figure 2). An atomic group mutagenesis analysis allowed us to further define the specific acceptor stem interactions. Purine analogues substituted into positions 72 and 73 of semi-synthetic $tRNA^{Pro}$ constructs revealed the importance of major groove functional groups at these sites (Figure 3). In particular, we find that the major groove elements N7 and O6 of G72 and N7 of A73 contribute significantly to aminoacylation catalytic efficiency. These results are in accordance with the major groove approach of class II synthetases to the top of the tRNA acceptor stem (*8*).

Arginine residues are commonly involved in specific RNA and DNA recognition. For example, the zinc-finger domain of the eukaryotic transcription factor Zif 268 contains several Arg residues that specifically bind to Gs of its target DNA (*33*). HIV-1 and HIV-2 tat proteins, as well as the RNA-binding domain of the small nuclear ribonucleoprotein U1A, have also been shown to recognize their target RNAs via Arg−G contacts (*34*−*36*). In the nucleotide binding regions of these proteins, an Arg side chain makes specific contacts

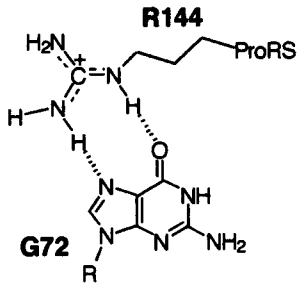

FIGURE 5: Proposed hydrogen bonding interaction between the critical motif 2 loop R144 of *E. coli* ProRS and the major groove of G72 in the acceptor stem of *E. coli* $tRNA^{Pro}$.

with the 6-keto oxygen and N7 of a G residue in the RNA. This type of major groove interaction has also been observed in the cocrystal structure of class II *E. coli* AspRS complexed to its cognate tRNA (*15*). In this case, the discriminator base G73 is contacted by both an Asp and an Arg residue from the motif 2 loop. Specifically, the N7 and O6 of G73 interact with Nη1 and Nη2 of the motif 2 loop Arg222. Similarly, our results support a model wherein one or both *E. coli* ProRS motif 2 loop Arg residues interact with major groove functional groups of G72 (Figure 5) and/or A73.

Species-specific differences in acceptor stem discrimination have also been observed in the case of other synthetase systems (*15*, *37*−*44*). For example, in the case of the class I tyrosyl-tRNA synthetase system, the human enzyme prefers a C1•G72-containing tRNA, whereas the bacterial enzyme requires a more conventional G1•C72 base pair (*40*). This system provides another striking example of how synthetases and tRNAs have coadapted through evolution, as a simple 39 amino acid peptide swap between enzymes was sufficient to switch recognition of the first base pair (*45*). However, for class II synthetases, it appears that additional regions outside the motif 2 loop contribute to acceptor stem selectivity. Indeed, the X-ray crystal structure of class IIb *E. coli* AspRS shows that in addition to the motif 2 loop, several additional loops and helices are in contact with the tRNA acceptor stem (*15*). As another example, class IIa *E. coli* threonyl-tRNA synthetase makes substantial contacts with its tRNA acceptor stem via both the catalytic domain and an N-terminal module designated as N2 (*32*). Thus, it is not too surprising that altering the motif 2 loop sequence alone in the class IIb lysine system (*46*) did not change acceptor stem base selectivity.

Our data obtained with the *E. coli* ProRS system support the presence of a specific hydrogen bonding interaction as illustrated in Figure 5. Moreover, our experiments with chimeric $tRNA^{Pro}$ constructs show that the *E. coli* enzyme can recognize the G72 and A73 bases in the context of a heterologous tRNA acceptor stem, as long as the appropriate D−anticodon domain interactions are also present (Figure 4). The phylogenetic analyses suggest that this interaction between arginine and the tRNA acceptor stem is likely to be conserved among all prokaryotic-like ProRSs. Our experiments also show that such a base-specific interaction with the acceptor stem is missing in the eukaryotic-like group of ProRSs. Experimental and phylogenetic data suggest that if contacts between the human $tRNA^{Pro}$ acceptor stem and the motif 2 loop of human ProRS indeed occur, they are likely to involve the RNA backbone and protein main chain amino groups. Taken together, these data provide additional support for species-specific differences in ProRS-$tRNA^{Pro}$ interactions and illustrate how synthetase active sites have coadapted to changes in tRNA acceptor stems through evolution.

ACKNOWLEDGMENT

We thank Dr. Kiyotaka Shiba for providing an updated sequence alignment of ProRSs and for helpful discussion and comments on the manuscript. We thank Dr. Penny J. Beuning for preparation of synthetic RNA oligonucleotides used in the atomic mutagenesis study and for critical reading of the manuscript.

REFERENCES

1. Hou, Y.-M. (1997) *Chem. Biol. 4*, 93−96.
2. Giegé, R., Sissler, M., and Florentz, C. (1998) *Nucleic Acids Res. 26*, 5017−5035.
3. Musier-Forsyth, K., and Schimmel, P. (1999) *Acc. Chem. Res. 32*, 368−375.
4. Eriani, G., Delarue, M., Poch, O., Gangloff, J., and Moras, D. (1990) *Nature 347*, 203−206.
5. Cusack, S., Härtlein, M., and Leberman, R. (1991) *Nucleic Acids Res. 19*, 3489−3498.
6. Rossmann, M. G., Moras, D., and Olsen, K. W. (1974) *Nature 250*, 194−199.
7. Cusack, S., Berthet-Colominas, C., Härtlein, M., Nassar, N., and Leberman, R. (1990) *Nature 347*, 249−255.
8. 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.
9. Cusack, S., Yaremchuk, A., Krikliviy, I., and Tukalo, M. (1998) *Structure 6*, 101−108.
10. Stehlin, C., Burke, B., Yang, F., Liu, H., Shiba, K., and Musier-Forsyth, K. (1998) *Biochemistry 37*, 8605−8613.
11. Liu, H., Peterson, R., Kessler, J., and Musier-Forsyth, K. (1995) *Nucleic Acids Res. 23*, 165−169.
12. Cavarelli, J., Rees, B., Ruff, M., Thierry, J. C., and Moras, D. (1993) *Nature 362*, 181−184.
13. Cusack, S., Yaremchuk, A., and Tukalo, M. (1996) *EMBO J. 15*, 2834−2842.
14. Goldgur, Y., Mosyak, L., Reshetnikova, L., Ankilova, V., Lavrik, O., Khodyreva, S., and Safro, M. (1997) *Structure 5*, 59−68.
15. Eiler, S., Dock-Bregeon, A., Moulinier, L., Thierry, J. C., and Moras, D. (1999) *EMBO J. 18*, 6532−6541.
16. Beuning, P. J., Gulotta, M., and Musier-Forsyth, K. (1997) *J. Am. Chem. Soc. 119*, 8397−8402.
17. Yap, L. P., Stehlin, C., and Musier-Forsyth, K. (1995) *Chem. Biol. 2*, 661−666.
18. Heacock, D. H., 2nd, Forsyth, C. J., Shiba, K., and Musier-Forsyth, K. (1996) *Bioorg. Chem. 24*, 273−289.
19. Stehlin, C., Heacock, D. H., 2nd, Liu, H., and Musier-Forsyth, K. (1997) *Biochemistry 36*, 2932−2938.
20. Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 488−492.
21. Horton, R. M., and Pease, L. R. (1991) in *Directed Mutagenesis* (McPherson, M. J., Ed.) pp 217−247, IRL Press, New York.
22. Fersht, A. R., Ashford, J. S., Bruton, C. J., Jakes, R., Koch, G. L., and Hartley, B. S. (1975) *Biochemistry 14*, 1−4.
23. Calendar, R., and Berg, P. (1966) *Biochemistry 5*, 1690−1695.
24. Hill, K., and Schimmel, P. (1989) *Biochemistry 28*, 2577−2586.
25. Liu, H., and Musier-Forsyth, K. (1994) *Biochemistry 33*, 12708−12714.
26. Sampson, J. R., and Uhlenbeck, O. C. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 1033−1037.
27. Cavarelli, J., Eriani, G., Rees, B., Ruff, M., Boeglin, M., Mitschler, A., Martin, F., Gangloff, J., Thierry, J. C., and Moras, D. (1994) *EMBO J. 13*, 327−337.
28. Eriani, G., and Gangloff, J. (1999) *J. Mol. Biol. 291*, 761−773.

29. McClain, W. H., Schneider, J., and Gabriel, K. (1994) *Nucleic Acids Res. 22*, 522−529.
30. Crothers, D. M., Seno, T., and Söll, D. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 3063−3067.
31. Liu, H., Yap, L.-P., and Musier-Forsyth, K. (1996) *J. Am. Chem. Soc. 118*, 2523−2524.
32. Sankaranarayanan, R., Dock-Bregeon, A. C., Romby, P., Caillet, J., Springer, M., Rees, B., Ehresmann, C., Ehresmann, B., and Moras, D. (1999) *Cell 97*, 371−381.
33. Pavletich, N. P., and Pabo, C. O. (1991) *Science 252*, 809−817.
34. Puglisi, J. D., Tan, R., Calnan, B. J., Frankel, A. D., and Williamson, J. R. (1992) *Science 257*, 76−80.
35. Oubridge, C., Ito, N., Evans, P. R., Teo, C.-H., and Nagai, K. (1994) *Nature 372*, 432−438.
36. Brodsky, A. S., and Williamson, J. R. (1997) *J. Mol. Biol. 267*, 624−639.
37. Lee, C. P., and RajBhandary, U. L. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 11378−11382.
38. Shiba, K., Schimmel, P., Motegi, H., and Noda, T. (1994) *J. Biol. Chem. 269*, 30049−30055.
39. Hipps, D., Shiba, K., Henderson, B., and Schimmel, P. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 5550−5552.
40. Quinn, C. L., Tao, N., and Schimmel, P. (1995) *Biochemistry 34*, 12489−12495.
41. Becker, H. D., Giegé, R., and Kern, D. (1996) *Biochemistry 35*, 7447−7458.
42. Nameki, N., Tamura, K., Asahara, H., and Hasegawa, K. (1997) *J. Mol. Biol. 268*, 640−647.
43. Shiba, K., Stello, T., Motegi, H., Noda, T., Musier-Forsyth, K., and Schimmel, P. (1997) *J. Biol. Chem. 272*, 22809−22816.
44. Mazauric, M. H., Keith, G., Logan, D., Kreutzer, R., Giegé, R., and Kern, D. (1998) *Eur. J. Biochem. 251*, 744−757.
45. Wakasugi, K., Quinn, C. L., Tao, N., and Schimmel, P. (1998) *EMBO J. 17*, 297−305.
46. Steer, B. A., and Schimmel, P. (1999) *Biochemistry 38*, 4965−4971.
47. Yap, L.-P., and Musier-Forsyth, K. (1995) *RNA 1*, 418−424.

BI001835P