tRNA Glycylation System from *Thermus thermophilus.* $tRNA^{Gly}$ Identity and Functional Interrelation with the Glycylation Systems from Other Phylae†

Marie-Hélène Mazauric, Hervé Roy, and Daniel Kern*

Unité Propre de Recherche 9002 "Structure des Macromolécules Biologiques et Mécanismes de Reconnaissance", Institut de Biologie Moléculaire et Cellulaire du Centre National de la Recherche Scientifique, 15, Rue René Descartes, 67084, Strasbourg, Cedex, France

Received June 17, 1999; Revised Manuscript Received July 14, 1999

ABSTRACT: The systems of tRNA glycylation belong to the most complex aminoacylation systems since neither the oligomeric structure of glycyl-tRNA synthetases (GlyRS) nor the discriminator bases in $tRNA^{Gly}$ are conserved in the phylae. To better understand the structure−function relationship in glycylation systems of various origins and the functional peculiarities related to their structural divergences, the elements in tRNA conferring its glycine identity in *Thermus thermophilus* were characterized and compared to those of other systems. Thermophilic identity is conferred by the G_1-C_{72}, C_2-G_{71}, G_3-C_{70}, and C_{50}-G_{64} pairs together with the G_{10}, U_{16}, C_{35}, and C_{36} single residues. In contrast to most other aminoacylation systems, the discriminator base is not directly involved in identity. Transplantation of these elements in $tRNA^{Asp}$ and $tRNA^{Phe}$ converts specificity toward glycine albeit conservation of nucleotide 73. Analysis of the functional interrelation of the identity elements shows coupling in synthetase recognition of the elements from anticodon and G_{10} whereas those from acceptor arm are recognized independently. Despite nondirect implication in identity, the discriminator base contributes cooperatively with C_{36} in specificity of glycylation. The link between the structural heterogeneity and the functional divergence of the glycylation systems and the phylogenic interrelation of these systems were approached by comparing the ability of GlyRSs of various phylae to glycylate heterologous $tRNA^{Gly}$. Dimeric GlyRSs from mammalian and archaebacteria acylate efficiently only eukaryotic and archaebacterial $tRNA^{Gly}$ with a discriminatory A_{73}, whereas tetrameric *Escherichia coli* GlyRS acylates only eubacterial $tRNA^{Gly}$ with a discriminatory U_{73}. In contrast, dimeric yeast GlyRS acylates efficiently both eukaryotic and archaebacterial $tRNA^{Gly}$ as well as peculiar prokaryotic isoacceptors. Species specificity is lost with the dimeric GlyRS from *Thermus thermophilus* that acylates efficiently eubacterial, archaebacterial, and eukaryotic $tRNA^{Gly}$. These features are discussed in the context of the evolution of the glycylation systems and the phylogenic interrelation of the organisms.

Synthesis of functional proteins relies on accurate aminoacylation of tRNAs. This crucial step is promoted by aminoacyl-tRNA synthetases[1] (aaRSs), which select the cognate tRNAs among 20 sets of isoacceptors and acylate their 3′ accepting end with the proper amino acid. Despite their similar function, aaRSs exhibit a high structural diversity. Most are homodimers, but monomers and tetramers are also present and their polypeptide chains exhibit large variabilities in sequences and sizes (reviewed in refs *1*−*4*). This heterogeneity is reinforced at the functional level by variabilities in the catalytic mechanism, proofreading processes, and noncanonical functions exerted by some of them (*1*−*4*). However, sequence alignments and comparison of 3D structures allowed partition of these enzymes in two classes each composed of 10 members, differing structurally by sequence signatures and conserved motifs and by the ATP binding domain and functionally by the initial position of tRNA acylation (*5*, *6*). Since synthetases of a given specificity belong unambiguously to the same class independently upon their origin, these enzymes may derive from two ancesters which have diverged at an early stage of evolution each toward 10 distinct specificities. Until now, this class ranking is just broken by one exception, that of the euryarchaeal LysRSs from *Methanococcus maripalutis*, *Methanococcus jannaschii*, and *Methanobacterium thermoautotrophicum*, which are deprived from the characteristic motifs of class 2 synthetases but possess an ATP-binding domain resembling the Rossman fold found in class 1 synthetases (*7*). Analysis of the archaebacterial genomes suggest additional examples of nonconservation of the class defining motifs in aaRSs such CysRS for which no ORF could be characterized in the genome from *M. jannashii* (*8*). However, this observation is also compatible with the lack of that aaRS which is then compensated by formation of the homologous aa-tRNA via conversion of the amino acid mischarged on $tRNA^{Cys}$ by a noncognate aaRS. Such indirect

†This work was supported by grants from Centre National de la Recherche Scientifique (CNRS) and from Université Louis Pasteur. M.H.M. was recipient of a grant from Association de la Recherche sur le Cancer, and H.R. is recipient of a grant from Ministère de la Recherche et de la Technologie.

* To whom correspondence should be addressed. Phone: 33 (0)3 88 41 70 92. Fax: 33 (0)3 88 60 22 18. E mail: kern@ibmc.u-strasbg.fr.

[1] Abbreviations: aaRS, aminoacyl-tRNA synthetase; the three letter code is used for amino acids, e.g., Gly for glycine and GlyRS for glycyl-tRNA synthetase; BD-cellulose, benzoyldiethylaminoethyl-cellulose; DEAE-cellulose, diethylaminoethyl-cellulose; Hepes, *N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid; PAGE, polyacrylamide gel electrophoresis.

10.1021/bi991392t CCC: $18.00
Published on Web 09/14/1999

pathway of aa-tRNA formation was demonstrated for tRNA glutaminylation and asparaginylation in various organisms deprived of GlnRS and AsnRS (e.g., refs *9*−*11*).

Among aaRSs of a given specificity exhibiting structural divergences in different organisms, the GlyRSs are probably the most intriguing. Their oligomeric structures are not conserved since being α_2 dimers in eukaryotes and archaebacteria and $\alpha_2\beta_2$ tetramers (exceptionally dimers) in eubacteria (*12*). The consensus motifs characterizing class 2 synthetases are degenerated or even absent. In the dimeric GlyRSs, the conserved Pro in motif 1 is substituted by Ser or Thr, so that this atypical motif could only be evidenced when the 3D structure of the enzyme from *T. thermophilus* was established (*13*). In tetrameric GlyRSs, consensus motifs 1 and 2 could until now not be evidenced (*14*−*17*), and sequence similarities in dimeric and tetrameric GlyRSs are apparently missing (*14*). The structural diversity of GlyRSs is reinforced by functional peculiarities. *E. coli* and mammalian enzymes exhibit a strong preference for charging $tRNA^{Gly}$ originated from same species (*14*, *16*, *17*). Since prokaryotic and eukaryotic $tRNA^{Gly}$ differ by their discriminator base (U in prokaryotes and A in eukaryotes; ref *18*), this property was related to an implication of this element in glycine identity. Species specificity is supported by the ability of *E. coli* GlyRS to glycylate human $tRNA^{Gly}$ minihelix after substitution of A_{73} by U and of mammalian GlyRS to glycylate *E. coli* $tRNA^{Gly}$ minihelix after substitution of U_{73} by A (*16*, *17*, *19*). However, the dimeric GlyRSs from *T. thermophilus* and yeast are deprived of species specificity despite conservation in the homologous isoaccepting $tRNA^{Gly}$ of the prokaryotic and eukaryotic characters, namely U_{73} and A_{73}, suggesting that in these systems the discriminator base may not contribute to identity (*14*) and that species specificity is not a conserved feature in glycylation systems.

To rationalize the functional divergences in glycylation systems and to understand their interrelation, the recognition of *T. thermophilus* $tRNA^{Gly}$ by the homologous GlyRS was investigated and the nucleotides defining the thermophilic glycine identity characterized. Results indicate that identity is conferred by nucleotides from the acceptor stem and the anticodon, but in contrast to the homologous sytems from *E. coli* and mammalian, mutation of the discriminator base does not affect glycylation efficiency. Transplantation of the glycine identity set in *T. thermophilus* $tRNA^{Phe}$ and $tRNA^{Asp}$, possessing, respectively, A_{73} and G_{73} results in tRNAs able to be glycylated by the thermophilic GlyRS. Nevertheless, the discriminator base is involved in selection of $tRNA^{Gly}$ in the thermophile, since G_{73} disfavors glycylation, probably by acting as an antideterminant. This effect is considerably enhanced by altering C_{36} from the anticodon. Finally, comparison of the charging properties of GlyRSs of the various phylae reveal that *T. thermophilus* GlyRS, deprived of species specific glycylation, contrasts with yeast, *E. coli*, *M. jannaschii*, and rat liver GlyRSs, which exhibit relaxed or strong species specificity. The results will be discussed in the context of evolution of the glycylation systems.

EXPERIMENTAL PROCEDURES

Materials, Enzymes, and tRNAs. Oligonucleotides were synthesized on an Applied Biosystem 381 A DNA synthesizer using the phosphoramidite method and purified by HPLC on a Nucleosil 120-5-C18 column (Bischoff Chromatography, Zymark, France). [^{14}C]- and [^{3}H]Gly (0.1 and 17.9 Ci/mmol) were from Amersham. Restriction enzymes, T4 polynucleotide kinase, and T7 DNA polymerase were from New England Biolabs and T4 DNA ligase from Boehringer. T7 RNA polymerase (16000 units mg^{-1}) of high grade was prepared as described (*20*). Unfractionated tRNA from *E. coli* was from Boehringer, and pure $tRNA^{Gly(GCC)}$ and $tRNA^{Gly(UCC)}$ from *E. coli* (20 and 25 nmol mg^{-1}) from Subriden. Partially purified $tRNA^{Gly}$ from *T. thermophilus* was obtained by benzoyldiethylaminoethyl-cellulose[1] (BD-cellulose) chromatography of RNA isolated by phenol extraction of the cells and pure $tRNA^{Gly(CCC)}$ (37 nmol mg^{-1}) by additionnal BD-cellulose and anion-exchange chromatographies. Enriched $tRNA^{Gly(CCC)}$ and $tRNA^{Gly(GCC)}$ from yeast (6 nmol mg^{-1}) were obtained by counter-current distribution of bulk tRNA followed by BD-cellulose and salting-out chromatographies. Enriched $tRNA^{Gly}$ from beef and rabbit livers (3 nmol mg^{-1}) and *Bombyx mori* $tRNA^{Gly(GCC)}$ (20 nmol mg^{-1}) were obtained by BD-cellulose and anion-exchange chromatographies of bulk RNA. Pure rabbit liver $tRNA^{Gly(CCC)}$ (35 nmol mg^{-1}) was obtained by a final denaturing PAGE. Protein extracts from yeast, *E. coli* and rat liver, deprived of nucleic acids by DEAE-cellulose chromatography were obtained as described (*14*). Unfractionated tRNA and crude protein extract from *M. jannaschii* were kind gifts from Pr. D. Söll. *T. thermophilus* GlyRS (124 units mg^{-1}) was purified from overproducing *E. coli* BL21 (DE3) pLysS strain as described (*14*).

Preparation of tRNA Transcripts. *T. thermophilus* wild-type or mutated $tRNA^{Gly}$, $tRNA^{Asp}$, and $tRNA^{Phe}$ were obtained by in vitro transcription of synthetic genes flanked upstream by the consensus promoter (−21 to −5) of T7 RNA polymerase followed by a TATA box (−4 to −1) and downstream by a *Bst*NI restriction site. The genes were constructed by shotgun ligation of 10 DNA fragments (16−24 mers) covering both strands and ligated in the *Hin*dIII and *Bam*HI sites of pUC18 as described (*20*). The recombined plasmids were isolated after transformation of the DH5F′ strain and the sequences of the genes verified (*21*). Plasmidic DNA was isolated at large scale using the alkaline method followed by centrifugation on a ClCs gradient (*22*).

In vitro transcriptions were conducted at 37 °C in reaction mixtures of 250 μL containing 40 mM Tris-HCl, pH 8.1, 22 mM $MgCl_2$, 5 mM DTE, 0.01% Triton X-100, 1 mM spermidine, 4 mM of each nucleotide triphosphate, 5 mM GMP, 0.1 μg mL^{-1} plasmidic DNA digested with *Bst*NI, and 7.5 μg of T7 RNA polymerase. Reactions were stopped after 3 h of incubation at 37 °C by phenol/chloroform extraction, and transcripts purified by denaturing PAGE. Transcripts of highest mobility were electroeluted and deprived from last traces of urea by gel-filtration on Sephadex G-25 (*20*); 0.5−1 mg of native $tRNA^{Gly(CCC)}$ transcript of a charging capacity of 28 nmol mg^{-1} was obtained. Concentrations of tRNA and DNA solutions were determined spectrophotometrically; one A_{260nm} unit cm^{-1} corresponds to 40 μg of tRNA and 50 μg of DNA.

Aminoacylation Reactions. The standard aminoacylation mixture (total volume 25−200 μL) contained 100 mM Hepes-Na[1], pH 7.2, 10 mM $MgCl_2$, 30 mM KCl, 2 mM ATP, and Gly either ^{3}H-labeled (20 μM, 1100 cpm $pmol^{-1}$) for

K_M determinations of tRNAs and transcripts or ^{14}C-labeled (100 μM; 50 cpm pmol^{-1}) for k_{cat} determinations. Concentrations of tRNA^{Gly}, transcript, or variants were in the range of 0.1−60 μM and that of GlyRS was 5 nM to 5 μM according to the tRNA tested. When necessary, GlyRS was diluted in 100 mM Hepes-Na, pH 7.2, 10% glycerol, 1 mg mL^{-1} bovine serum albumin, 5 mM 2-mercaptoethanol, and 0.1 mM diisopropylfluorophosphate and dithiothreitol. Initial rates were measured at 37 or 70 °C by determining the labeled aa-tRNA formed in 10−40 μL aliquots after various incubation times. The K_Ms were determined from Lineweaver and Burk plots; each value is an average of at least three independent determinations. Because of the requirement of amino acid of high specific activity, K_Ms for tRNA were determined with Gly concentrations equalling the K_M value. The k_{cat}s were determined independently with saturating substrates concentrations (10−100 K_M). Experimental errors on kinetic constants are within 10% of the indicated values. Analysis of kinetic data for multiple mutants was according to Fersht (*23*) using the formalism of Pütz et al. (*24*). It is recalled that independent recognition of elements is characterized by additive effects of the mutations and k_{cat}/K_M losses equaling the product of the losses of individual mutants, whereas cooperative and anticooperative recognitions of the elements lead to respectively decreased and increased losses as expected from additive effects of the mutations (*25*). Aminoacylations of tRNA^{Gly} by GlyRSs of various origins were conducted by initial rate measurements in the standard glycylation mixture containing saturating substrates concentrations as described (*14*).

RESULTS AND DISCUSSION

Aminoacylation of Various tRNA^{Gly} Species by T. thermophilus Gly-tRNA Synthetase. Two genes encoding tRNA^{Gly} were characterized in *T. thermophilus*; they are tRNA$^{Gly(GCC)}$ and tRNA$^{Gly(CCC)}$ (*26*, *27*). Both are glycylated with similar efficiency since the mixture of the two isoacceptors enriched by BD-cellulose chromatography is charged as well as pure tRNA$^{Gly(CCC)}$ (Table 1). Further, heterologous prokaryotic tRNA^{Gly} from *E. coli* and eukaryotic tRNA^{Gly} from yeast, mammalian, and *B. mori* are charged by the thermophilic GlyRS (Table 1). *E. coli* tRNA$^{Gly(UCC)}$ and tRNA$^{Gly(GCC)}$ and yeast tRNA$^{Gly(GCC)}$ and tRNA$^{Gly(CCC)}$ are charged 2-fold more to 2-fold less efficiently by the thermophilic GlyRS than *T. thermophilus* tRNA^{Gly}, whereas *B. mori* and mammalian tRNA^{Gly} are charged 2−10-fold less efficiently. The consensus sequence of the species well charged by *T. thermophilus* GlyRS, shows conservation of 14 nucleotides in addition to those common to all tRNAs (Figure 1A). These nucleotides may include the elements specifying glycine identity.

Characterization of the Identity Elements for Glycylation by T. thermophilus Gly-tRNA Synthetase. Since *T. thermophilus* tRNA$^{Gly(CCC)}$, either in the native form or overexpressed in *E. coli*, and its transcript are charged with similar efficiencies (Table 1), it can be concluded that the posttranscriptional modifications are not involved in glycylation. Therefore the glycine identity elements could be characterized by in vitro analysis of the charging capacity of variants with mutations within positions conserved in the consensus sequence (Figure 1A). Table 2 summarizes the results.

Table 1: Kinetic Constants of Aminoacylation of Various tRNA^{Gly} by Gly-tRNA Synthetase from *T. thermophilus*

origin and nature of tRNA^{Gly}	kinetic parameters[a] k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M (s^{-1} μM^{-1})	L[b]
Prokaryotic				
T. thermophilus				
tRNA$^{Gly(CCC)}$	0.40	0.20	2.00	1.0
tRNA^{Gly} [c]	0.30	0.10	3.00	0.7
transcript tRNA$^{Gly(CCC)}$	0.40	0.33	1.21	1.6
E. coli				
tRNA$^{Gly(UCC)}$	0.36	0.12	3.10	0.7
tRNA$^{Gly(GCC)}$	0.36	0.16	2.25	0.9
Eukaryotic				
S. cerevisiae				
tRNA$^{Gly(GCC)}$	0.55	0.12	4.74	0.4
tRNA$^{Gly(CCC)}$	0.45	0.4	1.12	1.8
B. mori				
tRNA$^{Gly(GCC)}$	0.13	0.55	0.24	8.3
Mammalian				
beef liver				
tRNA^{Gly} [c]	0.27	0.32	0.84	2.4
rabbit liver				
tRNA$^{Gly(GCC)}$	0.4	2.00	0.20	10.0

[a] The conditions are described in the Experimental Procedures. [b] L (loss) is defined as the ratio of k_{cat}/K_M of charging homologous tRNA^{Gly} over that of charging heterologous tRNA^{Gly}. [c] Mixture of tRNA^{Gly} isoacceptors enriched by BD-cellulose chromatography.

(1) Acceptor Stem. Mutations of the three first base pairs decrease glycylation efficiency, with the strongest effect for mutations of the G_1-C_{72} pair. Substitution of C_{72} by U, which is predicted to preserve pairing with G_1, decreases the efficiency 45-fold, and additional substitution of G_1 by A creating an A_1-U_{72} pair decreases it 2-fold more. Nonconservative substitution of this pair by C_1-G_{72} provokes a drastic loss of charging which is decreased 10^4-fold. Noticeable, the conservative substitutions ($C_{72} \rightarrow$ U and G_1-$C_{72} \rightarrow$ A−U) decrease k_{cat} and increase K_M by a similar factor, whereas the nonconservative substitution (G_1-$C_{72} \rightarrow$ C−G) affects k_{cat} 2 orders of magnitude more than K_M. These observations indicate implication of the first base pair of tRNA^{Gly} in recognition by the thermophilic GlyRS. The important loss provoked by the $C_{72} \rightarrow$ U mutation, which contrasts with the poor one provoked by the $G_1 \rightarrow$ A mutation suggests recognition in the pyrimidine ring of the NH_2 group at position 4, whereas recognition of G_1 involves common purine features.

Mutations of the second base pair moderately affect glycylation, essentially by increasing the K_M. However, mutation of C_2 by U, which is predicted to preserve pairing with G_{71}, decreases the efficiency ∼5-fold more than nonconservative substitution of the pair by G_2-C_{71}, suggesting an involvement of the NH_2 group at position 4 of C_{71} in recognition.

Nonconservative substitution of the third base pair by C_3-G_{70} provokes a loss of charging of ∼500 times. Since conservative substitution of C_{70} by U which is predicted to preserve pairing, decreases charging efficiency 10-fold, the NH_2 group of position 4 from the pyrimidine ring may be involved in recognition.

Efficient charging of prokaryotic and eukaryotic tRNA^{Gly} by *T. thermophilus* GlyRS despite nonconservation of the discriminator base (U in prokaryotes and A in eukaryotes) suggests that this nucleotide does not contribute to glycylation in the thermophile (Figure 1A). However, since this

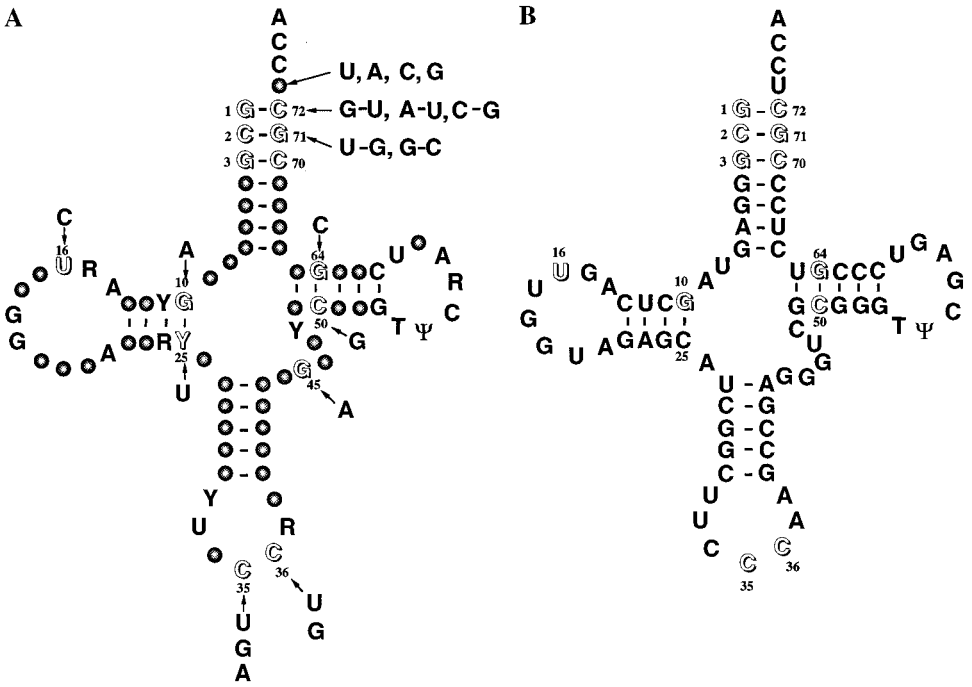

FIGURE 1: Sequence features of tRNA^{Gly} species and identity elements. (A) Consensus sequence of the tRNA^{Gly} of various origins the most efficiently charged by *T. thermophilus* Gly-tRNA synthetase and nature of the nucleotide substitutions. The sequence includes the nucleotides common to *T. thermophilus* tRNA$^{Gly(CCC)}$ and tRNA$^{Gly(GCC)}$, *E. coli* tRNA$^{Gly(UCC)}$, and tRNA$^{Gly(GCC)}$ and yeast tRNA$^{Gly(GCC)}$ and tRNA$^{Gly(CCC)}$: black dots, nonconserved positions; black characters, nucleotides conserved in all tRNAs; open characters, nucleotides conserved in the tRNA^{Gly} efficiently charged by *T. thermophilus* GlyRS. R, purine; Y, pyrimidine. The posttranscriptional modifications are not indicated except ribothymidine 54 and pseudouridine 55. Sequence data and numbering of the positions are according to ref *18*. The nature of the substitutions are indicated in black characters and emphasized by arrows. (B) The identity elements in *T. thermophilus* tRNA$^{Gly(CCC)}$. The elements defining the glycine identity are in open characters.

nucleotide is involved in *E. coli*, mammalian and yeast glycine identities (*14*, *16*, *17*, *28*, *29*), its possible role in thermophilic identity was investigated. Replacement of U_{73} by C or A is without significant effect, but substitution by G decreases glycylation efficiency ~60-fold by affecting equally k_{cat} and K_M. The absence of effect of the nonconservative $U_{73} \rightarrow A$ substitution which alters all chemical groups from U suggests that no chemical group of the base is involved in recognition and argues for no direct involvement of the discriminator base in thermophilic identity. Thus, the loss provoked by the $U_{73} \rightarrow G$ mutation may result from an antidetermination effect provoked by peculiar chemical groups of G.

(2) Anticodon. Because of the degenerency of the third base of glycine codons, nucleotide 34 in tRNA^{Gly} is not conserved and may not be involved in glycine identity. This is not the case of the conserved C_{35} and C_{36}, which mutations lead to decrease of glycylation efficiency. The losses in activity of 3 orders of magnitude provoked by the $C_{35} \rightarrow U$ and $C_{36} \rightarrow U$ mutations suggest involvement of the NH_2 group at position 4 from the pyrimidine rings in recognition. The most important effects are observed when C_{35} is substituted by A or U and C_{36} by G (losses of 1500 and 5000). Nonconservative $C_{36} \rightarrow G$ substitution affects 12-fold more glycylation efficiency than the $C_{35} \rightarrow G$ mutation, suggesting that the NH_2 or O groups at positions 2 and 6 in G can partially replace essential groups of C_{35} but not of C_{36} for recognition; alternatively G may provoke different antidetermination effects related to distinct protein contextes around nucleotides 35 and 36. The mutations at positions 35 and 36 affect K_Ms by about 1 order of magnitude more than k_{cat}s.

(3) D- and T-Arms and Loops and Variable Region. The conserved G_{10} from the D-arm in the consensus sequence of tRNA^{Gly} is paired with C_{25} or U_{25}. To maintain base pairing, we substituted the G_{10}-C_{25} pair in tRNA$^{Gly(CCC)}$ transcript by A_{10}-U_{25}. This substitution decreases charging efficiency 43-fold, essentially by increasing K_M. Substitutions of the other nucleotides conserved in the consensus sequence affect only poorly glycylation or are without effect. Nonconservative substitution of the C_{50}- G_{64} pair from the T-arm decreases efficiency 4-fold and substitution of U_{16} by C in the D-arm decreases it 5-fold. Substitution of G_{45} by A in the variable region is without significant effect.

Altogether, the glycine identity in *T. thermophilus* tRNA is mainly conferred by the three first base pairs from the acceptor stem and by C_{35} and C_{36} from the anticodon. In the acceptor stem, the first and third base pairs have a prevalent contribution, but only the first pair is conserved in the tRNA^{Gly} of various origins (*18*); G_{10}, U_{16}, and the C_{50} -G_{64} pair from D- and T-arms and loops contribute marginally to identity (Figure 1B). However, direct implication of G_{10} in identity is difficult to prove, since being probably involved in the triple (G_{10}-C_{25})-G_{45} interaction, its substitution can affect the tertiary structure of tRNA and indirectly affect charging efficiency. Finally, it should be pointed out that prediction of implication of peculiar chemical base groups in recognition from kinetic effects of the mutations assumes no alteration of the tertiary structure of tRNA.

Transplantation of the Elements Conferring Glycine Identity in T. thermophilus tRNA^{Asp} and tRNA^{Phe}. To verify the completeness of the glycine identity set, it was transplantated in tRNA^{Asp} and tRNA^{Phe} of *T. thermophilus* and the acquisition of the new specificity by the chimeric tRNAs

Table 2: Kinetic Constants of Aminoacylation of *T. thermophilus* $tRNA^{Gly}$, $tRNA^{Asp}$, and $tRNA^{Phe}$ Variants by the Homologous Gly-tRNA Synthetase and Effect of Temperature[a]

	kinetic parameters								
	37 °C				70 °C				
tRNA	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M ($s^{-1}\,\mu M^{-1}$)	L^d	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M ($s^{-1}\,\mu M^{-1}$)	L^d	$L_{37°C}/L_{70°C}$
$tRNA^{Gly}$ Modified									
$tRNA^{Gly}$ *T. th.*	0.4	0.19	2.1		2	1.14	1.8		
$tRNA^{Gly}$ *T. th.*[b]	0.3	0.4	0.75		2	1.4	1.4		
$tRNA^{Gly}$ Transcripts									
wild-type variants	0.43	0.3	1.43	1.0	1.4	1.5	1.0	1.0	1
Acceptor Stem									
G_1-C_{72} → G_1-U_{72}	0.04	1.25	0.032	45			nd		
A_1-U_{72}	0.025	1.5	0.016	90	0.12	2	0.06	16	5.6
C_1-G_{72}	0.0008	5	0.00016	8940	0.01	25	0.0004	2500	3.5
C_2-G_{71} → U_2-G_{71}	0.12	2.3	0.052	28			nd		
G_2-C_{71}	0.3	1	0.3	5			nd		
G_3−C_{70}→G_3−U_{70}	0.23	1.6	0.143	10			nd		
C_3-G_{70}	0.01	3.3	0.003	476	0.12	4	0.03	33	14
Discriminator Base									
U_{73} → A_{73}	0.4	0.43	0.93	1.5	1.8	0.5	3.6	0.3	5
C_{73}	0.25	0.5	0.5	2.9	1.6	1	1.6	0.6	4.8
G_{73}	0.05	2	0.025	57	0.24	16	0.015	66	0.9
Anticodon Loop									
C_{35} → A_{35}	0.05	62	0.0008	1800			nd		
G_{35}	0.11	31	0.0035	408			nd		
U_{35}	0.02	25	0.0008	1800	0.2	100	0.002	500	3.6
C_{36} → G_{36}	0.01	35	0.0003	4700	0.3	40	0.0075	133	35
U_{36}	0.04	30	0.001	1430			nd		
D-Stem and Loop									
U_{16} → C_{16}	0.24	0,8	0.3	4.8	1	0.58	1.7	0.6	8
G_{10}-C_{25} → A_{10}-U_{25}	0.18	5.5	0.033	43	0.3	10	0.03	33	1.3
Variable Loop									
G_{45}→A_{45}	0.2	0.2	1	1.4	1.14	0,7	1.6	0.6	2.3
T-Stem									
C_{50}-G_{64} → G_{50}–C_{64}	0.2	0.6	0.33	4.3	1	0.5	2	0.5	8.6
Chimeric tRNA									
$tRNA^{Asp→Gly(U73)}$ [c]	0.2	0.3	0.66	2.2	1	0.3	3.3	0.3	7.3
U_{73} → A_{73}	0.12	0.2	0.6	2.4			nd		
C_{73}	0.018	0.22	0.082	17.4			nd		
G_{73}	0.005	0.15	0.034	42			nd		
$tRNA^{Phe→Gly(U73)}$ [c]	0.3	1.14	0.26	5.5	1	2	0.5	2	2.8
U_{73} → A_{73}	0.6	0.62	0.97	1.47			nd		
C_{73}	0.078	1	0.078	18.3			nd		
G_{73}	0.02	1.43	0.014	102			nd		

[a] The k_{cat} and K_M were determined as described in the Experimental Procedures; k_{cat} were determined with saturating concentrations of ligands. [b] *T. thermophilus* $tRNA^{Gly}$ overproduced in *E. coli*. [c] $tRNA^{Asp→Gly(U73)}$ and $tRNA^{Phe→Gly(U73)}$ refer to *T. thermophilus* $tRNA^{Asp}$ and $tRNA^{Phe}$ transplantated with the glycine identity elements and U_{73}. [d] The losses of catalytic efficiencies are expressed as k_{cat}/K_M of the wild-type $tRNA^{Gly}$ transcript over k_{cat}/K_M of the variant; nd, not determined.

was tested. These tRNAs are not charged by *T. thermophilus* GlyRS. They were chosen as a framework for glycine identity elements despite variations in the number of nucleotides in the D-loop (9 and 8 nucleotides in $tRNA^{Asp}$ and $tRNA^{Phe}$ and 7 or 8 in the various $tRNA^{Gly}$) because they contain part of the putative glycine identity set and both contain, like $tRNA^{Gly}$, five nucleotides in the variable loop. The $tRNA^{Asp→Gly}$ and $tRNA^{Phe→Gly}$ variants (Figure 2) were created by introducing in the $tRNA^{Asp}$ and $tRNA^{Phe}$ transcripts the missing nucleotides to complete the set of glycine identity. To prevent a defavorable context for expression of this identity, U_{73} was introduced on the discriminatory positions. These chimeric tRNAs exhibit glycine identity and are charged, respectively, only 5- and 2-fold less efficiently by the thermophilic GlyRS than the $tRNA^{Gly}$ transcript (Table 2). The sligthly lower glycylation efficiency of the $tRNA^{Phe→Gly}$ chimer in which G_{50}-C_{64} was conserved, compared to the $tRNA^{Asp→Gly}$ chimer where the C_{50}→G_{64} pair is present, confirms contribution of this pair to thermophilic glycine identity (Table 2).

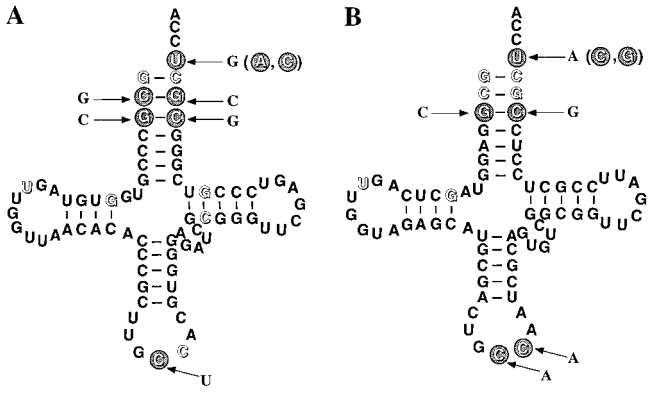

FIGURE 2: Transplantation of the glycine identity elements in *T. thermophilus* $tRNA^{Asp}$ (A) and $tRNA^{Phe}$ (B). The black characters are the nucleotides present in the nonmutated tRNA and the open circled characters are the nucleotides introduced in the chimeric tRNA. The arrows emphasize the substitutions in the chimeric tRNA. The nucleotides in bracket indicate additionnal substitutions of the discriminator base.

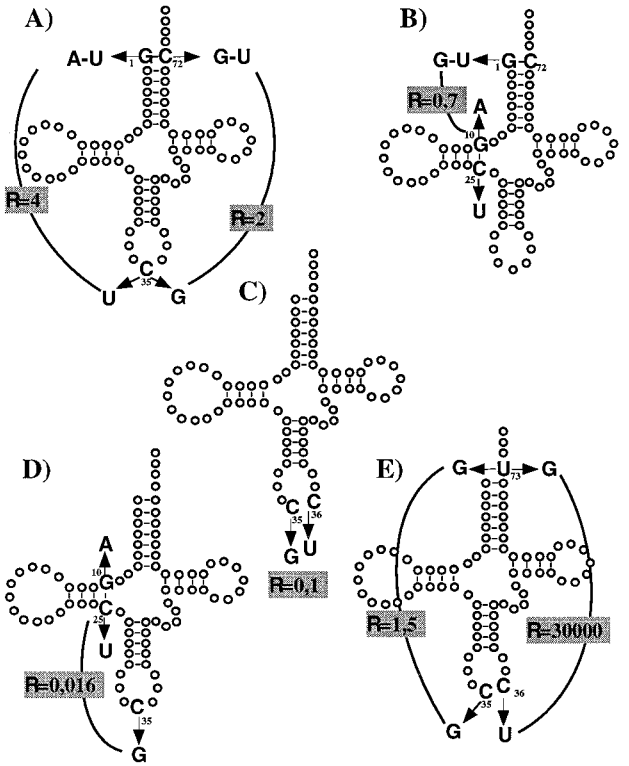

FIGURE 3: Creation of double mutants in *T. thermophilus* $tRNA^{Gly}$. The mutated nucleotides are (A) the first base pair from acceptor arm and C_{35}; (B) G_{10}-C_{25} and G_1-C_{72}; (C) C_{35} and C_{36}; (D) G_{10}-C_{25} and C_{35}; (E) U_{73} and C_{35} or C_{36}. The substituted nucleotides are indicated in black characters and emphasized with arrows. $R = L_{exp(dm)}/L_{cal(dm)}$ (see legend to Table 3).

Nature of the Discriminator Base and Glycylation of the Chimeric $tRNA^{Asp\rightarrow Gly}$ and $tRNA^{Phe\rightarrow Gly}$. A possible role of the discriminator base to glycine identity in the transplantated $tRNA^{Asp\rightarrow Gly}$ and $tRNA^{Phe\rightarrow Gly}$ was searched (Figure 2). Substitution of U_{73} by A and G provokes comparable effects as in $tRNA^{Gly}$, namely A is without significant effect whereas G decreases glycylation efficiency by about 2 orders of magnitude (Table 2). In contrast, substituting U_{73} by C decreases charging efficiency of the chimera ~5 times more than that of $tRNA^{Gly}$ (Table 2) indicating that the effect of C on position 73 is context dependent but not those of A and G.

Interrelation between the Identity Elements in T. thermophilus $tRNA^{Gly}$. Communication between identity elements was searched by analyzing the charging properties of multiple mutants of the $tRNA^{Gly}$ transcript. To measure accurately the charging efficiency of variants substituted at multiple positions, the multiple mutants were created by combining mutations inducing only moderate effects (Figure 3).

The variants combining mutations of C_{35} and of the G_1-C_{72} pair (U_{35} and A_1-U_{72} or G_{35} and G_1-U_{72}, Figure 3A) have losses comparable to the product of the losses provoked by the single mutations (Table 3). Thus, mutations of the first base pair from the acceptor arm and of the central position from the anticodon act independently. Additive effects are also provoked when both G_{10}-C_{25} and G_1-C_{72} pairs are substituted by A_{10}-U_{25} and G_1-U_{72} pairs (Figure 3B). In contrast, when C_{35} and C_{36} from the anticodon are both replaced by G_{35} and U_{36}, the loss of charging is 10-fold lower than expected from additivity (Figure 3C), and when substituting simultaneously C_{35} and the G_{10}-C_{25} pair by G_{35} and A_{10}-U_{25}, it is about 2 orders of magnitude lower than expected from additive effects of the single mutations (Figure 3D, Table 3). This characterizes anticooperativity of the effects. Simultaneous mutations of the discriminator base U_{73} and either C_{35} or C_{36} show distinct effects (Figure 3E). When U_{73} and C_{35} are together substituted by G the loss of charging indicates additive effects of the individual mutations. In contrast, when U_{73} is substituted by G and C_{36} by U, the effects are strongly cooperative, since the loss exceeds 30000-fold that expected from additive effects of the mutations (Table 3).

Effect of Temperature on the Specificity of tRNA Glycylation. The optimal growth temperature of *T. thermophilus* is around 75 °C, and it was proposed that in thermophiles high-temperature promotes specificity of tRNA aminoacylation (*30*). Since the transcript is well charged at 70 °C, the effect of temperature on the specificity of GlyRS was tested by comparing the efficiencies of aminoacylation of variants at 37 and 70 °C (Table 2). Temperature is without significant effect on the charging efficiency of native $tRNA^{Gly(CCC)}$ and its transcript although with k_{cat} and K_M increased each ~5-fold when temperature increases. In contrast, the efficiency of glycylation of most variants and of the $tRNA^{Asp\rightarrow Gly}$ and $tRNA^{Phe\rightarrow Gly}$ chimera decrease 3−8-fold when temperature increases from 37 to 70 °C, and those of the G_3-$C_{70} \rightarrow C_3$-G_{70} and $C_{36} \rightarrow G_{36}$ mutants decrease 35- and 14-fold. These effects are mainly k_{cat} promoted when temperature increases. Only the G_{10}-$C_{25} \rightarrow A_{10}$-$U_{25}$ and $U_{73} \rightarrow G_{73}$ mutants exhibit similar losses at 37 and 70 °C. Since mutation of most identity elements affects less charging efficiency at 70 than 37 °C, specificity of GlyRS decreases when temperature increases. This behavior resembles that described for the aspartate system from *T. thermophilus* (*20*) and may be related to a decreased stability of the complexes between the synthetase and the altered tRNA at high temperature.

The comparable k_{cat}/K_M values at 37 and 70 °C for charging either native $tRNA^{Gly}$ or its transcript indicate that the Michaelis complexes with modified or unmodified tRNA are formed with similar second-order rate constants at both temperatures and that formation of the transition state requires the same activation energy along the temperature scale. Thus, the same mechanistic step determines the rate of tRNA charging at 37 and 70 °C assuming no compensation of different variables, e.g., decrease of affinity for the substrates and increase of k_{cat}. In contrast, the increased efficiency for charging variants, in particular the $C_{36} \rightarrow G_{36}$ and the G_3-$C_{70} \rightarrow C_3$-G_{70} mutants, when temperature increases, agrees with decrease of the activation energy in formation of the transition state as a probable consequence from variation of the nature of the rate-limiting step of tRNA charging.

Cross-Glycylations Involving Partners of Different Origins. Table 4 shows the relative rate constants of cross-glycylations involving $tRNA^{Gly}$ and GlyRSs of different phylae. *E. coli* GlyRS acylates efficiently only eubacterial $tRNA^{Gly}$, e.g., the various *E. coli* isoacceptors and *T. thermophilus* $tRNA^{Gly}$, whereas the various yeast, mammalian, and archaebacterial $tRNA^{Gly}$ isoacceptors are charged 2−3 orders of magnitude slower and yeast $tRNA^{Gly(UCC)}$ is not charged at all. In contrast, *T. thermophilus* GlyRS acylates efficiently eubacterial, archaebacterial, and eukary-

Table 3: Comparison of Effects of Single and Multiple Mutations on tRNA Glycylation by *T. thermophilus* Gly-tRNA Synthetase[a]

	single mutants		double mutants			
	$L_{(sm)}$[b]					
$tRNA^{Gly}$ variant, variant1/variant2	L_1[b]	L_2[b]	$L_{cal(dm)}$[c] (L_1L_2)	$(k_{cat}/K_M)_{exp(dm)}$[d] (s^{-1} μM^{-1})	$L_{exp(dm)}$[e]	R[f]
anticodon/acceptor-stem						
U_{35}/A_1-U_{72}	1800	90	162 000	2×10^{-6}	715 000	4
G_{35}/G_1-U_{72}	400	45	18 000	4×10^{-5}	36 000	2
acceptor stem/D-stem						
G_1-U_{72}/A_{10}-U_{25}	45	45	2000	1×10^{-3}	1400	0.7
anticodon/anticodon						
G_{35}/U_{36}	400	1400	560 000	2×10^{-5}	71 500	0.10
anticodon/D-stem						
G_{35}/A_{10}-U_{25}	400	45	18 000	5×10^{-3}	300	0.016
anticodon/discriminator						
G_{35}/G_{73}	60	400	24 000	4×10^{-5}	36 000	1.5
U_{36}/G_{73}	60	1400	84 000	6×10^{-10}	24×10^8	30 000

[a] The double mutants were created by combining mutations of variants 1 and 2. [b] L_1 and L_2 are the losses of the single mutants [$L_{(sm)}$] of variants 1 and 2 reported in Table 2. [c] $L_{cal(dm)}$ refers to the calculated loss of the double mutant assuming independent effects of the single mutations. [d] $(k_{cat}/K_M)_{exp(dm)}$ refers to the experimental k_{cat}/K_M of the double mutant. [e] $L_{exp(dm)}$ refers to the experimental loss of the double mutant (k_{cat}/K_M of the wild-type (Table 2) over k_{cat}/K_M of the double mutant). [f] R refers to $L_{exp(dm)}/L_{cal(dm)}$.

Table 4: Aminoacylation of $tRNA^{Gly}$ of Various Phylae by *E. coli*, *T. thermophilus*, Yeast, Rabbit Liver, and *M. jannaschii* Gly-tRNA Synthetases

	origin of Gly-tRNA synthetase,				
origin of $tRNA^{Gly}$[a]	*E. coli*	*T. thermophilus*	yeast	rabbit liver	*M. jannaschii*
	Normalized Initial Rates of tRNA Charging[b]				
T. thermophilus					
$tRNA^{Gly(CCC)}$	0.59	1.00	0.13	0.00	0.02
E. coli					
$tRNA^{Gly(GCC)}$	1.00	0.68	0.05	0.00	0.03
$tRNA^{Gly(UCC)}$	1.00	0.73	1.00	0.02	0.03
Yeast					
$tRNA^{Gly(GCC)}$	0.02	0.82	0.50	1.00	0.80
$tRNA^{Gly(UCC)}$	0.00	0.02	1.00	0.50	0.10
$tRNA^{Gly(CCC)}$	0.03	0.40	0.70	0.80	0.50
Mammalian Rabbit Liver					
$tRNA^{Gly}$ (F1)	0.007	0.57	1.08	1.00	0.50
$tRNA^{Gly}$ (F6)	0.003	0.55	1.08	0.85	0.50
Beef Liver					
$tRNA^{Gly}$	0.007	0.20	1.29	1.00	0.50
M. jannaschii					
$tRNA^{Gly}$	0.001	0.15	0.40	0.60	1.00

[a] The $tRNA^{Gly}$ were obtained as described in the Experimental Procedures. F1 and F6 refer to $tRNA^{Gly}$ containing fractions eluted on BD-cellulose. [b] The initial rates of tRNA charging are expressed as relative values of specific activities of charging homologous $tRNA^{Gly}$; relative value: 1 = 1.4, 1.0, 0.24, 0.02, and 0.0001 nmol mg^{-1} min^{-1}, respectively, for *E. coli*, *T. thermophilus*, yeast, rat liver, and *M. jannaschii* GlyRSs.

otic $tRNA^{Gly}$; only yeast $tRNA^{Gly(UCC)}$ is poorly charged. Yeast GlyRS acylates mammalian and archaebacterial $tRNA^{Gly}$ as efficiently as its own isoacceptors; *E. coli* $tRNA^{Gly(UCC)}$ is also well charged by yeast GlyRS, but *E. coli* $tRNA^{Gly(GCC)}$ and *T. thermophilus* $tRNA^{Gly(CCC)}$ are charged about 1 order of magnitude slower. Mammalian GlyRS acylates efficiently all eukaryotic and archaebacterial $tRNA^{Gly}$, but eubacterial $tRNA^{Gly}$s are not or only poorly charged by this enzyme. Finally, *M. jannaschii* GlyRS acylates eukaryotic $tRNA^{Gly}$ as efficiently as its cognate tRNA, except yeast $tRNA^{Gly(UCC)}$, which is charged about 10-fold slower; eubacterial $tRNA^{Gly}$ from *E. coli* and *T. thermophilus* are charged 2 orders of magnitude slower than the homologous tRNA by the archaebacterial GlyRS.

GENERAL DISCUSSION

Glycine Identity in T. thermophilus. In most aminoacylation systems, the posttranscriptional modifications are not involved in tRNA charging and sometimes transcripts are even better charged than the native tRNA, probably because the absence of tRNA modifications facilitates adaptability of the complex and increases its functional competence (*20*, *31*). Exceptions were reported for $tRNA^{Glu}$ and $tRNA^{Lys}$ from *E. coli*, where modifications at the first anticodon position ($mcm^5s^2U_{34}$ and $mnm^5s^2U_{34}$) contribute strongly to k_{cat} (*32*−*34*), for $tRNA^{Ile(GAU)}$ where the isopentenyl group on A_{37} contributes to K_M and k_{cat} (*35*), and for $tRNA^{Ile(CAU)}$, where conversion of C_{34} to lysidine shifts specificity for Met toward Ile (*36*). In the thermophilic $tRNA^{Gly}$, the posttranscriptional modifications are not involved in aminoacylation, since the native tRNA from *T. thermophilus* as well as that overexpressed in *E. coli*, which differs by its modification pattern (unpublished results) and the nonmodified transcript are all aminoacylated with the same kinetic constants. A similar behavior was reported for the glycylation system from *E. coli* (*28*); in yeast, however, the modifications of $tRNA^{Gly(GCC)}$ increase charging efficiency 30-fold, probably by stabilizing the tRNA conformation involved in recognition (*28*).

Major elements conferring identity in *T. thermophilus* $tRNA^{Gly}$ are located in the acceptor stem and in the anticodon

while additional elements located elsewhere in the molecule contribute less strongly to identity; they are G_1-C_{72}, C_2-G_{71}, and G_3-C_{70} from acceptor arm, C_{35} and C_{36} from anticodon, G_{10} and U_{16} from D-stem and loop, and the C_{50}-G_{64} pair from T-arm (Figure 1B). Contributions of the G_1-C_{72} pair from acceptor stem and of C_{35} and C_{36} from anticodon to identity prevail. The loss provoked by the G_{10} mutation is more difficult to interpret since this nucleotide takes part in the triple (G_{10}-C_{25})-G_{45} interaction and in tRNA folding and thus contributes indirectly to tRNA identity. However, direct implication in recognition cannot be excluded as suggested in the yeast aspartate system where the 3D structure of the complex reveals contact of G_{10} with amino acid residues of the loop connecting the anticodon binding domain to the catalytic core of the enzyme (*37*).

The absence of effect when U_{73} is substituted by A either in the $tRNA^{Gly}$ transcript or in the $tRNA^{Asp\rightarrow Gly}$ and $tRNA^{Phe\rightarrow Gly}$ chimera, argue for a nonimplication of the discriminator base in thermophilic glycine identity. Further, efficient glycylation of eukaryotic and archaebacterial $tRNA^{Gly}$ (Table 1) shows that non implication of U_{73} in glycine identity is independent from tRNA context. As a consequence, the important impairment of $tRNA^{Gly}$ charging and that of $tRNA^{Asp\rightarrow Gly}$ and $tRNA^{Phe\rightarrow Gly}$ chimera with G_{73} are provoked by an antideterminant effect by G and not by the privation of essential recognition elements on U. A peculiar group of the purine ring of G absent in A, such as NH_2 at position 2, may alter functional interaction of the amino acid accepting end of tRNA with GlyRS.

Dynamic of the Gly tRNA Synthetase-$tRNA^{Gly}$ Interaction in T. thermophilus. Substitution of the identity elements from the anticodon and D-arm affect K_M about 1 order of magnitude more than k_{cat}, in contrast to substitution of the three first base pairs from acceptor arm that affects either equally both kinetic constants or k_{cat} more than K_M. Thus, binding capacity of tRNA to GlyRS is mostly conferred by C_{35}, C_{36}, and G_{10}, and the kinetic competence of the complex is conferred by elements from the acceptor arm, which likely trigger the orientation of the CCA end in the catalytic center of the synthetase. Insight into the mechanism of functional adaptation of synthetase and tRNA is brought by analysis of multiple mutants. The anticooperative effects provoked when either C_{35} and C_{36} or C_{35} and G_{10} are together substituted indicate coupled binding of the elements from anticodon and G_{10}. In contrast, the additive effects provoked by mutating either G_{10} or U_{35} together with the first base pair suggests independent recognition of the acceptor stem. This contrasts with the yeast and *T. thermophilus* aspartate systems for which anticooperative binding of the anticodon and the acceptor arm were described (*20*, *24*). However, the cooperative effect provoked by U_{36} and G_{73} reveals a peculiar interrelation between a nucleotide from anticodon involved in identity and the discriminator base not involved in identity, which triggers specificity by preventing aminoacylation of noncognate tRNAs which possess a discriminatory G_{73} and a nucleotide different from C_{36} in the anticodon. Synergistic recognition of elements from anticodon and G_{10} and anticooperative effects provoked by mutating the discriminator base and C_{36} may be promoted by conformational changes of the complex mediated by the loop lying the anticodon binding domain to the core enzyme (*13*) as suggested in the aspartate system where residues of this loop were found in close contact with G_{10} from tRNA (*37*).

Glycine Identity in T. thermophilus, E. coli, and Yeast. The glycylation systems of various origins differ by the nature of the elements specifying identity and by their hierarchy in contribution to identity.

Hierarchy of Contribution of the Identity Elements to Glycylation Efficiency. (1) In Homologous Systems. The three first base pairs and C_{35} and C_{36} determine glycylation of *T. thermophilus*, *E. coli*, and yeast tRNAs although contribution to efficiency differs and implication of the discriminator base diverges widely (Table 4). The glycine systems differ first by the amplitude of the effects provoked by mutation of the identity elements. Strong losses occur in yeast compared to the modest ones found in the prokaryotic systems. However, except in a few cases, the losses provoked by mutations in *T. thermophilus* exceed those in *E. coli*. This contrasts with other identities where alteration of determinants often provokes higher losses in prokaryotic than in eukaryotic systems (*20*). The glycine systems differ also by the hierarchy of the determinants. In *E. coli*, contribution of C_{35} from anticodon is the strongest. C_{36} is next in strength, followed by the two first base pairs and the discriminator base, which contribute equally to identity, and the third base pair which has the lowest contribution. In *T. thermophilus*, C_{35} and C_{36} from anticodon are prevalent, followed by the first and third base pairs from acceptor stem; the contribution of the second pair is minor and the discriminator base is not involved in identity. In yeast, contributions of the discriminator base and of C_{35} and C_{36} from the anticodon are prevalent, and strength of the second base pair from the acceptor stem exceeds that of the first and third base pairs. Thus, the second and third base pairs are poorly involved in *E. coli* glycine identity, the second pair in that from *T. thermophilus* and the first pair in that from yeast. Interestingly, the first pair G_1-C_{73} is conserved in all $tRNA^{Gly}$ species except in mitochondria (*18*), whereas the second and third pairs are not conserved in prokaryotic and eukaryotic $tRNA^{Gly}$ [in yeast $tRNA^{Gly(UCC)}$ C_2-G_{71} is substituted by G_2-C_{71} and, in archaea and mycoplasma $tRNA^{Gly(UCC)}$ and in *B. mori* and human $tRNA^{Gly(GCC)}$, G_3-C_{70} is substituted by A_3−U_{70}; ref *18*). This suggests that contribution of the second and third base pairs is context dependent whereas that of the first base pair relates with the role of the discriminator base. Decreased contribution of the discriminator base to identity compensates increased involvement of the first base pair and vice-versa.

(2) In the Heterologous System. In charging of *E. coli* $tRNA^{Gly}$ by *T. thermophilus* GlyRS, contribution of the elements from anticodon is reinforced, whereas that of the elements from acceptor arm is decreased compared to the homologous thermophilic system. Table 4 shows that C_{35} and C_{36} from *E. coli* $tRNA^{Gly}$ constitute prevalent glycine identity elements whereas the G_1-C_{72} and C_2-G_{71} pairs are less important and in contrast to the homologous system, the G_3-C_{70} pair is not involved in glycylation by the thermophilic GlyRS. As in the homologous system, the discriminator base does not contribute to glycylation since A_{73} is without effect on charging but the antideterminant effect exerted by G_{73} is lower.

Contribution of the Discriminator Base to Identity. $tRNA^{Gly}$ of the various phylae differ by the nature of the discriminator base: eubacterial $tRNA^{Gly}$ contains U whereas

Table 5: Comparison of Effect of Substitution of the Identity Elements of $tRNA^{Gly}$ on Charging Efficiency in the Systems of *T. thermophilus*, *E. coli*, and Yeast

position in $tRNA^{Gly}$	glycylation system			
	T. thermophilus homologous[a]	heterologous[b]	*E. coli* homologous[a]	yeast homologous[a]
		L		
Discriminator Base				
$N_{73} \rightarrow A_{73}$	1.5	1	11	>5000
C_{73}	2.9	5	11	>5000
G_{73}	57	14	31	>5000
U_{73}	1		1	
Acceptor Stem				
G_1-$C_{72} \rightarrow G_1$-U_{72}	45	13	11	5
A_1-U_{72}	90	47	43	2
C_1-G_{72}	8940	nd	nd	nd
C_2-$G_{71} \rightarrow U_2$-G_{71}	28	nd	nd	nd
G_2-C_{71}	5	6	10	nd
C_2-A_{71}	nd	4	5	430
U_2-A_{71}	nd	3	9	400
G_3-$C_{70} \rightarrow G_3$-U_{70}	10	2	4	nd
A_3-U_{70}	nd	1	2	16
A_3-C_{70}	nd	1	1	16
C_3-G_{70}	476	nd	nd	17
D-Stem And Loop				
G_{10}-$C_{25} \rightarrow A_{10}$-$U_{25}$	43	nd	nd	nd
$U_{16} \rightarrow C_{16}$	4.8	nd	nd	nd
T-Stem				
C_{50}-$G_{64} \rightarrow G_{50}$-$C_{64}$	4.3	nd	nd	nd
Anticodon				
$C_{35} \rightarrow A_{35}$	1800	nd	nd	nd
G_{35}	408	>5000	1500	>5000
U_{35}	1800	1300	102	1500
$C_{36} \rightarrow A_{36}$	nd	>5000	41	1700
G_{36}	4700	nd	nd	nd
U_{36}	1430	740	21	>5000

[a] Homologous systems involving GlyRS and $tRNA^{Gly}$ from same origin. [b] Heterologous sytem involving *T. thermophilus* GlyRS and *E. coli* $tRNA^{Gly}$. The losses (*L*) of the homologous *T. thermophilus* system are from this study (Table 2) and those of the other systems from Nameky et al. (*28*); nd, not determined.

eukaryotic and archaebacterial $tRNA^{Gly}$ contain A; exceptionally, plant $tRNA^{Gly}$ contain C (*18*). Wide divergence of contribution of this base to identity superimposes to its structural heterogeneity since it can either be absent from the set of identity or contribute to charging at various extents and even determine species-specific glycylation. Substitution of U_{73} by A or C in *E. coli* $tRNA^{Gly}$ decreases glycylation efficiency by 1 order of magnitude and substitution by G 3-fold more (Table 5). Thus, this base contributes to the recruitment and charging of $tRNA^{Gly}$ in the eubacterial system although to a low extent. However, this contrasts with the unability of the synthetase to charge efficiently eukaryotic and archaebacterial $tRNA^{Gly}$ possessing A_{73} (Table 4) and with acquisition of its ability to charge the eukaryotic-accepting microhelix only after substitution of A_{73} by U (*16*, *17*), indicating strong contribution of the discriminator base. Thus, eubacterial species specificity of glycylation is promoted by the nature of the discriminator base but depends on tRNA context. In contrast, GlyRSs from higher eukaryotes and archaebacteria exhibit species specificity for $tRNA^{Gly}$ possessing A_{73}. Mammalian GlyRS aminoacylates efficiently only eukaryotic and archaebacterial $tRNA^{Gly}$ (Table 4) and becomes able to charge the accepting *E. coli* microhelix only after substitution of U_{73} by A (*16*, *17*) whereas *M. jannaschii* GlyRS charges efficiently only archaebacterial and eukaryotic $tRNA^{Gly}$ species (Table 4). Yeast GlyRS, which charges efficiently *T. thermophilus* $tRNA^{Gly(CCC)}$ and *E. coli* $tRNA^{Gly(UCC)}$, exhibits a relaxed species specificity which agrees with a moderate role of the discriminator base in glycylation (Table 4). Nevertheless, this role of A_{73} is also context dependent since its substitution in yeast $tRNA^{Gly(GCC)}$ by any other nucleotide decreases drastically charging by the cognate GlyRS (Table 5). Finally, efficient charging by *T. thermophilus* GlyRS of eukaryotic $tRNA^{Gly}$ and of the cognate $tRNA^{Gly}$ in which U_{73} is substituted by A or C (Table 4) excludes participation of the discriminator base in thermophilic glycine identity, but it is involved in selection of tRNA since G_{73} hinders drastically glycylation.

Purines are the most frequently found at the discriminator position in tRNAs. In *E. coli*, only $tRNA^{Gly}$ and $tRNA^{Cys}$ contain U_{73}. Interestingly, in $tRNA^{Cys}$ the discriminator base is also involved in identity (*38*, *39*), and this tRNA contains further the G_1-C_{72} and C_{35} glycine identity elements. However, $tRNA^{Cys}$ contains the 2-methylthio-*N*6-isopentenosyl modification in A_{37} which may act as an antideterminant for glycylation and improve specificity of cysteinylation.

The glycylation system is unique by the structural and functional variabilities of the discriminator base. In most aminoacylation systems, nucleotide 73 of tRNA is conserved in the various species, and except for *E. coli* glutamylation, its contribution to identity often prevails (*40*−*42*). Only exceptionally, this nucleotide differs in prokaryotic and eukaryotic tRNA isoacceptors or is not conserved in the isoacceptors from a given species: prokaryotic $tRNA^{His}$ contain C_{73} and the eukaryotic ones A_{73}; the $tRNA^{Arg}$ isoacceptors from *E. coli* contain G_{73} or A_{73} and those from yeast G_{73} or C_{73}; prokaryotic $tRNA^{Lys}$ contain A_{73}, archaebacterial $tRNA^{Lys}$ G_{73}, and eukaryotic $tRNA^{Lys}$ G_{73} or U_{73}. Contribution of the discriminator base to identity differs in

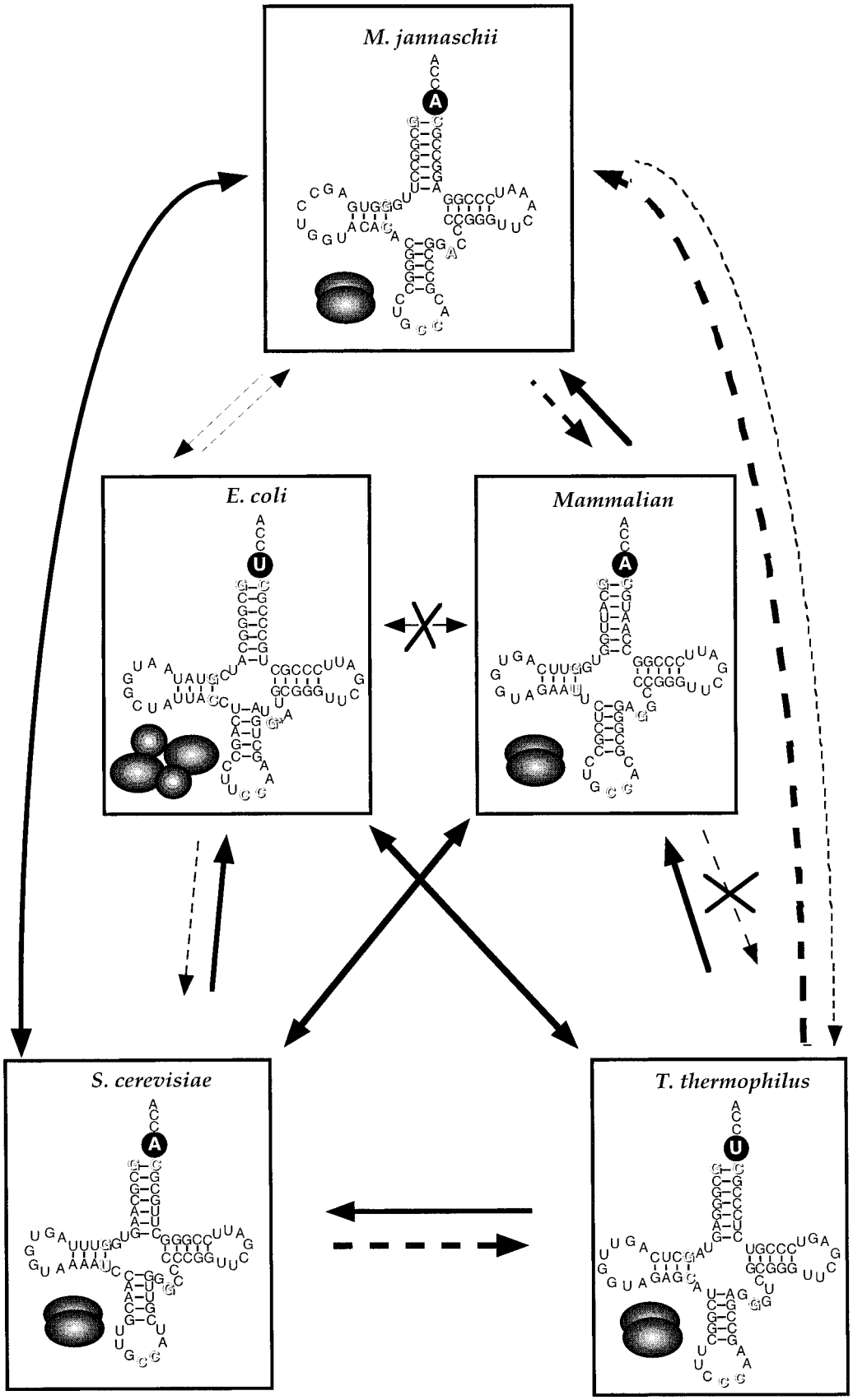

FIGURE 4: Schematic representation of the functional interrelation between the glycylation system from *T. thermophilus* with those from other eubacteria (*E. coli*), archaebacteria (*M. jannaschii*), and eukaryotes (mammalian and yeast) and interrelations between the glycylation systems of these organisms. The $tRNA^{Gly}$ the most efficiently aminoacylated by the heterologous GlyRS and the dimeric or tetrameric structure of the enzyme are shown. The tRNA sequences are without the posttranscriptional modifications; the discriminator bases are black boxed white characters; the positions conserved in the various $tRNA^{Gly}$ except those present in all tRNAs are open characters; (*) G_{45} can exceptionally be substituted by U. The arrows give a quantitative estimate of the efficiency of aminoacylation of $tRNA^{Gly}$ by the heterologous GlyRS. The bold arrows represent efficient charging, the bold dotted arrows decreased charging efficiency, the thin dotted arrows poor charging efficiency and the crossed dotted arrows indicate absence of significant charging.

these systems, but it never determines species-specific aminoacylation as in the glycine system. Nucleotide 73 is strongly implicated in histidylation in *E. coli* (*43*, *44*) but only weakly in yeast (*45*, *46*); arginine identities from yeast and *E. coli* do not involve the discriminator base, albeit in *E. coli*, purines are preferred over pyrimidines (*34*, *47*) and *E. coli* LysRS prefers A_{73} (*34*) whereas human LysRS charges well all $tRNA^{Lys}$ variants of nucleotide 73 (*48*). This suggests emergence of the aminoacylation specificities by distinct processes and agrees with different mechanisms of evolution of these systems.

Evolution of the tRNA Glycylation System. The structural heterogeneity of GlyRSs and the variability of the discriminator base in $tRNA^{Gly}$ determine functional divergences of the glycylation systems (Figure 4). The tetrameric eubacterial GlyRS acylates efficiently only eubacterial $tRNA^{Gly}$ with a discriminatory U whereas dimeric mammalian GlyRS acylates eukaryotic and archaebacterial $tRNA^{Gly}$ with a discriminatory A. The dimeric archaebacterial GlyRS acylates eukaryotic $tRNA^{Gly}$ as efficiently as the archaebacterial one but much less the eubacterial $tRNA^{Gly}$. In contrast, the dimeric *T. thermophilus* GlyRS is totally deprived of species specificity since it does not distinguish the heterologous tRNA species and acylates tRNA with either A or U at the discriminatory position. Finally, yeast GlyRS possesses a relaxed specificity by acylating various prokaryotic $tRNA^{Gly}$ species as efficiently as eukaryotic and archaebacterial $tRNA^{Gly}$. These peculiarities argue for a complex evolution of the glycylation system. Species specificity of the *E. coli* and mammalian GlyRSs agrees with coevolution of synthetase and tRNA which altered the two partners to preserve specificity. This property was probably acquired late since organisms in the lower part of the evolutionary scale, such *T. thermophilus* and yeast, are either deprived of species specificity or have a relaxed one. Ancestral GlyRS was probably dimeric and deprived of specificity for the discriminatory U_{73} of $tRNA^{Gly}$. A_{73} appeared in the ancester of archaebacteria which evolved toward eukaryotes, but in this phylae specificity for A_{73} was fully acquired only in higher eukaryotes after alteration of the dimeric synthetase. The eubacterial system conserved U_{73} and evolved independently by altering differently the synthetase which acquired a tetrameric structure and specificity for U_{73}. The elements which determine efficient glycylation in the homologous systems, namely G_1-C_{72}, G_{10}, C_{35}, and C_{36}, constitute probably the ancestral identity set which was preserved during evolution as well as the triad G_{10}-$C(U)_{25}$-G_{45} involved in tRNA folding (Figure 4). The existence of a dimeric GlyRS in *T. thermophilus* deprived of specificity for the discriminator base argues for an ancestral origin of the thermophilic glycine system and suggests that *T. thermophilus* is located at the borderline between eubacteria and archaebacteria. This agrees with the role of *T. thermophilus* as a link in evolution also suggested on the basis of the properties of its aspartylation system (*11*, *49*).

ACKNOWLEDGMENT

We thank Pr. D. Söll and co-workers (Yale University, New Haven, CT) for generous gifts of crude protein extracts and unfractionated tRNA from *M. jannaschii* and R. Giegé (IBMC, Strasbourg) for constant support and critical reading of the manuscript.

REFERENCES

1. Schimmel P. R., and Söll D. (1979) *Annu. Rev. Biochem. 48*, 602−648.
2. Lapointe J., and Giegé, R. (1991) in *Translation in Eukaryotes* (Trachsel, H., Ed.) pp 35−69, CRC Press Inc., Boca Raton, FL.
3. Kisselev, L. L., and Wolfson, A. D. (1994) *Prog. Nucleic Acids Res. 48*, 86−144.
4. Meinnel, L. L., Mechulam, Y., and Blanquet S. (1995) in *tRNA: Structure, Biosynthesis and Function* (Söll, D., and RajBhandary, U., Ed.) pp 251−292, American Society of Microbiology, Washington DC.
5. Cusack, S., Berthet-Colominas, C., Härtlein, M., Nassar, N., and Leberman, R. (1990) *Nature 347*, 249−255.
6. Eriani, G., Delarue, M., Poch, O., Gangloff, J., and Moras, D. (1990) *Nature 347*, 203−206.
7. Ibba, M., Morgan, S., Curnow, A. W., Pridmore, D. R., Vothknecht, U. T., Gardner, W., Lin, W., Woese, C. R., and Söll, D. (1997) *Science 278*, 1119−1122.
8. Bult et al. (1996) *Science 273*, 1058−1073.
9. Curnow, A. W., Hong, K. W., Yuan, R., Kim, S. I., Martins, O., Winkler, W., Henkin, T. M., and Söll, D. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 11819−11826.
10. Curnow, A. W., Ibba, M., and Söll, D. (1996) *Nature 382*, 589−590.
11. Becker, H. D., and Kern, D. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 12832−12837.
12. Mazauric, M.-H., Reinbolt, J., Lorber, B., Ebel, C., Keith, G., Giegé, R., and Kern, D. (1996) *Eur. J. Biochem. 241*, 814−826.
13. Logan, D. T., Mazauric, M.-H., Kern, D., and Moras, D. (1995) *EMBO J. 14*, 4156−4167.
14. Mazauric, M.-H., Keith, G., Logan, D., Kreutzer, R., Giegé, R., and Kern, D. (1997) *Eur. J. Biochem. 251*, 744−757.
15. Freist, W., Logan, D. T., and Gauss, D. H. (1996) *Biol. Chem. Hoppe-Seyler 337*, 343−356.
16. Shiba, K., Schimmel, P., Motegi, H., and Noda, T. (1994) *J. Biol. Chem. 269*, 30049−30055.
17. Hipps, D., Shiba, K., Henderson, B., and Schimmel, P. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 5550−5552.
18. Sprinzl, M., Steegborn, C., Hübel, F., and Steinberg, S. (1996) *Nucleic Acids Res. 24*, 68−72.
19. Francklyn, C., Shi, J.-P., and Schimmel, P. (1992) *Science 25*, 1121−1125.
20. Becker, H., D., Giegé, R., and Kern, D. (1996) *Biochemistry 35*, 7442−7458.
21. Tabor, S., and Richardson, C. C. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 4767−4771.
22. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular cloning: a laboratory manual*, 2nd ed., Cold spring Harbor Laboratory, Plainview, NY.
23. Fersht, A. R. (1985) *Enzymes, Structure and Mechanism*, Freeman, San Francisco.
24. Pütz, J., Puglisi, J. D., Florentz, C., and Giegé, R. (1993) *EMBO J. 12*, 2949−2957.
25. Midvan, A. S., Weber, D. J., and Kuliopulos, A. (1992) *Arch. Biochem. Biophys. 294*, 327−340.
26. Vogel, D. W., Hartmann, R. K., Kröger, B., Ulbrich, N., and Erdmann, V. A. (1987) *Biochem. Int. 14*, 167−175.
27. Weisshaar, M., Ahmadian, R., Sprinzl, M., Sato, M., Kushiro, A., and Tomita, K. (1990) *Nucleic Acids Res. 18*, 1902.
28. Nameki, N., Tamura, K., Asahara, H., and Hasegawa, T. (1997) *J. Mol. Biol. 268*, 640−647.
29. McClain, W. H., Foss, K., Jenkins, R. A., and Schneider, J. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 6147−6151.
30. Kumasawa, Y., Himeno, H. Miura, K., and Watanabe, K. (1991) *J. Biochem. 109*, 421−427.
31. Pütz, J., Puglisi, J., D., Florentz, C., and Giegé, R. (1991) *Science 252*, 1696−1699.
32. Kern, D., and Lapointe, J. (1979) *Biochemistry 18*, 5819−5826.
33. Sylvers, L. A., Rogers, K. C., Shimizu, M., Ohtsuka, E., and Söll, D. (1993) *Biochemistry 32*, 3836−3841.

34. Tamura, K., Himeno, H., Asahara, H., Hasegawa, T., and Shimizu, M. (1992) *Nucleic Acids Res. 20*, 2335−2339.
35. Nureki, O., Niimi, T., Muramatsu, T., Kanno, H., Kohno, T., Florentz, C., Giegé, R., and Yokoyama, S. (1994) *J. Mol. Biol. 236*, 710−724.
36. Muramatsu, T., Nishikawa, K., Nemoto, F., Kuchino, Y., Nishimura, S., Miyazawa, T., and Yokoyama, S. (1988) *Nature 336*, 179−181.
37. Cavarelli, J., Rees, B., Ruff, M., Thierry, J.-C., and Moras, D. (1993) *Nature 362*, 181−184.
38. Pallanck, L., Li, S., and Schulman, L. H. (1992) *J. Biol. Chem. 267*, 7221−7223.
39. McClain, W., H. (1993) *J. Biol. Chem. 268*, 19398−19402.
40. Giegé, R., Sissler, M., and Florentz, C. (1998) *Nucleic Acids Res. 26*, 5017−5035.
41. Sylvers, L. A., Rogers, K. C., Shimizu, M., Ohtsuka, E., and Söll, D. (1993) *Biochemistry 32*, 3836−3841.
42. Sekine, S., Nureki, O., Sakamoto, K., Niimi, T., Tateno, M., Go, M., Kohno, T., Brisson, A., Lapointe, J., and Yokoyama, S. (1996) *J. Mol. Biol. 256*, 685−700.
43. Himeno, H., Hasegawa, T., Ueda, T., Watanabe, K., Miura, K., and Shimizu, M. (1989) *Nucleic Acids Res. 17*, 7855−7863.
44. Yan, W., and Francklyn, C. (1994) *J. Biol. Chem. 269*, 10022−10027.
45. Rudinger, J., Florentz, C., and Giegé, R. (1994) *Nucleic Acids Res. 22*, 5031−5037.
46. Nameki, N., Asahara, H., Shimizu, M., Okada, N., and Himeno, H. (1995) *Nucleic Acids Res. 23*, 389−394.
47. Sissler, M., Giegé, R., and Florentz, C. (1996) *EMBO J. 15*, 5069−5076.
48. Shiba, K., Stello, T., Motegi, H., Noda, T., Musier-Forsyth, K., and Schimmel, P. (1997) *J. Biol. Chem. 272*, 22809−22816.
49. Becker, H. D., Reinbolt, J., Kreutzer, R., Giegé, R., and Kern, D. (1997) *Biochemistry 36*, 8785−8797.

BI991392T