Pergamon

Bioorganic & Medicinal Chemistry, Vol. 5, No. 6, pp 1011–1019, 1997

PII: S0968-0896(97)00039-4
0968-0896/97 $17.00 + 0.00

A Strategy of tRNA Recognition that Includes Determinants of RNA Structure

Christian S. Hamann and Ya-Ming Hou*

Department of Biochemistry and Molecular Pharmacology, Thomas Jefferson University, 233 South 10th Street, Philadelphia, PA 19107, U.S.A.

Abstract—Recognition of tRNAs by aminoacyl tRNA synthetases establishes the connection between amino acids and anticodon triplets of the genetic code. Although anticodons and nucleotides adjacent to the amino acid attachment site are generally important, the tertiary structural framework of tRNAs has recently been implicated to have a role in tRNA recognition. A G15:G48 tertiary hydrogen base pair of *E. coli* $tRNA^{Cys}$ is important for recognition of the tRNA by cysteine tRNA synthetase. This base pair is proposed to consist of N2:N3, rather than N1:O6, hydrogen bonds. The reproduction of the hydrogen pairing scheme of G15:G48 and the structural features related to this pairing has been recently achieved in the sequence framework of *E. coli* $tRNA^{Gly}$. This reproduction required an A13:A22 mismatch in the dihyrouridine stem. To determine if A13:A22 is a determinant of the structural features of G15:G48, we investigated the A15:U48 and A15:A48 variants of $tRNA^{Gly}$ which harbored specific substitutions of A13:A22. We show here that introduction of A13:A22 to both tRNA frameworks confers structural features similar to those of G15:G48 in *E. coli* $tRNA^{Cys}$. These structural features are accompanied by efficient recognition of both tRNAs by cysteine tRNA synthetase. Substitution of A13:A22 with U13:A22 alters the structural features at 15:48 and impairs tRNA recognition. The dependence on A13:A22 for tRNA recognition has a distinct similarity to that of *E. coli* $tRNA^{Cys}$ and to that of the G15:G48 variant of $tRNA^{Gly}$. The results have implications for the design and manipulation of RNA structural elements as the basis for tRNA recognition. © 1997 Elsevier Science Ltd.

Introduction

Recognition of tRNAs by aminoacyl tRNA synthetases establishes specific aminoacylation of tRNAs. This recognition provides the connection between amino acids and the trinucleotide sequences of the genetic code. While the 20 aminoacyl tRNA synthetases are divided into two classes based on two distinct tertiary structures,[1,2] all tRNAs are folded into an L.[3–5] In this L, long range tertiary hydrogen bonds are formed between conserved and semi-conserved nucleotides to connect distal parts of the tRNA molecule so as to place the acceptor stem and the anticodon at the opposite ends of the L. The most common strategy of tRNA recognition uses nucleotides either in the acceptor stem or in the anticodon, or both, as determinants for the specific interaction between a tRNA and its cognate synthetase.[6–9]

Determinants are defined as nucleotides in each tRNA that have the ability to determine the specificity of aminoacylation. Biochemical studies show that substitutions of nucleotides in the acceptor stem and in the anticodon have a deleterious effect on aminoacylation of most tRNAs.[6–9] Transfer of these nucleotides from one tRNA to another confers the latter with the aminoacylation specificity of the former. The ability of these nucleotides to establish aminoacylation specificity in a different tRNA sequence framework is largely due to the conservation of the L. Crystallographic analyses of several tRNA-synthetase complexes show that, within the constraint of the L, the presentation of the nucleotides in the acceptor stem and in the anticodon to synthetases is relatively conserved, and that specific functional groups of these nucleotides establish distinct RNA-protein contacts that confer the specificity of aminoacylation.[10–12] In this strategy of tRNA recognition, the conserved and semi-conserved nucleotides maintain the framework of the L and share structural features that are common to most tRNAs.

A variation of the common recognition strategy is to include unusual structural features of the L as determinants for tRNA recognition. We and others previously showed that U73 and the GCA anticodon of *E. coli* $tRNA^{Cys}$ are important for aminoacylation (Fig. 1).[13–16] In addition, we showed that an unusual G15:G48 tertiary hydrogen base pair unique to *E. coli* $tRNA^{Cys}$ is an important structural element that contributes to aminoacylation (Fig. 1).[14,17] The tertiary base pair between nucleotides at positions 15 and 48 connects the dihydrouridine (D) and variable loops of the tRNA to stabilize the folding of the L. The system of recognition of *E. coli* $tRNA^{Cys}$ by cysteine tRNA synthetase thus may include determinants of the L shaped structure.

Inspection of tRNA sequences shows that most tRNAs contain a pu15:py48 base pair.[18] The G15:C48 or A15:U48 base pair consists of two hydrogen bonds that involve N1 and O6 of G15, or N1 and N6 of A15.[3–5] In these pairings, G15 or A15 directly stacks on A14 and is

*To whom correspondence should be addressed. Telephone: 215-503-4480; Fax: 215-923-9162; E-mail: hou1@jeflin.tju.edu.

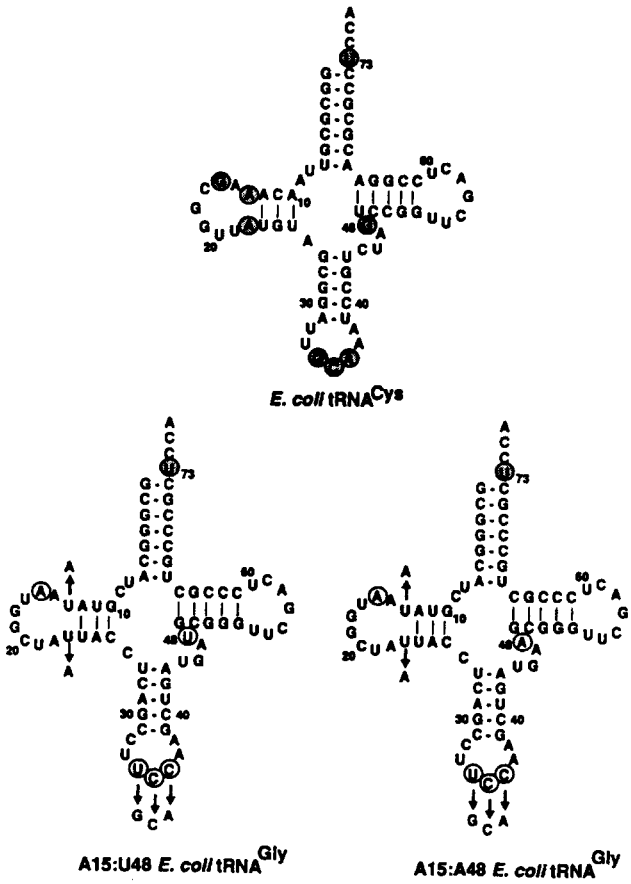


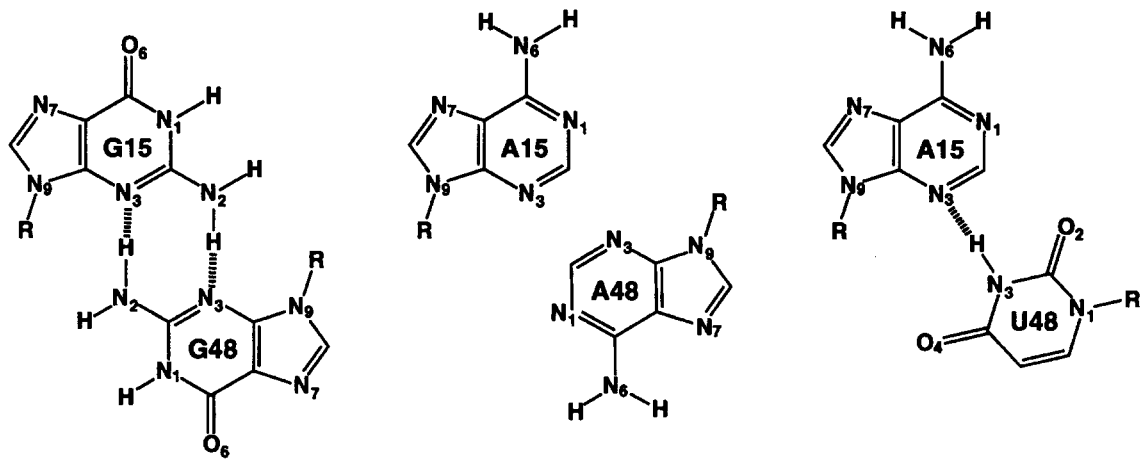
Figure 1. Sequence and cloverleaf structure of *E. coli* $tRNA^{Cys}$,[18] and of the A15:U48 and A14:A48 frameworks of *E. coli* $tRNA^{Gly}$. The latter two frameworks were derived from the sequence of *E. coli* $tRNA^{Gly/UCC}$.[18] Substitution of the UCC anticodon with GCA created the A15:U48 framework while additional substitution of U48 with A48 created the A15:A48 framework. The nucleotides A15:U48 and A15:A48 in the two frameworks are indicated by open circles while introduction of A13:A22 to replace the wild-type U13:U22 is indicated by arrows. In *E. coli* $tRNA^{Cys}$, the U73 nucleotide, the GCA anticodon, the G15:G48 tertiary base pair, and the A13:A22 mismatch that are important for aminoacylation with cysteine are indicated by shaded circles. U73 is shared in common between $tRNA^{Cys}$ and $tRNA^{Gly}$.

protected from chemical modification at the N7 position.[19,20] The G15:C48 and A15:U48 hydrogen pairings are superimposable and, in most cases, can accommodate substitutions with only a small effect on aminoacylation.[21,22] In contrast, alteration of G15:G48 to G15:C48 in *E. coli* $tRNA^{Cys}$ impairs aminoacylation.[14,17] The unusual structural features of G15:G48 include the accessibility of G15 and G48 to kethoxal (a chemical that attacks the N1 and N2 of guanine), and the accessibility of the N7 of G15 to dimethyl sulfate (DMS).[14,17] Based on these features, we proposed that the G15:G48 base pairing is mediated through the N2:N3 positions (Fig. 2). This pairing not only leaves N1 and N2 available for reaction with kethoxal, but also is predicted to swing G15 away from A14 to render N7 of G15 reactive to DMS. Although the N2:N3 pairing of G15:G48 was documented in a DNA dodecamer,[23] it has not been found in any RNAs of known structure.

In *E. coli* $tRNA^{Cys}$, while the structural features related to the N2:N3 pairing of G15:G48 are sensitive to substitutions of G15:G48 itself, they are also sensitive to substitutions of an A13:A22 mismatch in the D stem. In fact, in *E. coli* $tRNA^{Cys}$, substitution of A13:A22 can alter the hydrogen bonding of the G15:G48 base pair. These substitutions decrease the catalytic efficiency of aminoacylation (k_{cat}/K_m) and thus provide the basis for a structure–function correlation between G15:G48 and A13:A22 for aminoacylation of *E. coli* $tRNA^{Cys}$.[17,24]

Additional experiments implicated A13:A22 as a structural determinant for the N2:N3 pairing of G15:G48. In these experiments, we transferred G15:G48 into an unrelated tRNA sequence framework and showed that structural and functional features related to the specific base pairing of G15:G48 in *E. coli* $tRNA^{Cys}$ were recapitulated, provide that A13:A22 was present in the tRNA.[24] We chose the sequence framework of *E. coli* $tRNA^{Gly/UCC}$ because it already has U73 (Fig. 1). This tRNA (Gly01, Table 1) was not a substrate for the cysteine enzyme and its k_{cat}/K_m of aminoacylation with cysteine was four orders of magnitude below that of *E. coli* $tRNA^{Cys}$ (Cys01, Table 1). Conversion of the UCC anticodon to GCA created Gly06 which had significantly improved cysteine acceptance (relative k_{cat}^{+}/K_m = 0.30, Table 1). Into Gly06 we introduced G15:G48 to replace the wild-type A15:U48 and created variations at 13:22 that included U13:U22, A13:U22, U13:A22, and A13:A22. To each 13:22 variant, we generated an A21 and a U21 mutant to examine the significance of the unusual U21 found in *E. coli* $tRNA^{Cys}$. We created a total of eight variants of Gly06. Among these variants, those containing A13:A22, regardless of A21 or U21, successfully reproduced the chemical accessibilities of G15:G48 as those in *E. coli* $tRNA^{Cys}$. These variants were also fully competent for efficient aminoacylation with cysteine by cysteine tRNA synthetase.[24]

One question that arises from these studies is whether A13:A22 specifically alters the structure of G15:G48, or whether A13:A22 is responsible for structural features of G15:G48 that may be mimicked by other nucleotides. To address this question, we constructed eight variations of 13:22/21 in the A15:U48 and A15:A48 frameworks of $tRNA^{Gly}$ and examined if these variations recapitulated the structure–function relationship as that in the G15:G48 framework. We used chemical probes and kinetics of tRNA aminoacylation to investigate these variants. We prepared all tRNAs as T7 transcripts. Although these transcripts lacked modified bases, we previously showed that the lack of modification did not interfere with the ability of the cysteine enzyme to kinetically discriminate among tRNA variants.[14,17] The results show that both A15:U48 and A15:A48 have the ability to create the functional and structural features of G15:G48, and that this ability is dependent on A13:A22. In addition, mutants that contain U13:A22 alter the structural features at 15:48 and are defective for aminoacylation. The structure–



H-bond p15	**kethoxal +**	**DMS +**	**DMS +**
H-bond p48	**kethoxal +**	**DMS +**	**CMCT −**
stacking p15	**DMS +**	**DEPC +**	**DEPC +**

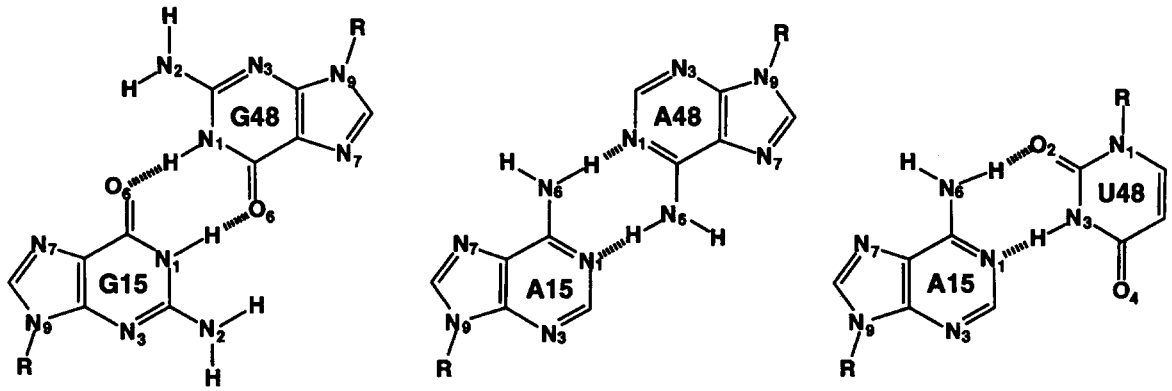
Figure 2. The proposed N2:N3 pairing of G15:G48 in *E. coli* $tRNA^{Cys}$, and the analogous 'pairing' for A15:A48 and A15:U48 in $tRNA^{Gly}$. Potential hydrogen bonds are indicated by dashed lines. Predictions of hydrogen donors or acceptors for their chemical accessibilities to kethoxal (N1 and N2 of G), DMS (N1 of A), and CMCT (N3 of U) are indicated, as are predictions of the N7 accessibility at position 15 to modification by DMS (N7 of G) or DEPC (N7 of A). The '+' symbol indicates predicted accessibility while '−' indicates predicted inaccessiblity. Predictions that were supported by experimental results are bold-faced while those that were not supported are plain.

function correlation between 15:48 and 13:22 is strikingly consistent with that of *E. coli* $tRNA^{Cys}$ and with that of the G15:G48 framework of $tRNA^{Gly}$. Together, these four frameworks provide evidence that RNA structural features can contribute to tRNA recognition.

Results

Kinetic analysis of tRNA mutants

Gly06, which contained the wild-type A15:U48 and A21 of $tRNA^{Gly}$, was the starting material for the A15:U48



H-bond p15	**kethoxal −**	**DMS −**	**DMS −**
H-bond p48	**kethoxal −**	**DMS −**	**CMCT −**
stacking p15	**DMS −**	**DEPC −**	**DEPC −**

Figure 3. The proposed N1:O6 pairing of G15:G48 in *E. coli* $tRNA^{Cys}$, and the analogous 'pairing' for A15:A48 and A15:U48 in $tRNA^{Gly}$. Potential hydrogen bonds are indicated by dashed lines. Predictions of hydrogen donors or acceptors for their chemical accessibilities to kethoxal (N1 and N2 of G), DMS (N1 of A), and CMCT (N3 of U) are indicated, so are predictions of the N7 accessibility at position 15 to methylation by DMS (N7 of G) or DEPC (N7 of A). The '+' symbol indicates predicted accessibility while '−' indicates predicted inaccessibility. Predictions that were supported by experimental results are bold-faced while those that were not supported are plain.

Table 1. Kinetic parameters for variants of A15:U48 $tRNA^{Gly/GCA}$

	Name	Nucleotide position			k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($M^{-1}s^{-1}$)	Relative to Cys01
		13	22	21				
G15:G48	Cys01	A	A	U	0.470 ± 0.090	2.24 ± 1.20	210,000	1.00
A15:U48	Gly01	U	U	A				3.0×10^{-5}
A15:U48	Gly06	U	U	A	0.944 ± 0.480	14.8 ± 8.1	63,500	0.30
	Gly07	U	U	U	0.725 ± 0.146	18.7 ± 4.8	38,800	0.18
	Gly15	A	U	A	0.378 ± 0.099	16.9 ± 6.7	22,400	0.11
	Gly30	A	U	U	0.415 ± 0.103	28.3 ± 3.3	14,700	0.070
	Gly16	U	A	A	0.177 ± 0.030	26.5 ± 3.2	6680	0.032
	Gly31	U	A	U	1.15 ± 0.36	13.4 ± 4.1	85,800	0.41
	Gly13	A	A	A	1.07 ± 0.34	3.62 ± 0.83	296,000	1.41
	Gly26	A	A	U	0.96 ± 0.67	5.14 ± 1.01	190,000	0.90

Kinetic parameters for the wild-type *E. coli* $tRNA^{Cys}$ (Cys01), and wild-type $tRNA^{Gly/UCC}$ (Gly01) are shown as a reference. Subsitution of the UCC anticodon in Gly01 with GCA created the starting A15:U48 framework (Gly06). Including Gly06, eight variations at 13:22/21 in this framework were created and the kinetic parameters of these tRNAs were compared to those of the wild type $tRNA^{Cys}$. Parameters k_{cat} and K_m were derived from the Michaelis–Menten equation of velocity vs. substrate concentration using the IGOR-Pro curve-fitting program (WaveMetrics, Inc., Lake Oswego, OR). Each parameter was the average of at least three independent determinations. Mutants (with A21 or U21) are arranged in the table according to the order of U13:U22, A13:U22, U13:A22, and A13:A22. Gly13 and Gly16, the most and least active mutants for aminoacylation with cysteine, respectively, are highlighted by shading.

framework (Fig. 1). Gly06 was a substrate for the cysteine enzyme, but its k_{cat}/K_m for aminoacylation with cysteine is 0.30 relative to that of $tRNA^{Cys}$. Including Gly06, eight tRNAs with variations at 13:22/21 were created by site-directed mutagenesis and were kinetically screened for the mutant that would have the wild-type activity for aminoacylation with cysteine. The results of kinetic analysis are shown in Table 1. In general, mutants containing U13:U22, A13:U22 or U13:A22 were defective for aminoacylation with cysteine, whereas mutants containing A13:A22 were fully active. The weakest mutant is Gly16, which contained U13:A22/A21 and had a k_{cat}/K_m of aminoacylation with cysteine 0.032 relative to that of $tRNA^{Cys}$. The most active mutant is Gly13, which contained A13:A22/A21 and had a k_{cat}/K_m as good as that of $tRNA^{Cys}$. Between Gly16 and Gly13, the single substitution of U13 with A13 improved k_{cat}/K_m by almost two orders of magnitude. In Gly13, substitution of A21 with U21 to create Gly26 had no major effect on aminoacylation with cysteine.

We next created the A15:A48 framework of $tRNA^{Gly}$ by introducing A48 to replace U48 (Fig. 1). The starting material, Gly08, had a k_{cat}/K_m for aminoacylation with cysteine 0.55 relative to that of $tRNA^{Cys}$ (Table 2). Including Gly08, eight variants with substitutions at 13:22/21 were determined for their kinetic parameters for aminoacylation with cysteine. The results are shown in Table 2. The behaviors of these variants are generally

Table 2. Kinetic parameters for variants of A15:A48 $tRNA^{Gly/GCA}$

	Name	Nucleotide position			k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($M^{-1}s^{-1}$)	Relative to Cys01
		13	22	21				
G15:G48	Cys01	A	A	U	0.470 ± 0.090	2.24 ± 1.20	210,000	1.00
A15:U48	Gly01	U	U	A				3.0×10^{-5}
A15:A48	Gly08	U	U	A	1.50 ± 0.26	13.1 ± 5.6	115,000	0.55
	Gly09	U	U	U	1.39 ± 0.43	18.6 ± 3.4	74,700	0.36
	Gly17	A	U	A	0.649 ± 0.189	15.6 ± 8.2	41,600	0.20
	Gly28	A	U	U	0.365 ± 0.082	29.1 ± 9.6	12,500	0.060
	Gly18	U	A	A	0.248 ± 0.105	25.0 ± 14.1	9920	0.047
	Gly29	U	A	U	0.769 ± 0.084	28.2 ± 5.7	27,300	0.13
	Gly14	A	A	A	1.06 ± 0.35	2.41 ± 0.71	440,000	2.1
	Gly27	A	A	U	1.05 ± 0.57	4.97 ± 1.20	211,000	1.0

Kinetic parameters for the wild-type *E. coli* $tRNA^{Cys}$ (Cys01), and wild-type $tRNA^{Gly/UCC}$ (Gly01) are shown as a reference. Subsitution of the UCC anticodon in Gly01 with GCA and substitution of U48 with A48 created the starting A15:A48 framework (Gly08). Including Gly08, eight variations at 13:22/21 in this framework were created and the kinetic parameters of these tRNAs were compared to those of the wild-type $tRNA^{Cys}$. Parameters k_{cat} and K_m were derived as described in the legend to Table 1. Gly14 and Gly18, the most and least active mutants for aminoacylation with cysteine, respectively, are highlighted by shading.

similar to those of the variants of the A15:U48 framework. While mutants containing U13:U22, A13:U22, and U13:A22 were defective for aminoacylation with cysteine, mutants containing A13:A22 were fully active. The weakest mutant is Gly18, which contained U13:A22/A21 and had a k_{cat}/K_m of aminoacylation with cysteine 0.047 relative to that of $tRNA^{Cys}$. The most active mutant is Gly14, which contained A13:A22/A21 and had a k_{cat}/K_m about the same as that of $tRNA^{Cys}$. Between Gly18 and Gly14, the single substitution of U13 with A13 elevated k_{cat}/K_m by almost two orders of magnitude. Again, substitution of A21 with U21 in Gly14 to create Gly27 had no major effect on aminoacylation.

Examination of kinetic parameters in Tables 1 and 2 shows that the effect of A13:A22 versus U13:A22 on aminoacylation is primarily on K_m. The most active mutant of the A15:U48 and of the A15:A48 frameworks (Gly13 and Gly14, respectively) maintain a K_m (3.6 and 2.4 μM, respectively) similar to that of $tRNA^{Cys}$ (2.2 μM), whereas the least active mutants (Gly16 and Gly18, respectively) have K_m values at least tenfold higher (26.5 and 25.0 μM, respectively). Because K_m reflects the binding of the tRNA to the synthetase, the most active mutants are likely to maintain the binding strength as $tRNA^{Cys}$ whereas the least active mutants harbor a defect in binding.

The most and least active mutants of the A15:U48 framework and of the A15:A48 framework of $tRNA^{Gly}$ share the same nucleotide substitutions as their respective counterparts of the G15:G48 framework.[24] In all cases, the most active mutant contains A13:A22/A21 whereas the least active mutant contains U13:A22/A21. This consistency suggests that all three frameworks of $tRNA^{Gly}$ may provide the same basis for discrimination by cysteine tRNA synthetase. Between the three frameworks of $tRNA^{Gly}$ and the framework of $tRNA^{Cys}$, however, the most active form of $tRNA^{Cys}$ is the wild-type, which contains A13:A22/U21. Substitution of U21 with A21 has a small effect on k_{cat}/K_m.[17] This indicates a small preference for U21 by $tRNA^{Cys}$. The least active mutant of $tRNA^{Cys}$ contains U13:A22/A21,[17] which is identical to that of the least active mutants of $tRNA^{Gly}$. Thus, except for the difference at position 21, nucleotide sequences for the most and least active variants of $tRNA^{Cys}$ and of the three frameworks of $tRNA^{Gly}$ are the same.

The quantitative effect on aminoacylation with cysteine by the most and least active mutants is also consistent among the four tRNA sequence frameworks. The A13:A22 containing mutants of the three frameworks of $tRNA^{Gly}$ all have comparable k_{cat}/K_m as that of $tRNA^{Cys}$. The U13:A22/A21 containg mutants of $tRNA^{Gly}$ are all defective for aminoacylation with cysteine by almost two orders of magnitude. The k_{cat}/K_m of the U13:A22/A21 containing mutant of $tRNA^{Cys}$ is 0.0067 relative to that of the wild-type, which is also in the range of two orders of magnitude.[17] Thus, within the precision of kinetic analysis, the most and least active mutants of the three frameworks of $tRNA^{Gly}$ quantitatively recreate the respective behaviors of those in $tRNA^{Cys}$.

Analysis of tRNA mutants with chemical probes

We previously used chemical probes to distinguish the most active from the least active mutants in the G15:G48 sequence framework of $tRNA^{Gly}$. While the most active mutants maintained the structural features at 15:48 as those in *E. coli* $tRNA^{Cys}$, the least active mutants had features reminiscent of those in the U13:A22 containing mutant of $tRNA^{Cys}$.[24] Briefly, the most active mutants in these two frameworks contained G15 and G48 accessible to kethoxal, and the N7 of G15 accessible to DMS.[14,17,24] In contrast, the least active mutants contained G15 and G48 inaccessible to kethoxal, and the N7 of G15 inaccessible to DMS.[14,17,24] The alteration of chemical accessibilities suggested a change in the hydrogen pairing scheme at 15:48. Based on structural modeling, we proposed an N1:O6 pairing for the least active mutants (Fig. 3). In this pairing, the N1 nitrogen of G15 and G48 served as a hydrogen donor for the base pair and therefore was inaccessible to kethoxal. Further, to form the N1:O6 pairing, the backbone of G15 in the N1:O6 pairing was rearranged relative to that of G48. This rearrangement placed G15 directly stacked with A14 and provided the basis for its resistance to DMS modification. In principle, the N1:O6 pairing of G15:G48 would align the base pair with that of the common G15:C48 or A15:U48 in most tRNAs.

With a similar probing strategy, we evaluated if the most and least active mutants of the A15:U48 and A15:A48 frameworks of $tRNA^{Gly}$ responded to chemical probes the same way as their counterparts in the G15:G48 framework. We used DMS to probe the hydrogen bonding potential of N1 of A15 and of A48 and used CMCT (1-cyclohexyl-3-(2-morpholinoethyl)-carbodiimide metho-*p*-toluene sulfonate) to probe N3 of U48.[25] Both modifications were mapped by primer extension with reverse transcriptase. Additionally, we used DEPC (diethyl pyrocarbonate) to probe the accessibility of N7 of A15. This modification was mapped by aniline scission followed by gel electrophoresis.[25] As a hypothesis, we proposed that the most active mutants would maintain the structural features of those in the N2:N3 pairing of G15:G48, and that the least active mutants would mimic those in the N1:O6 pairing of G15:G48.

The results of chemical probes generally supported the hypothesis. To interpret the results, we first modeled the base pairing of A15:U48 and of A15:A48 in the same backbone configuration as that of the N2:N3 base pairing of G15:G48 (Fig. 2). This modeling allowed us to determine if chemical modifications supported these base pairings. For the most active mutants, our modeling predicted one hydrogen bond between A15 and U48 and no apparent hydrogen bonds between A15

and A48. According to this model, the A15:U48 framework should have an accessible N1 of A15 but an inaccessible N3 of U48. This is partially supported by the lack of accessibility of U48 to CMCT in Gly13 (the most active mutant of the A15:U48 framework). In addition, our modeling predicted that both A15 and A48 in the A15:A48 framework contained an accessible N1. This is partially supported by the accessibility of A48 to DMS in Gly14 (the most active mutant of the A15:A48 framework). The strongest support for the model was the accessibility of DEPC to the N7 of A15 in both frameworks. This accessibility indicated that each A15 was in a distinctly different position from that of other tRNAs, where A15 was not modified.[19,20] To establish the validity of these results, we demonstrated in control experiments that nucleotides in the Watson–Crick base pairs were protected from modifications, whereas those in the anticodon loop were accessible (data not shown).

We performed similar experiments with the least active mutants and showed that the results were also supportive of our structural hypothesis. We modeled the base pairing of A15:U48 and A15:A48 in the same backbone configuration as that of the N1:O6 pairing of G15:G48. Our modeling predicted two hydrogen bonds for A15:U48 and two hydrogen bonds for A15:A48 that would involve the N1 and N6 of A15 in each base pair (Fig. 3). In this modeling, the A15:U48 framework should have an inaccessible N1 of A15 and an inaccessible N3 of U48. Both predictions were supported by the lack of accessibility of A15 to DMS and of U48 to CMCT in Gly16 (the least active mutant of the A15:U48 framework). In addition, our modeling predicted that A15 and A48 in the A15:A48 framework contained an inaccessible N1. This is partially supported by the lack of accessibility of A15 to DMS in Gly18 (the least active mutant of the A15:A48 framework). Again, the strongest support for the model was the protection of the N7 position of A15 from DEPC in both frameworks. This protection suggested that each A15 was involved in stacking interactions with A14, as observed for other tRNAs.[19,20]

We also investigated the structural difference between A13:A22 and U13:A22. Chemical probing with DMS showed that A13:A22 in the most active mutants (Gly13 and Gly14) was a mismatched base pair while U13:A22 in the least active mutants (Gly16 and Gly18) maintained a Watson–Crick base pair. The former was based on the accessibility of N1 of A22 to DMS while the latter was based on the inaccessibility of N1 of A22 to DMS (not shown). The observation that a mismatched A13:A22 and a Watson–Crick base paired U13:A22 were each correlated with specific structural features at 15:48 also reproduced what was observed in the G15:G48 framework of $tRNA^{Gly}$.[24]

Discussion

We have examined the relationship between structural features at 15:48 and variations at 13:22 in the A15:U48 and A15:A48 frameworks of $tRNA^{Gly}$.[24] This relationship shares a striking parallel with that of the recently characterized G15:G48 framework of $tRNA^{Gly}$. Regardless of the specific hydrogen base pairing at 15:48, we show in all three frameworks that A13:A22 confers unusual structural features to 15:48 that are largely eliminated upon substitution with U13:A22. The most significant feature of the A13:A22 containing mutants is the accessibility of N7 at position 15. In all other tRNAs that have been tested to date, N7 at position 15 is not accessible to chemical probes.[19,20] Together with the framework of *E. coli* $tRNA^{Cys}$, we have now demonstrated the correlation between A13:A22 and the accessibility of N7 at position 15 in four different tRNA sequence frameworks.[14,17,24] In all cases, tRNAs with A13:A22 and the accessibility of N7 at position 15 are efficiently aminoacylated with cysteine, whereas tRNAs with U13:A22 and the inaccessibility of N7 at position 15 are defective in k_{cat}/K_m by about two orders of magnitude. The kinetic and structural parallels between these four frameworks are strong, suggesting that A13:A22 is responsible for tRNA recognition by cysteine tRNA synthetase and that the accessibility of N7 at position 15 is a structural feature closely correlated with aminoacylation activity.

The ability of A13:A22 to determine the accessibility of N7 at position 15 and the ability of U13:A22 to eliminate this accessiblity provide the basis that a tertiary structural feature in a tRNA can be manipulated by another structural feature of the same molecule. In our model of *E. coli* $tRNA^{Cys}$, A13 of the A13:A22 mismach had the potential to base pair with A46 via two symmetrical hydrogen bonds between N1 and N6.[24] This base pairing would form an A13:A22:A46 base triple that would be stabilized by a base–backbone interaction at A22. However, upon substitution of A13 with U13 to create the U13:A22 base pair, we predicted that A46 would switch to base pair with A22 to maintain a U13:A22:A46 base triple. Because A46 is covalently linked to G48, the switch of A46 could indirectly influence the backbone configuration of G15 via the G15:G48 hydrogen bonds. According to this model, the A13:A22:A46 base triple may pose a constraint that allows the N2:N3 pairing of G15:G48, whereas the U13:A22:A46 base triple may pose a different constraint that allows the N1:O6 pairing of G15:G48. Thus, the ability of A13:A22 to determine the accessibiltiy of N7 at G15 and the ability of U13:A22 to eliminate this accessibility may be both mediated through A46.

In $tRNA^{Gly}$, we proposed an N2:N3 pairing for G15:G48 in the active mutants of the G15:G48 framework and an N1:O6 pairing in the least active mutants. The kinetic similarity between the G15:G48 framework of $tRNA^{Gly}$ and that of $tRNA^{Cys}$ suggests that the mechanisms of manipulating the base pairing scheme at 15:48 may be

the same. This mechanism may also facilitate manipulation of the A15:U48 and A15:A48 frameworks described here. In the latter two frameworks, we proposed only one hydrogen bond for A15:U48 and no hydrogen bond for A15:A48 in the active mutants. However, this does not mean that they cannot be stabilized by hydrogen interactions with a solvent molecule or with the adjacent sugar-phosphate backbone, or stabilized by general stacking interactions. Our failure to detect chemical accessibility at the N1 position of A15 in both cases supports the possibility that each N1 is a hydrogen bond acceptor. The putative hydrogen bonds at N1 may hold A15 and U48 together in one framework, and A15 and A48 together in the other framework, such that the accessiblity of N7 of A15 can be linked to nucleotide variations at 13:22 via the same proposed mechanism for the G15:G48 framework.

Our proposed mechanism for how nucleotides at 13:22 determines the structure at 15:48 has implications for the design and manipulation of an RNA tertiary base pair. RNAs have the capacity to form a variety of non-canonical tertiary hydrogen base pairs that are often important for the biological functions of RNAs. Some of these tertiary base pairs have been identified by crystallographic or NMR analyses.[26–34] For example, two unusual G:A base pairs in the core of the hammerhead ribozyme are important for the catalytic activity of the RNA.[26,27] However, we do not yet know the structural context which permits the formation of each of these tertiary base pairs. Biochemical studies of these base pairs that elucidate the structural determinants for their formation will provide necessary insight into higher order RNA structures.

The studies described here provide a foundation for more detailed investigation of the relationship between 13:22 and the structural features at 15:48. This relationship clearly is important for tRNA recognition by cysteine tRNA synthetase. One of the major questions that needs to be addressed is whether A13:A22 is the only mismatch that confers tRNA recognition. We previously showed that U13:U22 does not confer such a recognition[24] and it is because U13:U22 appears to form a homopyrimidine base pair with two hydrogen bonds.[35,36] Other mismatches such as G13:A22 or C13:C22 remain to be tested. A second question is whether there is a covariation relationship between nucleotides at 13:22:46 that may better define the structural elements for tRNA recognition. Both $tRNA^{Cys}$ and $tRNA^{Gly}$ contain A46 and this may limit the observation to A13:A22 as the only determinant for tRNA recognition. A covariation relationship at 13:22:46 may be established by in vitro selection of functional tRNA variants from random sequences at 13, 22 and 46. A third question is whether the relationship between 15:48 and 13:22:46 can be expanded to include novel combinations that improve upon the wild-type efficiency of tRNA recognition. These novel combinations may also be identified through in vitro selection of random sequences.

The structural relationship between 13:22 and 15:48, together with the U73 nucleotide and the GCA anticodon, constitute a strategy for tRNA recognition by cysteine tRNA synthetase. In this strategy, tertiary structural features of the tRNA play a role in aminoacylation. A similar strategy may be employed by alanine tRNA synthetase. In the latter, the G3:U70 base pair in the acceptor stem of *E. coli* $tRNA^{Ala}$ that is important for aminoacylation may have a dual function.[37,38] It may make direct contact with the synthetase.[39] Additionally, this secondary structural element may confer a helical irregularity to the acceptor stem that is recognized by the alanine enzyme.[40] While this model of a dual function for the G3:U70 base pair remains to be tested with further biochemical experiments, it appears that the ability of an RNA structural feature to influence the specificity and efficiency of aminoacylation may be more general than previously anticipated.

Materials and methods

Site-directed mutagenesis

The gene for the starting material Gly06 was constructed in plasmid pTFMa between the T7 RNA polymerase promoter and a *Bst*N1 restriction site.[14,17] This plasmid carries the F′ intergenic region and can be converted to single-stranded DNA. Site-directed mutagenesis was performed by the procedure of Kunkel.[41,42] Briefly, single-stranded uracil-containing pTFMa was obtained by amplifying the plasmid in CJ236 (*dut*$^-$, *ung*$^-$, *thi*$^-$, *recA-1*, pCJ105 (Cm^r)) and superinfecting the strain with M13K07. A mutagenic oligonucleotide was hybridized to the single-stranded DNA and primer-extended by T7 DNA polymerase. The completed double-stranded plasmid was then used to transform TG1 (*supE, hsdΔ5, thi*, Δ(*lac-proAB*, F′[*traD36, proAB, lacI*^{q}, ZΔM15]) and colonies containing the desired mutations were screened by sequencing analysis.

Preparation of tRNA transcripts

All tRNA transcripts were made by T7 RNA polymerase from tRNA genes in plasmid pTFMa. T7 RNA polymerase was purified from pAR1219/BL-21 according to published procedures.[43,44] Restriction of pTFMa with *Bst*N1 generated the template for T7 transcription.[45] Transcription was performed for 4 h at 37 °C in a 500 μL reaction that contained 40 mM Tris–HCl, pH 8.0, 24 mM $MgCl_2$, 1 mM spermidine, 5 mM dithiothreitol (DTT), 0.01% Triton X-100, 0.24 U/μL RNasin, 4 mM each NTPs, 16 mM GMP, and 84 U/μL T7 RNA polymerase. The concentration of the DNA template in the transcription reaction was 0.1 (g/μL. Full-length transcripts were purified by electrophoresis on a denaruting gel (12% polyacrylamide/7 M urea) and ethanol precipitated. Concentrations of tRNAs were

determined by absorption at 260 nm (1 OD_{260} = 0.04 μg/μL RNA).

Aminocylation with cysteine

The purified cysteine tRNA synthetase was prepared from the overproducer strain pKK107/JM109 as previously described.[46] Prior to aminoacylation assay, all tRNA transcripts were denatured (80 °C, 3 min) and annealed (37 °C, 30 min) in 10 mM Tris–HCl, 1 mM EDTA, and 20 mM $MgCl_2$. Aminoacylation was assayed at 37 °C in a 24-μL reaction that contained 20 mM KCl, 10 mM $MgCl_2$, 25 mM DTT, 2 mM ATP, 20 mM Tris-HCl, pH 7.5, 50 μM cysteine, 0.46 μM [^{35}S]cysteine, and 1–200 μM purified tRNA transcripts. Aminoacylation was initiated by the addition of purified cysteine tRNA synthetase (0.1–1.0 nM) and aliquots of 4 μL were removed to 20 μL of alkylation solution (0.24 M iodoacetic acid, 0.1 M sodium acetate, pH 5.0, in formamide) at appropriate time intervals.[14] After incubation for 30 min at 37 °C, 12 μL of the alkylation reaction was spotted onto a Whatman 3 MM filter pad, precipitated with 5% trichloroacetic acid, and washed as described.[47]

Chemical probes of hydrogen interactions

Chemical accessibilites to kethoxal (N1 and N2 of G), CMCT (N3 of U), and DMS (N1 of A) were determined by previously established procedures.[14,17,24,48] The buffer for kethoxal and DMS modification was 50 mM sodium cacodylate, pH 7.5, whereas the buffer for CMCT modification was 50 mM sodium borate, pH 8.0. Each tRNA (120 pmol) was either briefly denatured (80 °C, 3 min) and annealed to native condition in 10 mM $MgCl_2$, or briefly denatured (90 °C, 2 min) and snap-cooled to semi-native condition on ice in 1 mM EDTA. The native and semi-native tRNAs were then treated with kethoxal (1.8 mg/mL), CMCT (3.8 mg/mL), or DMS (0.4%) at room temperature for 5–20 min. The kethoxal reaction was quenched with 50 mM potassium borate, pH 8.0 (50 mM), followed by incubation on ice for 5 min. The CMCT reaction was quenched with 0.3 M sodium acetate, pH 6.0. The DMS reaction was quenched with 200 mM 2-mercaptoethanol, followed by incubation on ice for 5 min.[49] Modified tRNAs were recovered by ethanol precipition and resuspended in 5 μL of water. An appropriate oligonucleotide primer (5′-labeled with ^{32}P, ~10^5 cpm/μL, 2 μL) was added to the modified tRNA and the mixture was denatured (90 °C, 2 min) and snap-cooled on ice. AMV-reverse transcriptase (Life Sciences, 2 units) was added to initiate primer extension (in 50 mM Tris–HCl, pH 8.3, 6 mM $MgCl_2$, 40 mM KCl, 27 mM potassium borate, 0.33 mM each dNTP) at 42 °C for 30 min. Primer extension was stopped by the addition of 10 mM Tris–HCl (pH 8.0) and 10 mM EDTA, and the modified tRNA was passed through a C18 cartridge (Waters, Sep-Pak® Vac 1cc), washed with two 500 μL portions of water, and eluted with 500 μL of 7:7:6 acetonitrile:methanol:water. Eluates were dried in a Speed-Vac, resuspended, and analyzed by electrophoresis on a 10% polyacrylamide/7 M urea gel. Reverse transcriptase sequencing of unmodified tRNA and dideoxynucleotides (Pharmacia) was run in parallel to provide a reference to map the sites of modifications. Each modification reaction was performed at least two times.

Chemical probes of stacking interactions

Chemical accessibilites to DEPC (N7 of A) was determined by previously established procedures.[50] Each tRNA (3′ end labeled, ~2×10^5 cpm), mixed with 160 pmol of carrier tRNAs, was briefly denatured and annealed to either native (10 mM $MgCl_2$) or semi-denaturing (1 mM EDTA) conditions as described above in cacodylate buffer. DEPC was added to 20% at room temperature for 10 min. Modified tRNAs were recovered by ethanol precipitation, resuspended in 0.3 M sodium acetate, pH 6.0, and precipitated again. They were then treated with 1.0 M aniline/acetate buffer (60 °C for 10 min) to generate tRNA fragments that terminate at the site of modification. The aniline scission reaction was dried in a Speed-Vac, resuspended in 0.3 M sodium acetate, pH 6.0, and the products recovered by ethanol precipitation. These tRNA fragments were separated by electrophoresis on an 10% polyacrylamide/7 M urea gel, which was analyzed after autoradiography. Each tRNA was tested with DEPC modifications at least twice.

Acknowledgements

This work was supported in part by grant number GM47935 from the NIH (to Y.-M.H.), a TJU Institutional grant (to Y.-M.H.), and a grant from the Lucille P. Markey Charitable Trust.

We thank Dr Eric Westhof for the initial modeling of the G15:G48 tertiary base pair in *E. coli* $tRNA^{Cys}$ and thank Dr Virginia Rath for sharing unpublished results concerning the possibility of a 13:22:46 base triple.

References

1. Eriani, G.; Delarue, M.; Poch, O.; Gangloff, J.; Moras, D. *Nature* **1990**, *347*, 203.

2. Cusack, S.; Berthet-Colominas, C.; Hartlein, M.; Nassar, N.; Leberman, R. *Nature (London)* **1990**, *347*, 249.

3. Kim, S. H.; Suddath, F. L.; Quigley, G. J.; McPherson, A.; Sussman, J. L.; Wang, A. H. J.; Seeman, N. C.; Rich, A. *Science* **1974**, *185*, 435.

4. Robertus, J. D.; Ladner, J. E.; Finch, J. T.; Rhodes, D.; Brown, R. S.; Clark, B. F. C.; Klug, A. *Nature (London)* **1974**, *250*, 546.

5. Westhof, E.; Dumas, P.; Moras, D. *J. Mol. Biol.* **1985**, *184*, 119.

6. Schimmel, P. *Biochemistry* **1989**, *28*, 2747.

7. Normanly, J.; Abelson, J. *Ann. Rev. Biochem.* **1989**, *58*, 1029.

8. Schulman, L. H. *Prog. Nucleic Acid Res. Mol. Biol.* **1991**, *41*, 23.

9. Giegé, R.; Puglisi, J. D.; Florentz, C. *Prog. Nucleic Acid Res. Mol. Biol.* **1993**, *45*, 129.

10. Rould, M. A.; Perona, J. J.; Sll, D.; Steitz, T. A. *Science* **1989**, *246*, 1135.

11. Ruff, M.; Krishnaswamy, S.; Boeglin, M.; Poterszman, A.; Mitschler, A.; Podjarny, A.; Rees, B.; Thierry, J. C.; Moras, D. *Science* **1991**, *252*, 1682.

12. Biou, V.; Yaremchuck, A.; Tukalo, M.; Cusack, S. *Science* **1994**, *263*, 1404.

13. Pallanck, L.; Li, S.; Schulman, L. H. *J. Biol. Chem.* **1992**, *267*, 7221.

14. Hou, Y. M.; Westhof, E.; Giegé, R. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 6776.

15. Hou, Y. M.; Sterner, T.; Bhalla, R. *RNA* **1995**, *1*, 707.

16. Komatsoulis, G. A.; Abelson, J. *Biochemistry* **1993**, *32*, 7435.

17. Hou, Y. M. *Biochemistry* **1994**, *33*, 4677.

18. Steinberg, S.; Misch, A.; Sprinzl, M. *Nucleic Acids Res.* **1993**, *21*, 3011.

19. Romby, P.; Moras, D.; Dumas, P.; Ebel, J. P.; Giegé, R. *J. Mol. Biol.* **1987**, *195*, 193.

20. Theobald, A.; Springer, M.; Grunberg-Manago, M.; Ebel, J. P. Giegé, R. *Eur. J. Biochem.* **1988**, *175*, 511.

21. Sampson, J. R.; DiRenzo, A. B.; Behlen, L. S.; Uhlenbeck, O. C. *Biochemistry* **1990**, *29*, 2523.

22. Hou, Y. M.; Sterner, T.; Jansen, M. *Biochemistry* **1995**, *34*, 2978.

23. Wing, R.; Drew, H.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E. *Nature (London)* **1980**, *287*, 755.

24. Hamann, C. S.; Hou, Y. M. *Biochemistry* **1997** in press..

25. Ehresmann, C.; Baudin, P.; Mougel, M.; Romby, P.; Ebel, J. P.; Ehresmann, B. *Nucelic Acids Res.* **1987**, *15*, 9109.

26. Pley, H. W.; Flaherty, K. M.; McKay, D. B. *Nature (London)* **1994**, *372*, 68.

27. Scott, W. G.; Finch, J. T.; Klug, A. *Cell* **1995**, *81*, 991.

28. Jiang, F.; Kumar, R. A.; Jones, R. A.; Patel, D. J. *Nature (London)* **1996**, *382*, 183.

29. Heus, H. A.; Pardi, A. *Science* **1991**, *253*, 191.

30. Cate, J. H.; Gooding, A. R.; Podell, E.; Zhou, K.; Golden, B. L.; Kundrot, C. E.; Cech, T. R.; Doudna, J. A. *Science* **1996**, *273*, 1678.

31. Cate, J. H.; Gooding, A. R.; Podell, E.; Zhou, K.; Golden, B. L.; Szewczak, A. A.; Kundrot, C. E.; Cech, T. R.; Doudna, J. A. *Science* **1996**, *273*, 1696.

32. Wimberly, B.; Variani, G.; Tinoco, Jr., I. *Biochemistry* **1993**, *32*, 1078.

33. Puglisi, J. D.; Chen, L.; Frankel, A. D.; Williamson, J. R. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 3680.

34. Battiste, J. L.; Tan, R.; Frankel, A. D.; Williamson, J. R. *Biochemistry* **1994**, *33*, 2741.

35. Baeyens, K. J.; DeBondt, H. L.; Holbrook, S. R. *Structural Biology* **1995**, *2*, 56.

36. Wahl, M. C.; Rao, S. T.; Sundaralingam, M. *Nature Structural Biology* **1996**, *3*, 24.

37. Hou, Y. M.; Schimmel, P. *Nature* **1988**, *333*, 140.

38. McClain, W. H.; Foss, K. *Science* **1988**, *240*, 793.

39. Musier-Forsyth, K.; Usman, N.; Scaringe, S.; Doudna, J.; Green, R.; Schimmel, P. *Science* **1991**, *253*, 784.

40. Gabriel, K.; Schneider, J.; McClain, W. H. *Science* **1996**, *271*, 195.

41. Kunkel, T. A. *Proc, Natl. Acad. Sci. U.S.A.* **1985**, *82*, 488.

42. Kunkel, T. A.; Roberts, J. D.; Zakour, R. A. *Methods Emzymol.* **1987**, *154*, 367.

43. Grodberg, J.; Dunn, J. J. *J. Bacteriol.* **1988**, *170*, 1245.

44. Zawadzki, V.; Gross, H. J. *Nucleic Acids Res.* **1991**, *19*, 1948.

45. Sampson, J. R.; Uhlenbeck, O. C. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 1033.

46. Hou, Y. M.; Shiba, K.; Mottes, C.; Schimmel, P. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 976.

47. Schreier, A. A.; Schimmel, P. *Biochemistry* **1972**, *11*, 1582.

48. Moazed, D.; Stern, S.; Noller, H. F. *J. Mol. Biol.* **1986**, *187*, 399.

49. Pyle, A. M.; Cech, T. R. *Nature* **1991**, *350*, 628.

50. Peattie, D. A.; Gilbert, W. *Proc. Natl. Acad. Sci. USA* **1980**, *77*, 4679.

(*Received in U.S.A.* 10 *October* 1996; *accepted* 18 *February* 1997)