Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 6147–6151, July 1991
Biochemistry

Rapid determination of nucleotides that define tRNA^{Gly} acceptor identity

WILLIAM H. MCCLAIN, K. FOSS, R. A. JENKINS, AND JAY SCHNEIDER

Department of Bacteriology, University of Wisconsin, Madison, WI 53706–1567

Communicated by Sydney Brenner, May 3, 1991 (received for review April 10, 1991)

ABSTRACT Expression of the genetic code depends on the recognition of specific tRNAs by the enzymes that aminoacylate them. A computer comparison of tRNA sequences coupled with analysis of mutant nonsense-suppressor tRNAs has revealed the structural features that distinguish the acceptor identity of *Escherichia coli* tRNA^{Gly} from tRNAs that accept phenylalanine, arginine, lysine, and glutamine. On replacement of several nucleotides in the acceptor stem and anticodon of the latter tRNAs with tRNA^{Gly}-derived residues, the resulting molecules acquired a tRNA^{Gly} identity.

Genes specify the structures of proteins that are made in cells. Following transcription of a gene's sequence into mRNA, 20 types of tRNA molecules translate the codons into the amino acid sequence of the protein. Each acceptor type of tRNA has two specificities, one for the codon and the other for the amino acid. The specificity for the amino acid, termed tRNA acceptor identity, is dictated by some distinctive molecular structure in the tRNA that is manifested during its aminoacylation by the cognate aminoacyl-tRNA synthetase enzyme. The acceptor identity of a tRNA is the net result of two types of interactions, the productive interaction of the tRNA with the cognate aminoacyl-tRNA synthetase and the nonproductive interactions with all other aminoacyl-tRNA synthetases.

The structural basis of tRNA acceptor identity is being elucidated with mutant tRNAs that have acquired new identities following specific changes in a few nucleotides in their sequences. The tRNA mutants are relatively simple to construct with synthetic DNAs, and in some cases their sequences are designed in conjunction with computer analysis of tRNA sequences. Seven of 20 possible tRNA acceptor types have been characterized thus far, including tRNAs that accept serine (1, 2), methionine (3), alanine (4, 5), phenylalanine (6, 7), arginine (8, 9), glutamine (10), and lysine (11). Here we present a description of the determinants of *Escherichia coli* tRNA^{Gly}.

Sequence Comparisons and Genetic Analyses

We aligned 67 *E. coli* tRNA sequences (12) according to their common cloverleaf secondary structure so that the nucleotide residues at corresponding positions could be compared. A computer program (13) then identified the nucleotides that were conserved in the 5 tRNA^{Gly} isoacceptor sequences while differing from the nucleotides in the 62 other tRNA acceptor types. This revealed that no nucleotide at a single position distinguishes tRNA^{Gly} from all others. However, 15 of 176,700 possible two-nucleotide combinations (the combinations of 76 nucleotides taken 2 at a time from a tRNA^{Gly} composite sequence compared with each of 62 other tRNA sequences) were distinguishing characteristics of tRNA^{Gly} (Fig. 1). Although the uridine at position 73 (U73) was most prevalent and thus the best structural predictor, it must function with other nucleotides to discriminate tRNA^{Gly}. Fig. 2 shows the sequences of three tRNA^{Gly} isoacceptors and the location of residue U73 on the acceptor stem of the molecule. We show below that U73 and 5 other nucleotides revealed by the analysis (Fig. 1, residues C2, G3, C35, G71, and C70) are relevant.

If U73 contributes to tRNA^{Gly} acceptor identity, then substituting this residue in another tRNA molecule should confer a partial tRNA^{Gly} identity on it. We substituted U73 in four tRNAs (Phe, Arg, Lys, and Gln) containing varying nucleotides throughout their sequences (Fig. 3) not only to assure the presence of other tRNA^{Gly} determinants in a subset of the tRNAs but also to identify, by comparison, those determinants. Moreover, since residue A73 or G73 is a determinant of tRNAs that accept phenylalanine (6), arginine (11), lysine (11), or glutamine (10), the substitution of U73 into their sequences would potentiate the desired change. For an experimental analysis of amber suppressor tRNAs in *E. coli*, we combined the U73 substitution with a CUA anticodon, allowing the tRNA to translate the amber triplet, UAG, in an mRNA.

Genes for amber suppressor tRNAs containing U73 were inserted in expression plasmid pGFIB (pBR322-type; ref. 1) and separately transformed into *E. coli* cells carrying one of two amber mutant alleles, *trpA234* or *trpA15*. This allowed a rapid assessment of tRNA identity, since the insertion of glycine or alanine, but not phenylalanine, arginine, lysine, or glutamine, at *trpA234* conferred prototrophic growth to cells in minimal medium (14). Since none of the tRNAs examined here contained the G3·U70 determinant of tRNA^{Ala} (4, 5), the insertion of alanine was unlikely. The insertion of as little as 5% glycine as a secondary amino acid is sufficient for *trpA234* prototrophy (4), making this method quite sensitive. The *trpA15* allele accepts many different amino acids (14) and served as a check of suppressor tRNA activity with little regard to amino acid specificity. Unmutated versions of the four amber suppressor tRNA genes containing the wild-type nucleotide 73 were also combined with *trpA234* and *trpA15* for reference with the tRNA mutants.

When the U73 mutants of amber suppressor tRNA^{Phe}, tRNA^{Arg}, tRNA^{Lys}, and tRNA^{Gln} (named PHEU73, ARGU73, LYSU73, and GLNU73, respectively) were tested individually with *trpA234*, only two of the four mutants, PHEU73 and ARGU73, allowed cell growth on minimal medium, indicating that glycine was being inserted. The cells containing ARGU73 grew slowly, as if the suppressor tRNA were only partially active or inserting small amounts of glycine. The other two U73 tRNA mutants and the four tRNAs with the wild-type nucleotide 73 did not allow growth of *trpA234*-containing cells and presumably were not inserting significant amounts of glycine. Each of the four U73 tRNA mutants and the four unmutated tRNAs allowed *trpA15*-containing cells to grow on minimal medium, indicating that all of the suppressor tRNAs were active; however, slow growth was observed with ARGU73 and LYSU73, indicating reduced activity.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

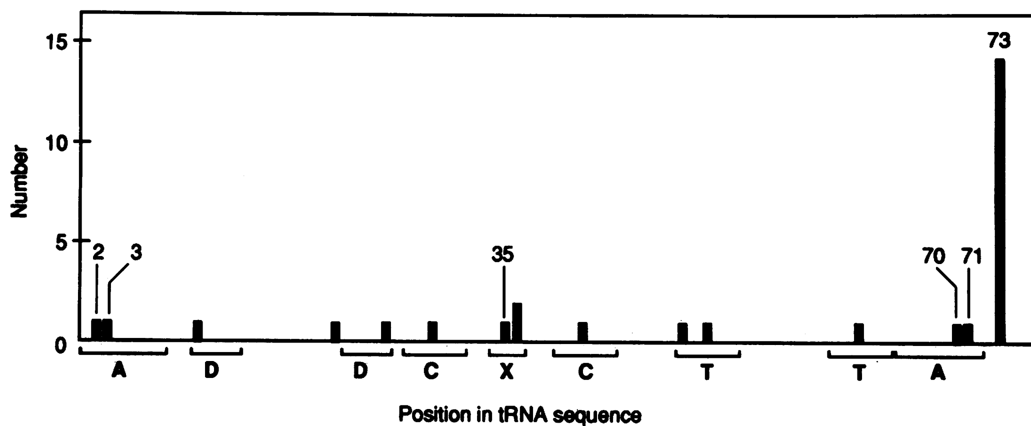

FIG. 1. Histogram of the two-residue discriminators (13) of *E. coli* tRNA^{Gly}. The tRNA^{Gly} cloverleaf regions are indicated as A, acceptor stem; C, anticodon stem; D, D stem; T, T stem; and X, anticodon. The discriminators were at residues 2, 73; 3, 73; 10, 73; 16, 73; 21, 73; 25, 73; 29, 73; 35, 36; 36, 73; 41, 73; 49, 73; 51, 73; 63, 73; 70, 73; and 71, 73.

A comparison (15) of transcribed tRNA sequences revealed that nucleotide pair C2·G71 is shared by the molecules of PHEU73, ARGU73, and all tRNA^{Gly} isoacceptors and is replaced by G2·C71 in LYSU73 and GLNU73 (Fig. 4, underlined residues). No other nucleotide or nucleotide pair shows this correlation. This suggests that C2·G71 is another determinant of tRNA^{Gly} identity. A new gene for amber suppressor tRNA^{Lys} containing U73 and C2·G71 was constructed and examined. The mutant tRNA, LYSU73C2·G71, allowed *trpA234*-containing cells to grow on minimal medium, indicating that glycine was being inserted. Thus, successive sequence comparisons, initially involving 67 tRNAs and then fewer, guided the construction of tRNA mutants and identified U73 and C2·G71 as tRNA^{Gly} determinants.

Biochemical Analysis

The magnitude of the change to tRNA^{Gly} identity was quantified for each mutant by sequencing suppressed dihydrofolate reductase protein. The protein was obtained from cells containing a second plasmid (pACYC184-type) expressing dihydrofolate reductase protein from an mRNA with an amber triplet near the protein's amino terminus (16). The identity of the mutant tRNA was reflected by the amino acid found in suppressed dihydrofolate reductase protein. An analysis revealed that amber suppressors PHEU73 and ARGU73 inserted 42% and 4% glycine, respectively, whereas amber suppressor LYSU73C2·G71 inserted 78% glycine. Neither LYSU73 nor GLNU73 inserted glycine (Table 1). While confirming the genetic studies, the protein data also indicated that other, uncharacterized determinants of tRNA^{Gly} identity existed, since the functional change was incomplete with every mutant.

We wished to confirm that U73 and C2·G71 contribute to tRNA^{Gly} identity and also examine the importance of other conserved residues in the acceptor stem of tRNA^{Gly}. Thus, genes with substitution mutations in amber suppressor GLY3 (Fig. 2*C*) were constructed and tested. This analysis corroborated the importance of U73 (mutant GLY3A73) and C2·G71 (mutant GLY3G2·C71) and also revealed that the first and third nucleotide pairs (mutants GLY3C1·G72 and GLY3C3·G70) contribute to tRNA^{Gly} identity, since each mutant tRNA degraded both the specificity for inserting glycine and the suppression efficiency for transmitting the growing protein chain beyond the nonsense triplet (Table 1). Mutation of the third nucleotide pair was less consequential, indicating that it is less important than residue 73 and the first two nucleotide pairs. Inconclusive results were obtained with respect to the fourth nucleotide pair (mutant GLYC4·G69), since only the suppression efficiency was affected, while the

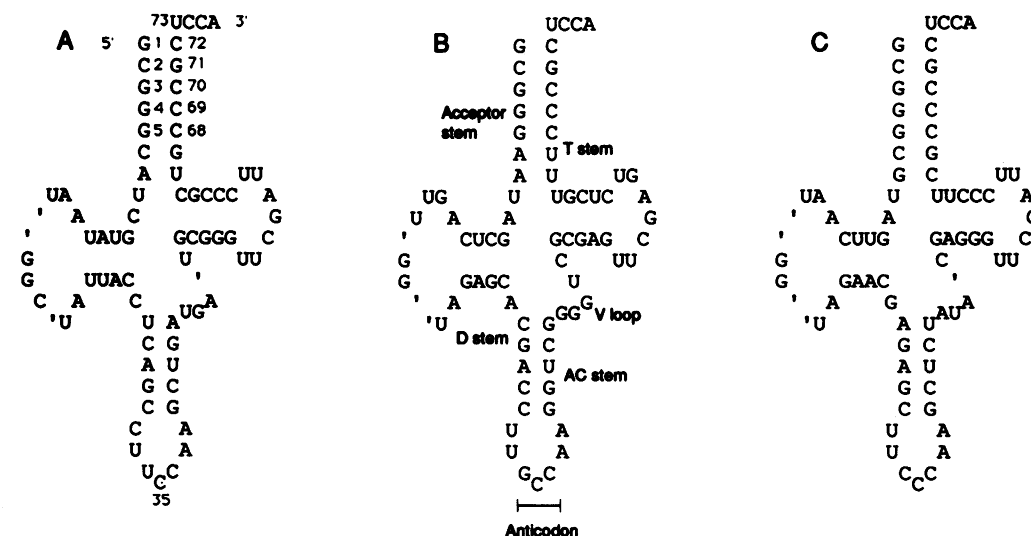

FIG. 2. Cloverleaf arrangement of nucleotide sequences of *E. coli* tRNA^{Gly} isoacceptors (12) corresponding to GLY1 (*A*), GLY2 (*B*), and GLY3 (*C*). Nucleotide modifications are not indicated. The ′ symbol represents an alignment gap. Amber or opal suppressor derivatives change the anticodon sequences to CUA or UCA, respectively.

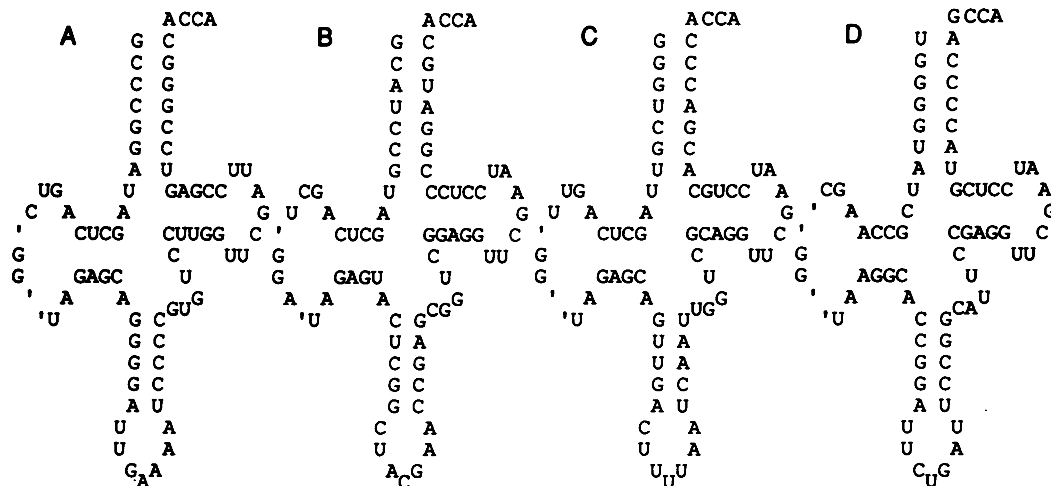

FIG. 3. Cloverleaf arrangement of nucleotide sequences of *E. coli* tRNAs (12) corresponding to phenylalanine (*A*), arginine (*B*), lysine (*C*), and glutamine (*D*). Nucleotide modifications are not indicated. The ′ symbol represents an alignment gap. Amber or opal suppressor derivatives change the anticodon sequences to CUA or UCA, respectively.

fifth nucleotide pair (mutant GLYC5·G68) was functionally insensitive to substitution (Table 1). These findings are consistent with the above results showing that other tRNAs changed to tRNA^{Gly} identity only when tRNA^{Gly}-derived residues were present at residue 73 and the first and second nucleotide pairs. The third nucleotide pair was less important in both types of mutants.

Others have shown that the second and third anticodon nucleotides (C35 and C36) can be important but are not essential for *in vitro* aminoacylation of *E. coli* tRNA^{Gly} (18, 19). We evaluated the role of C35 in tRNA^{Gly} acceptor identity with opal (anticodon UCA; C35) and amber (anticodon CUA; U35) suppressors of three different tRNA^{Gly} isoacceptors. The gene sequences were based on *E. coli* tRNA^{Gly} molecules containing wild-type anticodons of UCC (GLY1) or GCC (GLY2) or CCC (GLY3) (Fig. 2). When tested individually, suppressor forms of GLY2 and GLY3 gave only glycine in suppressed protein. However, while opal suppressor GLY1 gave only glycine, amber suppressor GLY1 gave 20% glycine and 74% glutamine (Table 1), demonstrating that tRNA^{Gly} identity had been damaged through changes in the molecule's productive and nonproductive interactions with glycyl- and glutaminyl-tRNA synthetases. The insertion of glycine by amber suppressor GLY3 (16) and of both glycine and glutamine by amber suppressor GLY1 has been observed (11, 16). Therefore, residue C35 is necessary in GLY1 tRNA. In contrast, residue C35 is not necessary for the productive interaction of GLY2 and GLY3 tRNAs with the glycyl-tRNA synthetase, and/or amber suppressors of these tRNAs are poorer substrates for other tRNA synthetases than they are for the Gly-tRNA synthetase. The other two anticodon nucleotides at residues 34 and 36 exhibit variation in wild-type and/or suppressor tRNAs without affecting specificity and thus assume less important roles in tRNA^{Gly} identity.

Dynamics of tRNA Identity

We have shown that U73, G1·C72, C2·G71, G3·C70, and C35 are determinants of tRNA^{Gly} identity. To gain more perspective on the relationship of structure to tRNA^{Gly} function and on the dynamics of tRNA^{Gly} identity, we have modified the genes of amber suppressors PHEU73, ARGU73, LYSU73, LYSU73C2·G71, and GLNU73 so that the resulting tRNA

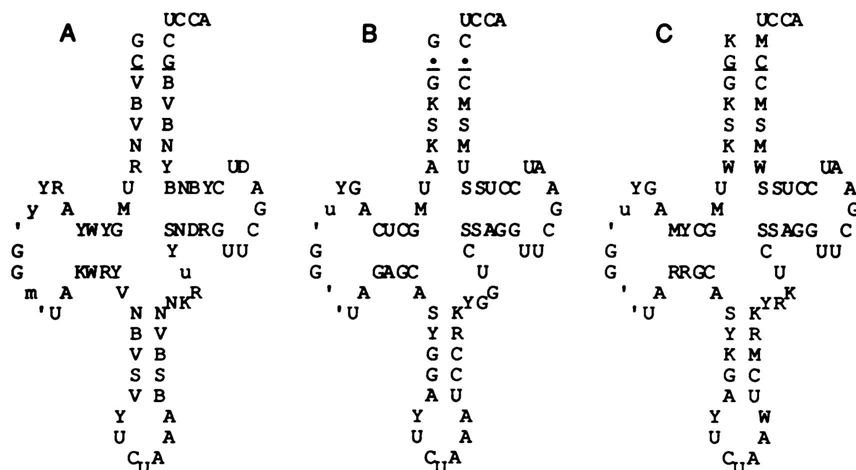

FIG. 4. Comparison of composite nucleotide sequences (15) representing "Gly-like" (PHEU73, ARGU73, and tRNA^{Gly} isoacceptors) (*A*), "non-Gly-like" (LYSU73 and GLNU73) (*C*), and common (*B*) nucleotides. The positions where the two composites do not contain a common nucleotide are underlined and marked with the • symbol in *B*. Incompletely specified nucleotides are designated as follows: M = A or C; R = A or G; W = A or U; S = C or G; Y = C or U; K = G or U; V = A, C, or G; H = A, C, or U; D = A, G, or U; B = C, G, or U; and N = A, C, G, or U. An alignment gap is indicated by the ′ symbol. Lowercase letter indicates that the position contains either an alignment gap or the indicated nucleotide.

Table 1. Functional properties of suppressor tRNAs

tRNA	Amino acid(s) in suppressed protein: Gly, %	Other (%)	Suppression efficiency, %
Amber suppressor			
PHEU73	42	Phe (42)	2.0
ARGU73	4*	Arg (66), Lys (7)	7.5
LYSU73		Lys (81)	7.3
LYSU73C2·G71	78	Lys (11)	19
GLNU73		Gln (93)	25
GLY1	20	Gln (76)	29
GLY2	93		10
GLY3	89		29
GLY3A73	26	Gln (21), Lys (14)†	0.3
GLY3C1·G72	64	Gln (19), Pro (18)	0.2
GLY3G2·C71	30	Gln (57)	1.0
GLY3C3·G70	79	Thr (10)	15
GLY3C4·G69	92		17
GLY3C5·G68	92		25
Opal suppressor			
GLY1	95		41
GLY2	98		4.4
GLY3	99		14
LYSU73C2·G71	99		19
GLNU73G1·C72C2·G71	94		6.9
PHEU73	48	Phe (40)	2.8
ARGU73	13	Arg (77)	5.5
LYSU73			1.1‡
GLNU73			0.5‡

The wild-type tRNA sequences are shown in Figs. 2 and 3. Methods for construction of mutant genes, purification and sequencing of dihydrofolate reductase protein, measurement of suppression efficiency, and analysis of data have been described (4, 6, 8, 11). The tRNAs were expressed from plasmid pGFIB and the proteins from plasmid pDAYQ (16) modified as described previously (11) or as described below. Most tRNA mutants were characterized for amino acid specificity at residue 10 of dihydrofolate reductase protein. We recently relocated the nonsense triplet to residue 3 to increase the sensitivity and rapidity of protein sequencing. The transcribed mRNA of the residue 3 mutants contained UAG (plasmid pD3am) or UGA (plasmid pD3op) at codon 3 combined with AUG, GUA, and GAU at codons 4, 10, and 11, respectively. The tRNAs characterized at residue 3 were amber suppressors GLY2, LYSU73C2·G71, and GLY3C3·G70 and opal suppressors LYSU73C2·G71 and GLNU73G1·C72C2·G71. To obtain sufficient suppressed protein from opal suppressor PHEU73, we used cells carrying a mutant protein release factor 2, allele *prfB1* (17). Amino acid yield < 5% was retained but not reported and the combined yield of reported residues was generally <100%. The Gln value includes ≤14% Glu. Cys was not analyzed. Suppression efficiency measurements used UAG allele *A16* or UGA allele *U4* in *lacI–Z* fusions; the reported value is the percentage of enzyme activity relative to a wild-type *I–Z* fusion (which averaged 190 units) and has not been corrected for the value of cells without a suppressor tRNA (0.00% for *A16* and 0.89% for *U4*).

*Confirmed genetically by suppression of *trpA234* auxotrophy.

†Also 13% Tyr and 6% each Thr, Pro, and Met.

‡Value similar to that in cells without an opal suppressor tRNA.

molecules contain additional $tRNA^{Gly}$ determinants. After introducing C35 in each gene by changing the anticodon from the amber- to the opal-suppressing form, we observed the following (Table 1). (*i*) The opal suppressors PHEU73 and ARGU73 inserted both glycine and either phenylalanine or arginine, respectively. A higher percentage of glycine was noted for opal suppressor ARGU73 in comparison with amber suppressor ARGU73. (*ii*) The opal suppressor LYSU73C2·G71 inserted only glycine. (*iii*) The opal suppressors LYSU73 and GLNU73 were judged inactive since their suppression efficiency was similar to *E. coli* cells without a suppressor tRNA. This suggests that the tRNA molecules of LYSU73 and GLNU73 lack sufficient determinants for either $tRNA^{Gly}$ identity or the respective old tRNA identity. We attribute the latter defects to the absence of U35, which determines $tRNA^{Lys}$ identity (11, 16) and $tRNA^{Gln}$ identity (10). For greater scrutiny, we constructed a mutant gene for $tRNA^{Gln}$ containing all of the $tRNA^{Gly}$ determinants. This mutant, opal suppressor GLNU73G1·C72C2·G71, inserted only glycine.

Several conditions are needed to change the identity from one tRNA acceptor type to another. The redesigned tRNA must contain all the appropriate nucleotides for the new acceptor function, while lacking residues for the old function. A quantitative switch can occur when determinants for both new and old functions reside in different nucleotide types at the same position in the two tRNA molecules. This can explain the complete switch to $tRNA^{Gly}$ identity observed with opal suppressors LYSU73C2·G71 and GLNU73G1·C72C2·G71, since both tRNA molecules contain U73, G1·C72, C2·G71, G3·C70, and C35 while also lacking determinants of either $tRNA^{Lys}$ (A73 and U35; ref. 11) or $tRNA^{Gln}$ (G73, U1·A72, G2·C71, and U35; ref. 10). The incomplete switch to $tRNA^{Gly}$ identity observed with opal suppressors PHEU73 and ARGU73 may reflect the interplay of several factors that are not mutually exclusive. (*i*) The modified tRNAs lack G3·C70 corresponding to $tRNA^{Gly}$ identity. Moreover, the resident 3·70 nucleotide pair in $tRNA^{Phe}$ and $tRNA^{Arg}$ could also contribute to their identity. (*ii*) The modified tRNA molecules retain known determinants corresponding to either $tRNA^{Phe}$ identity (e.g., U20; ref. 6) or $tRNA^{Arg}$ identity (e.g., A20; refs. 8 and 9), resulting in competition between identity systems (20, 21). (*iii*) The modified $tRNA^{Phe}$ and $tRNA^{Arg}$ molecules could contain steric or ionic structures that act negatively toward the $tRNA^{Gly}$-aminoacylating enzyme.

Conclusions

The complexity of $tRNA^{Gly}$ identity is manifested in the number and types of potential interactions of tRNAs with aminoacyl-tRNA synthetases and in the structural diversity of the $tRNA^{Gly}$ isoacceptors. Our belief that the acceptor stem and anticodon determine $tRNA^{Gly}$ identity rests not only on their necessity for maintaining $tRNA^{Gly}$ function but also on their sufficiency for switching several other tRNA identities to $tRNA^{Gly}$. They are the major determinants, but given the intricacies of the system, others probably exist.

We thank E. J. Murgola for the *trpA* strains; Y. Nakamura for the *prfB1* strain; K. L. Mittelstadt, S. Schwartz, C. Mattison, and J.-C. Wang for technical assistance; G. Roberts and Y.-M. Chen for comments on the manuscript; R. Niece and M. Strickland at the University of Wisconsin Biotechnology Center for oligodeoxynucleotide synthesis and HPLC analysis; and L. Williams at the University of Southern California Microchemical Core Laboratory for protein sequencing. This work was supported by National Institutes of Health Grant GM42123.

1. Normanly, J., Ogden, R. C., Horvath, S. J. & Abelson, J. (1986) *Nature (London)* **321**, 213–219.
2. Himeno, H., Hasegawa, T., Ueda, T., Watanabe, K. & Shimizu, M. (1990) *Nucleic Acids Res.* **18**, 6815–6819.
3. Schulman, L. H. & Pelka, H. (1988) *Science* **242**, 765–768.
4. McClain, W. H. & Foss, K. (1988) *Science* **240**, 793–796.
5. Hou, Y.-M. & Schimmel, P. (1988) *Nature (London)* **333**, 140–145.
6. McClain, W. H. & Foss, K. (1988) *J. Mol. Biol.* **202**, 697–709.
7. Sampson, J. R., DiRenzo, A. B., Behlen, L. S. & Uhlenbeck, O. C. (1989) *Science* **243**, 1363–1366.
8. McClain, W. H. & Foss, K. (1988) *Science* **241**, 1804–1807.
9. Schulman, L. H. & Pelka, H. (1989) *Science* **246**, 1595–1597.
10. Rould, M. A., Perona, J. J., Söll, D. & Steitz, T. A. (1989) *Science* **246**, 1135–1142.

11. McClain, W. H., Foss, K., Jenkins, R. A. & Schneider, J. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 9260–9264.
12. Sprinzl, M., Hartmann, T., Weber, J., Blank, J. & Zeidler, R. (1989) *Nucleic Acids Res.* **17,** 1–172.
13. McClain, W. H. & Nicholas, H. B., Jr. (1987) *J. Mol. Biol.* **194,** 635–642.
14. Murgola, E. J. (1985) *Annu. Rev. Genet.* **19,** 57–80.
15. Nicholas, H. B., Jr., Chen, Y.-M. & McClain, W. H. (1987) *Comput. Appl. Biosci.* **3,** 53.
16. Normanly, J., Kleina, L. G., Masson, J.-M., Abelson, J. & Miller, J. H. (1990) *J. Mol. Biol.* **213,** 719–726.
17. Kawakami, K., Inada, T. & Nakamura, Y. (1988) *J. Bacteriol.* **170,** 5378–5381.
18. Hill, C. W., Combriato, G. & Dolph, W. (1974) *J. Bacteriol.* **117,** 351–359.
19. Sabban, E. L. & Bhanot, O. S. (1982) *J. Biol. Chem.* **257,** 4796–4805.
20. Yarus, M. (1972) *Nature (London)* **239,** 106–108.
21. Swanson, R., Hoben, P., Sumner-Smith, M., Uemura, H., Watson, L. & Söll, D. (1988) *Science* **242,** 1548–1551.