Articles

Reconstruction of Quaternary Structures of Class II tRNA Synthetases by Rational Mutagenesis of a Conserved Domain†

Lluís Ribas de Pouplana‡ and Paul Schimmel*,‡

Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Received July 23, 1997; Revised Manuscript Received September 19, 1997⊗

ABSTRACT: Class II tRNA synthetases have long been known to have quaternary structures of α, α_2, $\alpha_2\beta_2$, and α_4, depending on the amino acid specificity and the organism from which the synthetase was isolated. Even the quaternary structures of enzymes for the same amino acid show variations in evolution. The basis for these variations has not been understood. We report here that sequence manipulations of a structural motif (motif 1) characteristic of all class II tRNA synthetases can generate most of the evolutionary diversity of quaternary forms of class II synthetases. Thus, the principles elucidated here for quaternary structure assembly may be general.

Aminoacyl-tRNA synthetases evolved to give rise to two distinct structural families (class I and class II), which were initially defined in terms of conserved sequence motifs (*1*) and, later, by the strict conservation of the active site fold in each class (*2*−*5*). Both families contain 10 enzymes that specifically recognize one amino acid and its corresponding tRNA and catalyze the covalent attachment of these two substrates to yield aminoacyl-tRNA. The specificity of each enzyme for its cognate substrates is essential for decoding genetic information. The enzymes achieve their specificities through variations in the amino acid binding site pocket and through the incorporation of idiosyncratic domains that present specific recognition sites for the cognate tRNAs (*4*, *6*, *7*).

Another source of variability among enzymes in the same class comes from their quaternary structures. Class I synthetases form monomeric or dimeric structures (*2*, *8*−*10*). Class II enzymes are more variable, and instances of monomeric, homodimeric, homotetrameric, and heterotetrameric ($\alpha_2\beta_2$) arrangements have been described (*6*, *11*−*16*). In some cases, such as alanyl-, glycyl-, and phenylalanyl-tRNA synthetases (AlaRS, GlyRS, and PheRS), different quaternary forms can be found across species, indicating that the evolution of quaternary structure can take place in the context of a well-defined protein architecture (*15*). The mechanism that allows these enzymes to evolve into new quaternary arrangements has not been understood.

The three-dimensional structures of six class II synthetases have been solved by crystallography methods (*3*, *5*, *14*, *17*−*19*). All of these are homodimeric enzymes with the exception of the heterotetrameric PheRS from *Thermus thermophilus*. These structures provide detailed information on the dimer interface of the enzymes. The dimer interface is formed by a consensus motif of the class II family known as motif 1. This motif is a helix−loop−strand that is followed by a large loop that is designated here as L1 (Figure 1). The whole region packs against its equivalent in the other monomer, connecting the two active sites of the dimeric proteins. This interaction extends over all the secondary structure elements of motif 1, including the loop L1 [which follows β-strand S1 (Figure 1)] where most of the structural variability in motif 1 occurs (*3*, *5*). As a result of this arrangement, each motif 1 contributes to the architecture of both active sites of the dimer as well as to the dimer interface. Cross-subunit contacts between motif 1 residues and the acceptor stem of the tRNA molecule bound by the second subunit can be observed in the crystal structure of the dimeric AspRS (*20*). This arrangement gives a functional rationale for the dimeric structure.

In the case of PheRS, a striking example of subunit interface conservation between two diverged (α and β) subunits is seen (*14*). Early work with the heterotetrameric *Escherichia coli* glycyl-tRNA synthetase suggested that this heterotetramer was operationally a homodimer, if one considered the α and β chains to be part of one polypeptide (*21*, *22*). This expectation was fulfilled in the subsequent discovery of *E. coli* glycyl-tRNA synthetase-like glycine enzymes that were α_2 dimers (*18*, *23*−*25*). Thus, the heterotetrameric species can be viewed as a variation on the homodimer structure.

Despite very low sequence conservation in this region among the enzymes of class II, the overall motif 1 interaction has been preserved throughout evolution. The question remains open as to how motif 1 evolved to allow for quaternary diversity without compromising the architecture of the active site. While the *T. thermophilus* enzyme has the α_2 dimer structure found for most class II enzymes (*26*), the *E. coli* enzyme is an α_4 tetramer (*27*) and the eukaryote enzyme is a monomer (*16*, *28*). Currently, no crystal-

† This work was supported by Grant GM15539 from the National Institutes of Health and by the National Foundation for Cancer Research. Ll.R.P. was supported by a postdoctoral fellowship from the Generalitat de Catalunya.

* To whom correspondence should be addressed.

‡ Present address: The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037.

⊗ Abstract published in *Advance ACS Abstracts,* November 15, 1997.

S0006-2960(97)01788-1 CCC: $14.00

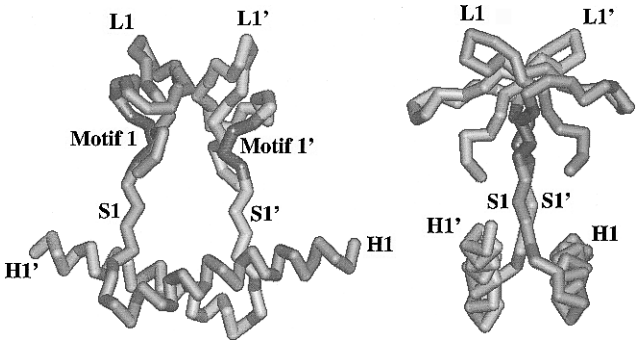

FIGURE 1: Backbone representation of the interactions of motif 1 regions at the dimerization surface of yeast AspRS (*5*). The backbones of subunits 1 and 2 are colored red and rust, respectively. The regions of the structure that harbor the consensus motif 1 sequences are labeled and colored in green. The three distinct regions of secondary structure of the domain are labeled accordingly: H1 and H1′, helix 1 from subunits 1 and 2, respectively; S1 and S1′, strand 1 from subunits 1 and 2, respectively; L1 and L1′, loop 1 from subunits 1 and 2, respectively.

lographic information is available for monomeric or homotetrameric enzymes which are represented by eukaryote (α monomer) or bacterial alanyl-tRNA synthetases (α_4 tetramer). All alanine enzymes contain the three class II consensus motifs (motifs 1, 2, and 3) at their N-terminal ends (between residues 1 and 250), which must be incorporated into class II active sites (*1, 29*−*31*). The *E. coli* enzyme is a tetramer of 875 amino acid polypeptides. Previous reports showed that the tetrameric structure can be disrupted into active monomers, if a C-terminal region of the enzyme is deleted (*12*). Such fragments maintain wild-type levels of activity and specificity with RNA microhelix substrates, which recapitulate the structure of the tRNA acceptor stem (*32*). Thus, the N-terminal fragments of AlaRS contain all the necessary elements for active site formation and tRNA recognition, despite the disruption of the tetramer interface. The removal of this element for oligomerization did not result in the formation of detectable dimer complexes. For example, a large deletion resulting in a 461 amino acid N-terminal fragment remained monomeric while maintaining significant activity both *in vivo* and *in vitro* (*12*).

That such fragments are monomeric is surprising given that they retain motif 1 at the N-terminus of the sequence, where it may be expected to be involved in oligomerization contacts. A possible explanation is that the idiosyncratic C-terminal region has a critical role in dimer formation and that the motif 1 interactions are too weak by themselves to sustain the dimeric structure and yet are strong enough to assemble a dimer of dimers to form a tetramer. This explanation is consistent with motif 1 of AlaRS being highly diverged from the motif 1 element of some of the other members of the class II tRNA synthetases (*1, 29*).

These features of *E. coli* alanyl-tRNA synthetase provide an opportunity to manipulate motif 1 in this enzyme and potentially find principles that are general for the entire class II tRNA synthetase family. Our results show that motif 1 is an important regulator of quaternary structure evolution in class II synthetases, because specific variations of its sequence alone, in the context of the *E. coli* protein, generate α, α_2, and α_4 quaternary forms. Thus, variations in this motif provide a structural explanation for the variability in quaternary structures observed among members of the class II synthetases.

MATERIALS AND METHODS

Sequence and Structure Analysis. All class II tRNA synthetase sequences were obtained from the Swissprot database (*33*). The predictions of coiled-coil regions were performed with the program COILS and were based on the presence in the sequence of heptad-repeat signature motifs (*34*). Sequence alignments were first performed using CLUSTALW (*35*) and then edited manually. The three-dimensional coordinates used were from the Protein Database at Brookhaven (*36*). Structural superimpositions were done manually using QUANTA (Molecular Simulations, Waltham, MA), and structure-based alignments were done manually on the basis of the result of the superimpositions.

Mutagenesis and in Vivo Complementation Assays. The gene *alaS* coding for *E. coli* alanyl-tRNA synthetase [harbored by plasmid pBSKS$^+$_*alaS* (*30*)] was mutagenized using the uracil-incorporation method (*37*). The region around the mutation site was sequenced using dideoxy sequencing methods (*38*) and subcloned back into the original plasmid to ensure that no other mutations were incorporated into the gene.

The complementation assays were carried out as described (*12*) using the *alaS* null strain W3110 (*lacI*^{q} *recA* Δ1 Kan^r *alaS*Δ2) maintained by plasmid pMJ901 (Tet^r marker) that expresses full-length AlaRS and which has a temperature-sensitive replicon (*30, 49*). At the elevated temperature of 42 °C, plasmid pMJ901 is lost and the cells cease to grow. When an additional plasmid is introduced that encodes for a mutant AlaRS, growth is rescued at 42 °C only if the mutation does not significantly disrupt enzyme activity.

Protein Expression and Purification. In order to optimize levels of protein expression and improve purification efficiency, the gene (*alaS*) coding for full-length AlaRS and a fragment of the gene coding for the N-terminal 459 amino acids (*alaS*_*459*) were fused to a C-terminal six-histidine extension in the context of plasmid pQE-70 (Amp^r marker) (Quiagen, Chatsworth, CA) to generate plasmids pQE-*alaS*-6H and pQE-*alaS*_459-6H.

Plasmid pQE-*alaS*-6H was constructed by introducing an *Sph*I site at the 5′ end of the gene and a *Bam*HI site at the 3′ end by site-directed mutagenesis (*37*). The 5′ end modifications involved mutating the two bases 5′ to the ATG start codon from TT to GC and the base immediately 3′ to the ATG codon from A to C. This last change resulted in a sequence change at position 2 from a serine codon (AGC) to an arginine codon (GCG). These three mutations generated a new *Sph*I restriction site (GCATGC).

The 3′ modifications in the *alaS* gene substituted the original stop codon (TAA) with a glycine codon (GGA) and mutated the three following bases from TAT to TCC, resulting in the introduction of a new *Bam*HI restriction site (GGATCC). The fragment containing the modified *alaS* gene (an *Sph*I−*Bam*HI restriction fragment) was purified by standard techniques (*40*). The fragment was then subcloned into plasmid pQE-70 that was previously digested with the same restriction enzymes. This procedure resulted in the fusion of the *alaS* gene to a 3′ end extension coding for 10 residues (Gly-Ser-Arg-Ser-His-His-His-His-His-His) followed by a TAA stop codon.

Plasmid pQE-*alaS*_459-6H was created to fuse the N-terminal 459 region of AlaRS to the same six-histidine extension used for the whole enzyme. In order to truncate

the gene at codon 459, the two original codons (for Gly459 and Tyr460) (GGCTAT) were mutated to GGATCC, thereby introducing a new *Bam*HI site at this position. These changes maintained a Gly codon at position 459 but changed the next codon to Ser. The *Sph*I and *Bam*HI fragment containing the region of the *alaS* gene coding for the N-terminal 459 residues of AlaRS was isolated by gel electrophoresis and elution (*40*). The fragment was then subcloned into plasmid pQE-70, which had been digested with the same restriction enzymes. This procedure resulted in the fusion of the desired fragment to a 3′ end extension coding for nine residues (Ser-Arg-Ser-His-His-His-His-His-His) followed by a stop codon TAA.

The expression product of the *alaS* gene fused to the described C-terminal six-His extension was designated as AlaRS-6H, and the expression product of the *alaS* gene fragment fused to the C-terminal extension was designated as AlaRS$_{-}$459-6H. Wild-type and mutant AlaRS-6H were purified from the W3110 *alaS* null strain (*41*) that had been transformed with wild-type or mutant versions of plasmids pQE-*alaS*-6H. The genomic copy of *alaS* of this strain is disrupted with a Kan^r marker, and the cells contain a copy of the *lacI*^{q} gene that ensures tight control of the lac promoter of pQE-70-based plasmids in the absence of the IPTG strain (*41*). The growth of the W3110 *alaS* null strain was maintained by plasmid pT461 (Tet^r marker) which encodes the active N-terminal 461 amino acid fragment of AlaRS (*42*). Because the pT461-expressed enzyme is a monomer that lacks the six-His tag and has a polypeptide size that is roughly half that of the full-length enzyme, the pQE-*alaS*-6H-expressed enzymes were well resolved from the 461mer.

In the case of AlaRS$_{-}$459-6H, this enzyme was purified from *E. coli* strain TG1 (*supE hsd* Δ*5 thi* Δ(*lac-pro*AB) F′ [*traD*36 *pro*AB+ *lacI*^{q} *lac*ZΔM15]) that had been transformed with pQE-*alaS*$_{-}$459-6H. In this system a second plasmid, pREP4 (Quiagen), containing a copy of the *lacI*^{q} gene and a Kan^r marker, was used in conjunction with the pQE-based plasmids to ensure tight control of the lac promoter of pQE in the absence of IPTG. This enzyme was used as a standard in our molecular weight determinations.

The purification protocol was essentially the same for both full-length and truncated proteins. Cells expressing AlaRS-6H proteins were grown in LB medium (5 g/L NaCl, 5 g/L yeast extract, 10 g/L casein) containing 50 μg/mL ampicillin, 25 μg/mL kanamycin, and 12 μg/mL tetracycline. Cells containing the pQE$_{-}$459-6H plasmid were grown in LB medium containing 50 μg/mL ampicillin and 25 μg/mL kanamycin. Typically 2 L of culture was grown at 37 °C to an optical density of ∼0.9 at 600 nm. IPTG was then added to a final concentration of 1 mM, and after 4−6 h of further growth cells were harvested by centrifugation. Cells were resuspended in 40 mL of lysis buffer [50 mM $NaHPO_3$, 500 mM NaCl, 10 mM imidazole (pH 7.4), 10% glycerol] and lysed in a French press. The lysed crude extract was then centrifuged at 30K rpm for 1 h at 4 °C, and the resulting supernatant was diluted 4-fold in lysis buffer.

This solution was loaded at 2 mL/min into a 15 mL column containing Niquel−nitrilotriacetic acid (Ni−NTA) attached to a Sepharose resin (Quiagen). The column had previously been equilibrated with lysis buffer. The column was washed with ∼100 column volumes of wash buffer [identical to lysis buffer except for an increased concentration of imidazole (30 mM)], with a flow of 1 mL/min.

The proteins bound to the column after the washing step were eluted using a gradient of imidazole (50−500 mM) in lysis buffer. Fractions (2 mL) were collected and analyzed by SDS−PAGE and staining with Coomassie blue (*40*). This method typically yields between 15 and 20 mg/L AlaRS-6H proteins and 70−100 mg/L AlaRS$_{-}$459-6H, with purity levels of 90−95%. The purified proteins were concentrated to ∼5 mg/mL in 50% glycerol and stored at −20 °C.

Kinetic Analysis. Alanine adenylate synthesis was measured at 25 °C in a thermostated water bath in 100 mM Tris-HCl (pH 8.0), 2 mM ATP, 2 mM pyrophosphate, 10 mM KF, 2 mM alanine, 10 mM β-mercaptoethanol, and 5 mM $MgCl_2$ as previously described (*43*). Aminoacylation activity was measured at 25 °C in a thermostated water bath in 50 mM Hepes (pH 7.5), 20 μM alanine, 4 mM ATP, 20 mM KCl, 10 mM $MgCl_2$, 20 mM β-mercaptoethanol, and 0.1 mg/mL bovine serum albumin as previously described (*44*, *45*). Microhelix RNA substrates were synthesized chemically using published procedures (*46*). All phosphoramidites used were from Chemgenes (Waltham, MA). To ensure proper folding, RNA microhelix substrates were heated to 65 °C in water and cooled in the absence of magnesium prior to addition to the aminoacylation reaction.

Enzyme concentrations were typically 10 nM as determined by UV absorbance at 280 nm and Bradford assays (*40*). Mutant enzymes were assayed at concentrations (10−50 nM) that gave aminoacylation rates more than 10-fold over background values. As a control, we confirmed that the six-His extension had no effect upon the kinetic characteristics of AlaRS or the 459mer fragment with respect to adenylate formation or microhelix and full-length tRNA charging.

Gel Filtration Chromatography. Gel filtration analysis was done using a Superose-6 FPLC gel filtration column (Pharmacia, Uppsala, Sweden) in wash buffer. Typically 100 μL of purified protein was loaded into the column at a flow speed of 0.25 mL/min. The OD_{280} of the column eluate was recorded, and 1 mL fractions were collected for SDS−PAGE and kinetic analysis. The relationship between molecular weight and elution time was calibrated with molecular weight standards (Pharmacia).

RESULTS

Sequence Analysis of AlaRS Sequences and Class II Structures. There are several examples of class II tRNA synthetases whose quaternary structures differ from species to species. For example, glycyl-tRNA synthetase generally is dimeric but is an $\alpha_2\beta_2$ heterotetramer in *E. coli* and in some other bacteria (*18*, *21*, *47*). Phenylalanyl-tRNA synthetase is also an $\alpha_2\beta_2$ heterotetramer in many organisms, but in the mitochondria of *Saccharomyces cerevisiae* it is a monomer (*48*). On the other hand, alanyl-tRNA synthetase is a monomer, homodimer, or homotetramer, depending on the organism (*vide supra*). The reasons for these cross-species differences are unclear. We reasoned that, given the role of motif 1 in dimerization shown by the crystal structures of class II enzymes, evolutionary changes in quaternary structure may be reflected in sequence changes in this region.

Multiple sequence alignments showed that the most conserved residues in the vicinity of the motif 1 consensus sequence encompassed the segment region from P34 to K55 in *E. coli* alanyl-tRNA synthetase (Figure 2). Based on the position of the conserved P34 in the motif 1 consensus

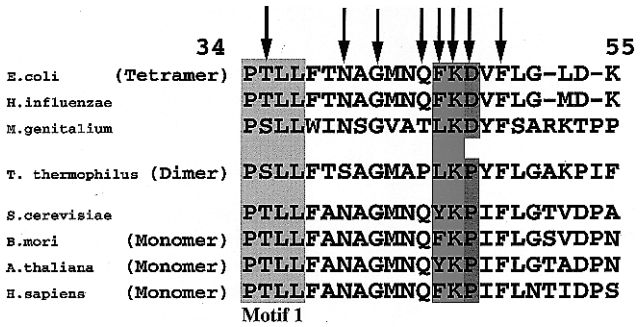

FIGURE 2: Alignment of AlaRS sequences in the motif 1 region. The residue numbering corresponds to the *E. coli* sequence. The motif 1 consensus sequence is shadowed in gray. Arrows indicate positions investigated in this work. Positions 46 and 47 are shadowed in red, and position 48 is shadowed in gray and blue to highlight the distribution of side chain types at this position. The quaternary structure (monomer, homodimer, or honotetramer) is indicated for those cases where it is known.

sequence, the segment identified coincides with the L1 loop (Figure 1) and extends approximately 20 residues from P34. In the light of the three-dimensional structures of other class II enzymes, eight residues were selected as of potential functional importance (Figure 2). These included T35, N40, G42, Q45, F46, K47, D48, and F50 in the *E. coli* enzyme. In the region analyzed for our experiments, only one position (D48 in *E. coli*) showed a consistent difference in amino acid side chain between enzymes of known quaternary structure. This position is occupied by an Asp in *E. coli* AlaRS (a tetramer) or Asn in other highly related bacterial sequences and by Pro in the rest of prokaryotes, archaea, and eukaryotes including human and yeast AlaRS (monomers).

We also analyzed the C-terminal sequences of alanyl-tRNA synthetases, because this part of the sequence harbors an element for oligomerization between Gly699 and Glu808 of the *E. coli* protein (*12*). We noted that this region of the sequence of the *E. coli* protein was strongly predicted to form a coiled-coil structure (Figure 3). In contrast, this coiled-coil tendency mostly disappears in the eukaryote enzymes. Thus, the absence of this predicted coiled-coil structure correlates with the known monomeric quaternary structure of the *Bombyx mori* and human enzymes. It is thus plausible that the oligomerization of AlaRS depends on contacts in the C-terminal region, where the C-terminal oligomerization element is located, and in the N-terminal region, where motif 1 is found.

Initial Mutagenesis Experiments. The tRNA synthetases typically carry out aminoacylation in a two-step reaction:

$$E + AA + ATP \rightarrow E(AA\text{-}AMP) + PP_i \quad (1)$$

$$E(AA\text{-}AMP) + tRNA \rightarrow AA\text{-}tRNA + AMP + E \quad (2)$$

In the first reaction (adenylate synthesis), the amino acid AA is activated to an enzyme-bound aminoacyl adenylate (AA-AMP), while in the second reaction the aminoacyl moiety is transferred to the 3′ end of the tRNA to give AA-tRNA (*6, 50*). We used assays of adenylate synthesis (eq 1) and of aminoacylation of microhelices (eqs 1 and 2) to test the functional effects of our mutations. We used microhelix substrates (Figure 4a) because of the known role of motif 1 in acceptor−helix interactions (*20*). Thus, the assays were designed to be sensitive to perturbations of these interactions.

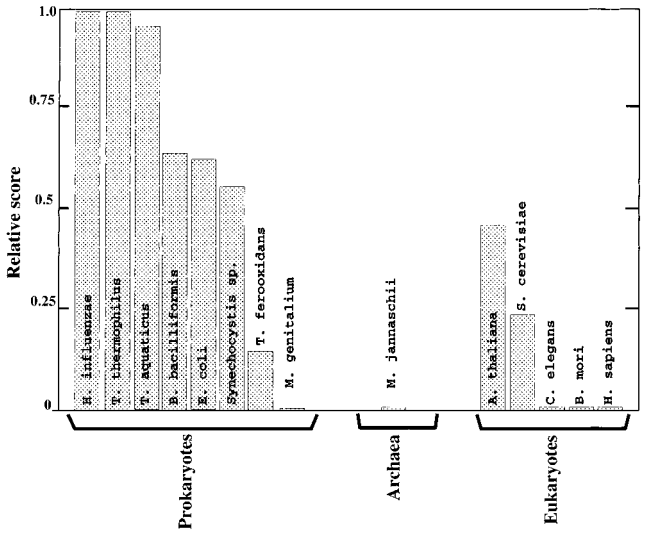

FIGURE 3: Coiled-coil predictions for the C-terminal regions of AlaRS sequences. The scores were obtained by multiplying the probability value for a coiled coil in the region [obtained using the program COILS (*34*)] by the total number of residues predicted to form a coiled coil and were normalized to the highest value (*H. influenzae*, 95% probability of coiled coil in a region of 47 amino acids).

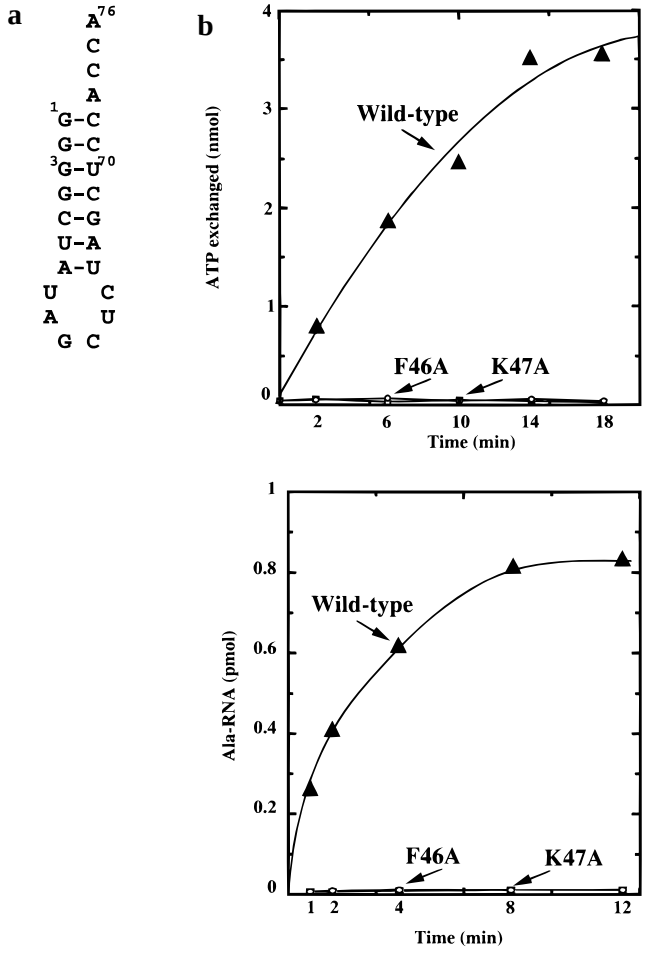

FIGURE 4: (a) Sequence and schematic representation of the microhelix^{Ala} used in this work. (b) Amino acid activation and microhelix charging activities for wild-type, F46A, and K47A AlaRS. Both mutant enzymes show over 100-fold decreases in both activities with respect to wild-type AlaRS.

Table 1: Results of Alanine Scanning Mutagenesis of the Motif 1 Region of *E. coli* AlaRS

enzyme	complementation of *alaS* null strain[a]	expression in null strain[b]	RNA microhelix charging activity[c]	adenylate synthesis activity[c]
wt	+	+	100	100
T35A	+	+	nd	nd
F38A	+	+	nd	nd
N40A	+	+	nd	nd
G42A	+	+	nd	nd
M43A	+	+	nd	nd
Q45A	+	+	nd	nd
F46A	−	+	<1	<1
K47A	−	+	<1	<1
D48P	+	+	70	80
F50A	+	+	nd	nd

[a] Complementation assay was as described in Materials and Methods. [b] Expression of mutant enzymes was checked by Western blot analysis of crude extracts with polyclonal anti-AlaRS antibodies (*42*). [c] Kinetic values determined with purified enzymes at pH 7.5, 25 °C, are relative to values for the wild-type enzyme.

The purpose of our initial experiments was to use mutagenesis to explore the idea that the predicted motif 1 region of AlaRS was similar to that of other class II tRNA synthetases. In particular, motif 1 in the dimeric class II enzymes contributes to both the active site and the dimerization interface. We therefore imagined that substitutions in the predicted motif 1 would affect catalytic activity and, additionally or alternatively, the oligomerization phenotype. Eight residues were selected for mutagenesis on the basis of their vicinity to the consensus sequence PTLL (*1, 29*) and their high level of conservation among sequences of alanyl-tRNA synthetases (Table 1, Figure 2). Their relative positions with respect to the consensus motif 1 sequence was considered because, in all three-dimensional structures of class II tRNA synthetases, functional contacts at the site for adenylate synthesis are provided by a large β-hairpin that follows the consensus sequence of the motif. Thus, assuming that the region following motif 1 of alanyl-tRNA synthetase has an equivalent role, mutations in its sequence should, by analogy, affect the adenylate synthesis activity of the enzyme.

Mutants were constructed as described in Materials and Methods, and their capacity to support growth of an *alaS* null strain was tested. This strain has a deletion of the gene *alaS* from the chromosome, and cells are maintained by an *alaS*-encoding plasmid that has a temperature-sensitive replicon. Because the plasmid-borne replicon is defective at 42 °C, no growth occurs unless a second plasmid encoding an active alanyl-tRNA synthetase is introduced. Thus, this system provides a way to test whether mutant enzymes have an activity that is sufficient to sustain cell growth and rescue the temperature-sensitive phenotype.

Mutations at two positions—F46A and K47A—resulted in the noncomplementation phenotype (Table 1). We verified that these mutant proteins were stable and accumulated in the cell, so that the reason for their noncomplementation phenotype was not the trivial one of being unstable and subject to degradation within the cell. The two mutants were then purified by Ni−NTA affinity chromatography, and their amino acid activation and aminoacylation activities were analyzed as described in Materials and Methods. As shown in Figure 4b, each of the two mutant enzymes is dramatically impaired for each of these catalytic activities. This result is consistent with the predicted motif 1 region playing an important role in catalysis, by analogy to its role in other class II enzymes.

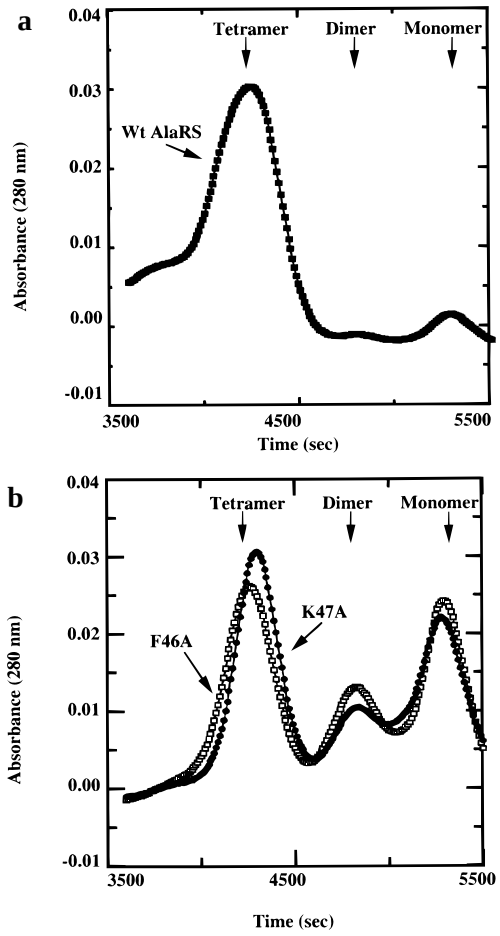

FIGURE 5: (a) Elution profile of wild-type AlaRS and molecular weight markers from the gel filtration chromatographic Superose 6 column. (b) Elution profile of F46A and K47A AlaRS under the same conditions as in (a).

Analysis of Quaternary Structures. The quaternary structures of the two mutant enzymes were analyzed by gel filtration chromatography on a Superose 6 column. Each of the mutant enzymes had a different size-exclusion chromatogram than that of the wild-type enzyme (Figure 5). The elution profile of wild-type alanyl-tRNA synthetase (AlaRS-6H) had a single main peak with an elution time that corresponded to an approximate molecular weight of 400 000. This molecular weight is close to the expected molecular weight of an AlaRS tetramer (~395 000). A small, but reproducible, peak was also observed with a calculated molecular weight of 95 000. This peak had protein with an aminoacylation activity that was proportionally equivalent to that of the main elution peak (data not shown). We concluded that this second peak contained monomeric AlaRS.

The mutant enzymes, on the other hand, produced a dramatically different elution profile. Here, three clear elution peaks of comparable heights can be observed. Polypeptide of the same molecular weight as that of the subunit of AlaRS was determined in all peaks by SDS−PAGE. The molecular weights estimated for the protein species of each elution peak closely approximate the values expected for monomeric, dimeric, and tetrameric AlaRS. Thus, point mutations in the predicted motif 1 of alanyl-

tRNA synthetase yield three of the quaternary structures observed for class II enzymes.

These results thus show that the predicted motif 1 plays a role in subunit interactions and catalysis, as is the case for other class II enzymes.

Analysis of D48P Alanyl-tRNA Synthetase. The cross-subunit interactions, at the motif 1 region of the active sites of the homodimeric class II enzymes of known structures, suggest a critical functional role for the subunit interfaces in these enzymes. The reduced activities and altered quaternary structures of both the F46A and K47A mutant alanine enzymes suggest that activity and quaternary structure are linked together, at least with respect to mutations in motif 1. In contrast, eukaryote alanyl-tRNA synthetases are monomeric proteins. These observations suggest that it should be possible to obtain a mutation in motif 1 of the *E. coli* enzyme which would alter the quaternary structure but not affect the active site architecture or the recognition of the tRNA acceptor stem. This kind of mutant would thus establish the principle that different quaternary structures of naturally occurring class II enzymes could be generated by sequence variations in motif 1, together with the presence or absence of an additional domain that built other subunit interactions such as a coiled-coil structure.

Encouraged by our ability to disrupt the formation of tetramers through mutations at the motif 1 region, we envisaged the possibility of obtaining the same disruption without significantly affecting the activity of the protein with microhelix substrates. Such a mutant would give the necessary evolutionary connection of a functional monomeric alanyl-tRNA synthetase (eukaryote-like) active site with its dimeric or tetrameric relatives in class II. For this purpose, we searched for sequence differences between motif 1 of eukaryote and prokaryote AlaRS. Only one position around the motif 1 sequences of AlaRS enzymes clearly showed a species-specific conservation pattern. Position 48 is either an aspartate or an asparagine in tetrameric prokaryotes and is a proline in eukaryotes (Figure 2). The nature of the side chain separates the two known quaternary forms of the enzyme, indicating potential structural differences in motif 1. Also, D48 is adjacent to K47, which is strictly conserved across all species and is critical for enzyme activity.

A D48P substitution was created in AlaRS-6H. Complementation assays, as well as *in vitro* kinetic analysis, showed this enzyme to be almost indistinguishable from the wild-type protein (data not shown).

In contrast, gel filtration revealed an elution pattern similar to that found for mutants F46A and K47A (Figure 6). Thus, the D48P substitution does not significantly affect the catalytic properties of *E. coli* AlaRS with microhelices but selectively disrupts the tetramer−dimer−monomer distribution. (This distribution was concentration independent over the range of 2 μM to 20 nM, suggesting that it is generated *in vivo* during folding.) This mutant bacterial protein demonstrates the possibility of altering motif-1-dependent subunit interactions without disrupting the structure of the active site. Thus, the phenotype of this mutant protein is like that of the eukaryotic alanine enzymes in that full-length monomers are generated which maintain full acceptor stem recognition even with disruption of motif 1-interactions.

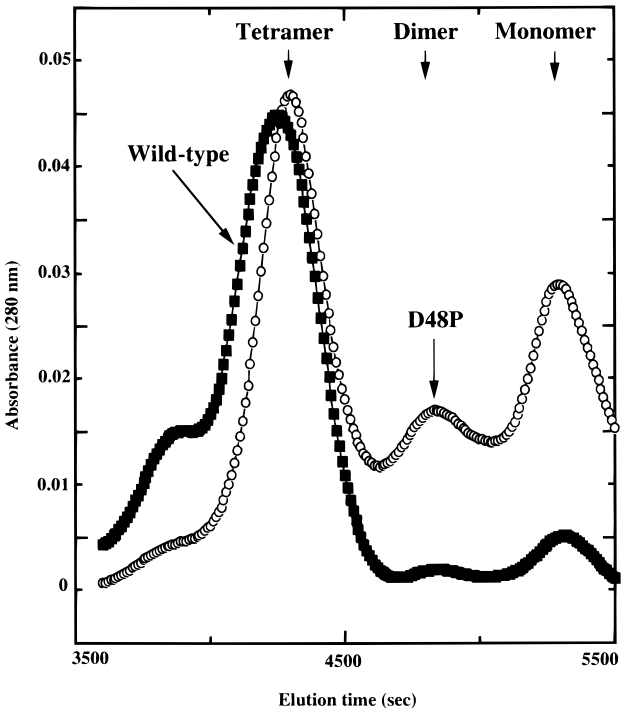

FIGURE 6: Elution profile of wild-type and D48P AlaRS under the same conditions as in Figure 5.

DISCUSSION

The gel filtration result with the D48P mutant (Figure 6), showing that monomers, dimers, and tetramers of *E. coli* alanyl-tRNA synthetase are present, is similar to that seen when dimethyl suberimidate cross-linking of alanyl-tRNA synthetase was carried out, and the cross-linked species were resolved by SDS−PAGE (*27*). In those cross-linking studies, monomers, dimers, and tetramers were also seen. Thus, the native α_4 enzyme can be viewed as a dimer of dimers. Earlier work identified a C-terminal region as important for subunit oligomerization (*12, 41*). The observations reported here identify a second region—motif 1—as also essential for the native tetrameric quaternary structure.

Previous mutation analysis by Eriani et al. (*51*) of motif 1 in a class II enzyme investigated the role of the conserved proline [P35 in the *E. coli* enzyme (Figure 1)] in yeast aspartyl-tRNA synthetase. In this system, direct enzyme-substrate contacts, such as those with ATP or $tRNA^{Asp}$, are largely confined to residues within a single subunit. [An exception is the interaction of the backbone phosphate of the first nucleotide of $tRNA^{Asp}$ bound to one subunit with a lysine amino group from the opposite subunit (*20*).] While an aspartyl-tRNA synthetase monomer was not obtained in the study of Eriani et al. (*51*), the analysis nonetheless gave support to the idea that catalytic activity was sensitive to the integrity of the dimer interface. For example, the Pro → Gly change in motif 1 (P273G in the aspartyl-tRNA synthetase sequence) reduced both the adenylate synthesis and the charging activity. This observation is consistent with the ATP bound at one subunit being indirectly linked to the conserved proline of the opposite subunit, by cross-subunit interactions of residues that are adjacent to this proline (*20*). All the available three-dimensional structures for class II synthetases suggest the existence of similar cross-subunit networks linking the structure of each motif 1 to the active site of the adjacent subunit (*3, 5, 14, 17−19*).

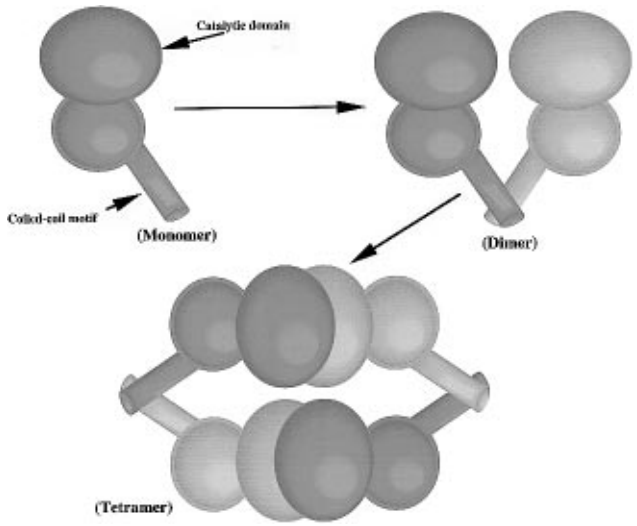

FIGURE 7: Tetramerization model for AlaRS. We propose a stepwise mechanism that requires initially the formation of a coiled-coil interaction through the C-terminal sequences to form an initial dimer. Interactions at the motif 1 region of the active site would be involved in the assembly of two dimers to form the final tetramer.

Unlike aspartyl-tRNA synthetase and other class II enzymes, *E. coli* alanyl-tRNA synthetase is a tetramer with an idiosyncratic C-terminal oligomerization domain that is essential for formation of the tetrameric structure. Moreover, monomers rather than dimers form when this domain is deleted (*12*). The results reported here show that, while motif 1 of alanyl-tRNA synthetase is not sufficient by itself to generate dimers or higher order oligomers, it has a key role in quaternary structure assembly when combined with the C-terminal domain. Possibly the coiled-coil interactions of the C-terminal domains bring together the subunits and increase the local concentration of motif 1 units. This enhancement of the local concentration could compensate for intersubunit interactions of motif 1 that are weaker than those seen with other class I enzymes (Figure 7).

The tetrameric and monomeric alanine enzymes are distinguished by the D *versus* P difference at position 48 of motif 1 (using the numbering system for the *E. coli* protein). Our experiments suggest that this difference and the missing C-terminal oligomerization domain are much of the explanation for the monomeric nature of eukaryotic alanyl-tRNA synthetases. Given the seeming importance of the motif-1-dependent dimer interface for the function of class II enzymes, the monomeric state of eukaryote alanyl-tRNA synthetases and of yeast mitochondrial phenylalanyl tRNA synthetase was striking.

A recent report on the characterization of *T. thermophilus* AlaRS (*26*) is consistent with this analysis. The sequence of this enzyme contains a proline residue at position 48 (*26*), a divergence from its bacterial counterparts. However, the C-terminal region of its sequence is predicted to contain a coiled coil. [The prediction is as strong as that for *Haemophilus influenzae* AlaRS (Figure 3).] The quaternary structure of *T. thermophilus* AlaRS is dimeric (*26*), as predicted from our analysis because the enzyme lacks the second oligomerization element that is located in motif 1.

The motif 1 dimer interface may normally be designed to have an indirect but functionally important conformational effect on the active site, as suggested by the results on yeast aspartyl-tRNA synthetase (*51*). Thus, the monomeric class II species must have a sequence design that generates the active conformation of the catalytic site independent of the motif 1 subunit interactions. If so, then our results on the *E. coli* alanine enzyme suggest that the P48 substitution in motif 1 of eukaryote alanyl-tRNA synthetases is part of that design.

The F46A and K47A substitutions in AlaRS resulted in sharply diminished catalytic activity and a perturbation of the quaternary structure (Figures 4 and 5). While the correlation between diminished activity and a perturbed quaternary structure suggested that the two functions were linked, the high activity of the D48P mutant protein implies that this correlation was coincidental and not causal. However, the sharply diminished adenylate synthesis activity of the F46A and K47A mutant enzymes is consistent with this region forming part of the ATP binding site, as seen in the structures of other class II enzymes (*3*, *5*, *17*).

Our experiments suggest that the diverse quaternary structures seen with class II enzymes are determined in large part by variations in the sequence of motif 1. Although a second domain, such as the C-terminal oligomerization domain of bacterial alanyl-tRNA synthetase, may be required to generate tetramers, the behavior of the F46A, K47A, and D48P mutant proteins showed that stable tetrameric structures require a specific motif 1 sequence (Figures 4−6). Thus, rather than being a structural element for building dimeric class II enzymes, this work establishes that motif 1 is part of the general scheme for building monomeric, dimeric, and tetrameric class II synthetases.

ACKNOWLEDGMENT

We are grateful to Dr. Douglas Buechter and Dr. Niranjan Sardesai for helpful discussions. We thank an anonymous reviewer for calling our attention to the analysis of the *T. thermophilus* sequence by Lechler et al. (1997), which strengthened the conclusions presented here.

REFERENCES

1. Eriani, G., Delarue, M., Poch, O., Gangloff, J., and Moras, D. (1990) *Nature 347*, 203−206.
2. Rould, M. A., Perona, J. J., Söll, D., and Steitz, T. A. (1989) *Science 246*, 1135−1142.
3. Cusack, S., Berthet-Colominas, C., Härtlein, M., Nassar, N., and Leberman, R. (1990) *Nature 347*, 249−255.
4. Cusack, S., Härtlein, M., and Leberman, R. (1991) *Nucleic Acids Res. 19*, 3489−3498.
5. Ruff, M., Krishnaswamy, S., Boeglin, M., Poterszman, A., Mitschler, A., Podjarny, A., Rees, B., Thierry, J. C., and Moras, D. (1991) *Science 252*, 1682−1689.
6. Moras, D. (1992) *Trends Biochem. Sci. 17*, 159−164.
7. Schimmel, P., and Ribas de Pouplana, L. (1995) *Cell 81*, 983−986.
8. Brick, P., Bhat, T. N., and Blow, D. M. (1988) *J. Mol. Biol. 208*, 83−98.
9. Doublié, S., Bricogne, G., Gilmore, C., and Carter, C. W., Jr. (1995) *Structure 3*, 17−31.
10. Nureki, O., Vassylyev, D. G., Katayanagi, K., Shimizu, T., Sekine, S.-i., Kigawa, T., Miyazawa, T., Yokoyama, S., and Morikawa, K. (1995) *Science 267*, 1958−1965.
11. Ostrem, D. L., and Berg, P. (1970)) *Proc. Natl. Acad. Sci. U.S.A. 67*, 1967−1974.
12. Jasin, M., Regan, L., and Schimmel, P. (1983) *Nature 306*, 441−447.
13. Schimmel, P. (1990) *Adv. Enzymol. Relat. Areas Mol. Biol. 63*, 233−270.
14. Mosyak, L., Reshetnikova, L., Goldgur, Y., and Delarue, M. (1995) *Nat. Struct. Biol. 2*, 537−547.

15. Cusack, S. (1995) in *Tracing biological evolution in protein and gene structures* (Go, M., and Schimmel, P., Eds.) pp 37−52, Elsevier Science, London.
16. Shiba, K., Ripmaster, T., Suzuki, N., Nichols, R., Plotz, P., Noda, T., and Schimmel, P. (1995) *Biochemistry 34*, 10340−10349.
17. Arnez, J. G., Harris, D. C., Mitschler, A., Rees, B., Francklyn, C. S., and Moras, D. (1995) *EMBO J. 14*, 4143−4155.
18. Logan, D. T., Mazauric, M. H., Kern, D., and Moras, D. (1995) *EMBO J. 14*, 4156−4167.
19. Onesti, S., Miller, A. D., and Brick, P. (1995) *Structure 3*, 163−176.
20. Cavarelli, J., Rees, B., Ruff, M., Thierry, J. C., and Moras, D. (1993) *Nature 362*, 181−184.
21. Toth, M. J., and Schimmel, P. (1986) *J. Biol. Chem. 261*, 6643−6646.
22. Toth, M. J., and Schimmel, P. (1990) *J. Biol. Chem. 265*, 1005−1009.
23. Ge, Q., Trieu, E. P., and Targoff, I. N. (1994) *J. Biol. Chem. 269*, 28790−28797.
24. Wagar, E. A., Giese, M. J., Yasin, B., and Pang, M. (1995) *J. Bacteriol. 177*, 5179−5185.
25. Shiba, K., Schimmel, P., Motegi, H., and Noda, T. (1994) *J. Biol. Chem. 269*, 30049−30055.
26. Lechler, A., Martin, A., Zuleeg, T., Limmer, S., and Kreutzer, R. (1997) *Nucleic Acids Res. 25*, 2737−2744.
27. Putney, S. D., Sauer, R. T., and Schimmel, P. (1981) *J. Biol. Chem. 256*, 198−204.
28. Chang, P. K., and Dignam, J. D. (1990) *J. Biol. Chem. 265*, 20898−20906.
29. Ribas de Pouplana, L., Buechter, D. D., Davis, M. W., and Schimmel, P. (1993) *Protein Sci. 2*, 2259−2262.
30. Davis, M. W., Buechter, D. D., and Schimmel, P. (1994) *Biochemistry 33*, 9904−9910.
31. Shi, J. P., Musier-Forsyth, K., and Schimmel, P. (1994) *Biochemistry 33*, 5312−5318.
32. Buechter, D. D., and Schimmel, P. (1993) *Biochemistry 32*, 5267−5272.
33. Benson, D., Boguski, M., Lipman, D. J., and Ostell, J. (1994) *Nucleic Acids Res. 22*, 3441−3444.
34. Berger, B., Wilson, D. B., Wolf, E., Tonchev, T., Milla, M., and Kim, P. S. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 8259−8263.
35. Higgins, D. G., and Sharp, P. M. (1988) *Gene 73*, 237−244.
36. Bernstein, F. C., Koetzle, T. G., Williams, G. J. B., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., and Tasumi, T. (1977) *J. Mol. Biol. 112*, 535−542.
37. Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 488−492.
38. Barrell, B. G., Anderson, S., Bankier, A. T., de Bruijn, M. H. L., Chen, E., Coulson, A., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R., and Young, I. G. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 3164−3166.
39. Jasin, M., and Schimmel, P. (1984) *J. Bacteriol. 159*, 783−786.
40. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
41. Jasin, M., Regan, L., and Schimmel, P. (1984) *Cell 36*, 1089−1095.
42. Regan, L. (1986) Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA.
43. Calendar, R., and Berg, P. (1966) *Biochemistry 5*, 1690−1695.
44. Schreier, A. A., and Schimmel, P. R. (1972) *Biochemistry 11*, 1582−1589.
45. Hill, K., and Schimmel, P. (1989) *Biochemistry 28*, 2577−2586.
46. Francklyn, C., Shi, J. P., and Schimmel, P. (1992) *Science 255*, 1121−1125.
47. Webster, T. A., Gibson, B. W., Keng, T., Biemann, K., and Schimmel, P. (1983) *J. Biol. Chem. 258*, 10637−10641.
48. Sanni, A. W. P., Boulanger, Y., Ebel, J. P., and Fasiolo, F. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 8387−8391.
49. Jasin, M., Regan, L., and Schimmel, P. (1985) *J. Biol. Chem. 260*, 2226−2230.
50. Schimmel, P. R., and Söll, D. (1979) *Annu. Rev. Biochem. 48*, 601−648.
51. Eriani, G., Cavarelli, J., Martin, F., Dirheimer, G., Moras, D., and Gangloff, J. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 10816−10820.

BI971788+