An Isolated Class II Aminoacyl-tRNA Synthetase Insertion Domain Is Functional in Amino Acid Editing*

Received for publication, August 29, 2003, and in revised form, September 27, 2003
Published, JBC Papers in Press, October 6, 2003, DOI 10.1074/jbc.M309627200

Fai-Chu Wong, Penny J. Beuning‡, Carmen Silvers, and Karin Musier-Forsyth§

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

Aminoacyl-tRNA synthetases are responsible for activating specific amino acids and transferring them onto cognate tRNA molecules. Due to the similarity in many amino acid side chains, certain synthetases misactivate non-cognate amino acids to an extent that would be detrimental to protein synthesis if left uncorrected. To ensure accurate translation of the genetic code, some synthetases therefore utilize editing mechanisms to hydrolyze non-cognate products. Previously class II Escherichia coli proline-tRNA synthetase (ProRS) was shown to exhibit pre- and post-transfer editing activity, hydrolyzing a misactivated alanine-adenylate (AlaAMP) and a mischarged Ala-tRNAPro variant, respectively. Residues critical for the editing activity (Asp-350 and Lys-279) are found in a novel insertion domain (INS) positioned between motifs 2 and 3 of the class defining microhelix-Pro variants but not cognate Pro-tRNAPro. Thus, we demonstrate for the first time that an independently folded class II synthetase editing domain and a previously identified homolog can catalyze a hydrolytic editing reaction.

Aminoacyl-tRNA synthetases are the family of enzymes responsible for activating cognate amino acids with ATP to form aminoacyl-adenylates and subsequently transferring the activated amino acids onto their corresponding tRNAs (1). The aminoacylated or charged tRNA is then delivered to the ribosome for use in protein synthesis. Each aminoacyl-tRNA synthetase must select and activate its cognate amino acid from the cellular pool of 20 different proteinaceous amino acids. High fidelity in this selection process is necessary to ensure faithful protein translation as an accumulation of mistakes in the aminoacylation process will eventually lead to cell death (2–4).

Some amino acids have chemical structures that closely resemble each other, making accurate discrimination by synthetases difficult. In particular, smaller non-cognate amino acids may enter into the active sites of synthetases and be misactivated and subsequently transferred onto the wrong tRNA. Some synthetases are capable of correcting mistakes in amino acid selection through hydrolysis of the misactivated aminoacyl-adenylate prior to amino acid transfer to the tRNA (pretransfer editing) and hydrolysis of the ester linkage of the misacylated tRNA (post-transfer editing) (5).

Editing has been well documented in the case of class I synthetases. For example, class I isoleucine-tRNA synthetase (IleRS)¹ and valine-tRNA synthetase (ValRS) activate and edit the non-cognate amino acids leucine and threonine, respectively (6, 7). More recently, class I leucine-tRNA synthetase (LeuRS) has also been shown to misactivate and edit isoleucine as well as a series of non-standard amino acids (8–11). In all of these class I enzymes, a structural domain distinct from the aminoacylation active site known as connective polypeptide 1 (CP1) (12) is responsible for the editing activity (9, 11, 13, 14). Biochemical and structural studies have also identified editing activities in class II synthetases, including editing of serine by threonine-tRNA synthetase (ThrRS) (15, 16), editing of alanine by proline-tRNA synthetase (ProRS) (17, 18), and editing of glycine and serine by alanine-tRNA synthetase (AlaRS) (19, 20).

Synthetases are modular proteins composed of domains that have distinct functional roles (21, 22). The core catalytic domain is responsible for amino acid activation and tRNA acceptor stem docking. It has been proposed that ancestral aminoacyl-tRNA synthetases consisted only of the core catalytic domain, while the anticodon-binding domain was recruited later to improve the binding and discrimination of tRNA substrates (21–23). The catalytic domain alone of some synthetases is capable of aminoacylating tRNA or microhelix substrates that mimic the acceptor stem of full-length tRNA (24, 25), supporting such an evolutionary scenario. In some cases, an editing domain may have been recruited to improve amino acid specificity and eliminate non-cognate product formation (22, 23). “Extra” domains such as the editing domains of certain class I synthetases have been cloned and expressed as independent functional domains. In particular, the independently

¹ The abbreviations used are: IleRS, isoleucine-tRNA synthetase; ValRS, valine-tRNA synthetase; LeuRS, leucine-tRNA synthetase; CP1, connective polypeptide 1; ThrRS, threonine-tRNA synthetase; ProRS, proline-tRNA synthetase; AlaRS, alanine-tRNA synthetase; INS, insertion domain; MBP, maltose-binding protein; HI1434, H. influenzae YbaK protein.
expressed CP1 domains from IleRS and ValRS were shown to deacylate Val-tRNAVal and Thr-tRNAThr, respectively (14). Interestingly the CP1 domain of LeuRS has been shown to play a critical role in group I intron splicing (26).

Editing domains found in class II aminoacyl-tRNA synthetases are significantly different from the CP1 domain responsible for editing in class I synthetases. Moreover, in the case of class II enzymes, editing domains identified to date are believed to be distinct from each other, although there is weak homology between the editing domains of ThrRS and AlaRS (15, 20). Based on sequence alignments, class II ProRSs can be divided into two distinct groups (27–29). The “prokaryotic-like” group contains synthetases from bacteria and eukaryotic mitochondrial enzymes, whereas the “eukaryotic-like” group contains synthetases from Eukaryotes, Archaea, and Bacteria. In the case of Escherichia coli ProRS, a representative member of the prokaryotic-like group, the editing active site is a novel insertion domain (INS) located between motifs 2 and 3, which together constitute the aminocylating active site (18). Surprisingly this ~200-amino acid INS is absent from eukaryotic-like ProRS. This observation suggests that either the INS domain was recruited late in evolution to enhance the performance of the prokaryotic-like enzymes (23) or that it was present early in both groupings of ProRS and then lost from the eukaryotic-like group, which generally exhibits higher amino acid specificity (30). Whereas the three-dimensional structures of two members of the eukaryotic-like ProRS group have been solved by x-ray crystallography (27, 31, 32), the structure of a bacterial ProRS containing the INS domain is not yet known.

Previously we have shown that E. coli ProRS is capable of hydrolyzing non-cognate Ala-AMP in a tRNA-independent pre-transfer editing reaction and of deacylating Ala-tRNAPro via post-transfer editing (17). Using alanine-scanning mutagenesis, we have also identified residues within the INS domain that are critical for both pre- and post-transfer editing (18). More recently, ProRS from the bacterium Aquifex aeolicus was shown to possess post-transfer editing activity, whereas a mutant lacking 117 residues of the insertion domain also lacked detectable deacylation activity (33). Taken together with the mutagenesis data obtained with the E. coli enzyme, these data provide strong support for the role of the prokaryotic INS domain in amino acid editing by ProRS.

By analogy to the class I CP1 domain, we hypothesized that the INS domain alone might be capable of hydrolytic editing activity. To test this hypothesis, we cloned and expressed the E. coli INS domain and tested its post-transfer editing capability. In addition, we deleted the INS from full-length E. coli ProRS to yield ΔINS-ProRS. We also tested the Haemophilus influenzae YbaK protein, previously reported to have high homology to the bacterial ProRS INS domain (34, 35), for editing activity. The data presented here provide further support for the functional role of the INS domain in translational editing and demonstrate for the first time that an independently folded domain derived from a class II synthetase can catalyze hydrolytic editing.

EXPERIMENTAL PROCEDURES

Plasmid Construction—A plasmid encoding ΔINS-ProRS was constructed from plasmid pCS-M1S, which encodes wild-type E. coli ProRS with an N-terminal histidine tag (28). This plasmid contains a single KpnI restriction site at the N terminus of the INS domain. A second KpnI restriction site was introduced at the C-terminal end of the INS domain using the QuickChange site-directed mutagenesis kit (Stratagene) and the following primers: 5’-GGCCGGTCTGCGTACGGA-3’ and 5’-CATTTTCCAGTGCAGGGTCCGCTGC-3’. After digesting with KpnI to excise the insertion domain, the digested plasmid was then ligated to yield the ΔINS-ProRS with amino acid residues 249–418 of wild-type E. coli INS domain deleted.

The plasmid containing the E. coli INS domain was prepared by PCR amplifying the INS fragment from full-length plasmid pCS-M1S using 5’-CCGCGGTGTCCAGGGTCCG-3’ and 5’-CCCGAGTTTGTTACGGCCATCTTTACCCC-3’. After digesting with BamHI and HindIII, the PCR fragments were ligated into the corresponding sites of the pMAL-c2E vector (New England Biolabs) immediately following the maltose-binding protein (MBP) coding region. Following ligation, plasmids were transformed into XL-1 Blue supercompetent cells (Stratagene). Results of all cloning and definition steps were confirmed by automated DNA sequencing (Microchemical Facility, University of Minnesota).

RNA Preparation—Wild-type E. coli tRNAPro and a G1:C72/U70-tRNAPro triple mutant were prepared by in vitro transcription from BstNI-linearized plasmids using T7 RNA polymerase as described previously (28, 36). The G1:C72/U70-microhelixPro variant was synthesized using automated chemical RNA synthesis as described previously (37, 38). Mischarged G1:C72/U70-tRNAPro and G1:C72/U70-microhelixPro were used for in vitro deacylation assays and were prepared as described previously (17, 18). All reactions contained 2 units/ml inorganic pyrophosphatase, and reactions were quenched by the addition of acetic acid to 1% final concentration. Purified charged tRNA and microhelix were quantified by scintillation counting.

Protein Preparation—Expression of wild-type and truncated ΔINS-ProRS was induced with 1 mM isopropl-β-D-thiogalactopyranoside in SG13009(pREP4) or BL21(DE3) plysE competent cells. His-tagged proteins were then purified and prepared using a Talon colo affinity resin (Takelab) as described previously (39). The MBP-INS fusion protein was purified using amylose affinity chromatography according to the manufacturer’s protocol (New England Biolabs). In brief, sonicated protein solution was loaded onto the amylose affinity column followed by extensive washing with 15 column volumes of buffer (20 mM Tris-HCl, pH 7.4, 200 mM NaCl, 1 mM EDTA). The MBP-INS fusion protein was then eluted with column buffer containing 10 mM maltose. Fractions containing the fusion protein were identified via SDS-PAGE, pooled, and subjected to cleavage by 0.1% enterokinase (New England Biolabs). The latter was performed at 37 °C for 1 h followed by incubation at 25 °C overnight. Following the removal of maltose from the cleaved fusion protein by hydroxyapatite chromatography (Bio-Rad), the solution was loaded onto a second amylose column. The cleaved INS domain was then eluted using column buffer, and fractions containing the purified protein were pooled and concentrated using Centricron 10 microconcentrators (Amicon). Purified enzyme concentrations were determined either by active site titration (full-length and ΔINS-ProRS) or by Bradford protein assay (INS) (Bio-Rad). Concentrated protein was stored in column buffer with the addition of an equal volume of 80% glycerol.

Plasmid pCYB2_H1143 encoding H. influenzae YbaK protein (H11434) was a gift from Prof. Osnat Herzberg (University of Maryland). The H11434 is expressed as a fusion protein consisting of an N-terminal H11434 polypeptide followed by a self-splitting intein and a chitin binding domain. Protein purification was performed essentially as described previously (35) using the IMPACT^™ system (New England Biolabs). In brief, protein expression was induced in the presence of 1 mM isopropl-β-D-thiogalactopyranoside, and cell pellets were harvested after 4 h of growth at 37 °C. Cell pellets were then lysed by sonication and clarified by centrifugation. Cell-free lysate was then passed through a chitin column. After washing with column buffer (20 mM HEPES, pH 8.0, 500 mM NaCl, 0.1 mM EDTA, and 0.1% Triton), inter- mediated cleavage was initiated by the addition of column buffer containing 50 mM dithiothreitol. Protein was then eluted with column buffer, and fractions containing the YbaK protein were pooled and concentrated using Centriprep 10 and Centricon 10 concentrators (Amicon). Concentrated protein was stored in column buffer containing an equal volume of 80% glycerol.

Circular Dichroism (CD) Spectroscopy—CD spectra of proteins were obtained at room temperature using a J-710 spectropolarimeter (Jasco). Protein samples at a concentration of 0.5 mg/ml in 50 mM HEPES, pH 7.0 were analyzed using a 0.1-cm-path length cuvette, and spectra were corrected over six scan cycles. The results of the UV-Vis absorption and cuvette were performed using the published conditions (43). The amino acid concentrations used were 0.05–2 mM (proline) and 25–500 mM (alanine), and the final...
enzyme concentrations were 1 mM (with proline) and 20 mM (with alanine). Kinetic parameters were determined from Lineweaver-Burk plots and represent the average of three determinations. Editing assays were performed using 2.5 nM wild-type ProRS and the cloned INS domain in comparison to full-length E. coli ProRS as indicated by a dotted line. AC indicates the anticodon-binding domain, and the numbers indicate the amino acid positions in ProRS. The arrow points to the cleavage site in the MBP-INS fusion construct.

**RESULTS**

Expression and Stability of Truncated ProRS Constructs—Fig. 1 shows a schematic illustration of the domain architecture of full-length E. coli ProRS (top). Two additional proteins derived from full-length E. coli ProRS were constructed for this work. ΔINS-ProRS, a variant with a deletion from residues 249–418 in full-length E. coli ProRS as indicated by a dotted line. AC indicates the anticodon-binding domain, and the numbers indicate the amino acid positions in ProRS. The arrow points to the cleavage site in the MBP-INS fusion construct.

**Expression and Stability of Truncated ProRS Constructs—**Fig. 1 shows a schematic illustration of the domain architecture of full-length E. coli ProRS (top). Two additional proteins derived from full-length E. coli ProRS were constructed for this work. ΔINS-ProRS, a variant with a deletion from residues 249–418 in full-length E. coli ProRS as indicated by a dotted line. AC indicates the anticodon-binding domain, and the numbers indicate the amino acid positions in ProRS. The arrow points to the cleavage site in the MBP-INS fusion construct.

**Expression and Stability of Truncated ProRS Constructs—**Fig. 1 shows a schematic illustration of the domain architecture of full-length E. coli ProRS (top). Two additional proteins derived from full-length E. coli ProRS were constructed for this work. ΔINS-ProRS, a variant with a deletion from residues 249–418 in full-length E. coli ProRS as indicated by a dotted line. AC indicates the anticodon-binding domain, and the numbers indicate the amino acid positions in ProRS. The arrow points to the cleavage site in the MBP-INS fusion construct.

**Expression and Stability of Truncated ProRS Constructs—**Fig. 1 shows a schematic illustration of the domain architecture of full-length E. coli ProRS (top). Two additional proteins derived from full-length E. coli ProRS were constructed for this work. ΔINS-ProRS, a variant with a deletion from residues 249–418 in full-length E. coli ProRS as indicated by a dotted line. AC indicates the anticodon-binding domain, and the numbers indicate the amino acid positions in ProRS. The arrow points to the cleavage site in the MBP-INS fusion construct.

**Expression and Stability of Truncated ProRS Constructs—**Fig. 1 shows a schematic illustration of the domain architecture of full-length E. coli ProRS (top). Two additional proteins derived from full-length E. coli ProRS were constructed for this work. ΔINS-ProRS, a variant with a deletion from residues 249–418 in full-length E. coli ProRS as indicated by a dotted line. AC indicates the anticodon-binding domain, and the numbers indicate the amino acid positions in ProRS. The arrow points to the cleavage site in the MBP-INS fusion construct.
**Fig. 2. CD spectra of wild-type *E. coli* ProRS and truncated constructs.** Protein samples were measured at a concentration of 0.5 mg/ml in 50 mM HEPES, pH 7.0 at room temperature using a 0.1-cm-path length cuvette. Spectra were accumulated over six scans. A shows full-length *E. coli* ProRS (solid curve) and ΔINS-ProRS (dashed curve). B shows the INS domain (solid curve) and K279A-INS (dashed curve). The inset to A shows a spectrum calculated by subtraction of the INS spectrum from that of the full-length protein (WT-INS). WT, wild type; deg, degrees.
rified *E. coli* AlaRS and 0.5 μM tRNA<sup>Ala</sup>.

**Activation of Cognate Proline, Cognate tRNA Aminoacylation, and Mischarging**—Under standard reaction conditions, no activation of proline to form the aminoacyl-adenylate is observed with ΔINS-ProRS. Therefore, although the ΔINS-ProRS protein appeared to be globally folded with secondary structural elements that resemble those of full-length ProRS (Fig. 2A), we speculate that deletion of the INS domain without insertion of a flexible linker sequence may cause misfolding of the adjacent (motifs 2 and 3) active site structure and thus prevent the cognate amino acid from being properly bound and activated. As expected, cognate proline activation was not detected with the independently cloned INS domain. Addition of the INS domain to the ΔINS-ProRS construct was also attempted, but no stimulation of proline activation was detected under any conditions used.

Both the INS domain and ΔINS-ProRS were tested for their ability to aminoacylate *E. coli* tRNA<sup>Pro</sup> with proline. No tRNA charging activity was detected with either protein. This is not surprising given the proline activation data described above. In addition, no non-cognate alanine mischarging activity was detected using as high as 2 μM ΔINS-ProRS and 10 μM tRNA<sup>Pro</sup>.

**Pretransfer Editing**—ATP hydrolysis in the presence of non-cognate amino acid is indicative of pretransfer editing. In the presence of 500 mM alanine, ATP hydrolysis catalyzed by wild-type ProRS was readily detected at room temperature (Fig. 3). Under the same conditions, no pretransfer editing activity was detected with ΔINS-ProRS (Fig. 3) despite the observation that the deletion construct was capable of misactivating non-cognate alanine. Addition of purified INS (up to 3 μM) did not stimulate editing by ΔINS-ProRS (Fig. 3).

**Post-transfer Editing by INS and ΔINS-ProRS**—Wild-type, ΔINS-ProRS, and the cloned INS were assayed for their ability to deacylate mischarged Ala-tRNA<sup>Pro</sup> or Ala-microhelix<sup>Pro</sup> variants (Fig. 4). For these experiments, a triple mutation was introduced into the acceptor stem of tRNA<sup>Pro</sup> and microhelix<sup>Pro</sup> to stimulate alanine charging by *E. coli* AlaRS as described previously (17). No deacylation of the mischarged G1:C72/U70-Ala-tRNA<sup>Pro</sup> was detected in the presence of up to 5 μM ΔINS-ProRS or INS. Under the same conditions, the mischarged Ala-tRNA<sup>Pro</sup> variant was readily deacylated by wild-type *E. coli* ProRS (data not shown) (17). We also tested the ProRS deletion constructs for their ability to deacylate the mischarged G1:C72/U70-Ala-microhelix<sup>Pro</sup>. With this substrate, the INS domain alone was found to catalyze deacylation at rates similar to wild-type *E. coli* ProRS (data not shown). No deacylation of Ala-microhelix<sup>Pro</sup> was detected in the presence of ΔINS-ProRS or INS. Under the same conditions, the mischarged Ala-tRNA<sup>Pro</sup> variant was readily deacylated by wild-type *E. coli* ProRS (data not shown) (17).

**Post-transfer Editing by *H. influenzae* YbaK Protein**—The prokaryotic ProRS INS domain shows no sequence homol-
As an independent protein. In addition, we cloned and expressed the 248-amino acid INS domain of ProRS is functional when separated from the rest of the enzyme, we cloned and expressed the 248-amino acid INS domain of ProRS. The stably folded 211-amino acid CP1Ile and 275-amino acid CP1 Val segments are capable of performing specific post-transfer editing of Val-tRNA^{Val} and Thr-tRNA^{Val}, respectively (14). In contrast, the 260-amino acid editing domain of E. coli LeuRS (CP1_{Leu}), which also appears to fold into a stable protein, was reported to lack editing function (9). Thus, in this case, it was proposed that the CP1 domain is active only in the context of the full-length LeuRS enzyme. The present work is the first study to examine the activity of an isolated editing domain from a class II synthetase.

To determine whether the editing domain of class II E. coli ProRS is functional when separated from the rest of the enzyme, we cloned and expressed the 248-amino acid INS domain as an independent protein. In addition, ΔINS-ProRS, a mutant construct with 86% of the INS domain deleted, was also examined for its aminoacylation and editing activities (Fig. 1). The ΔINS-ProRS deletion construct failed to activate cognate proline or non-cognate glycine (data not shown). The crystal structure of Thermus thermophilus ProRS complexed to the prolyl-adenylate intermediate suggests that residues that align with Glu-410, His-413, and Phe-415, which are missing in our ΔINS-ProRS construct, are critical for prolyl-adenylate formation (44). Since the deleted INS domain lies between motifs 2 and 3, which constitute the amino acid binding pocket, it is also possible that the large 169-amino acid deletion results in local misfolding of the active site. This structural change must be quite subtle since the active site is sufficiently formed to be able to misactivate alanine, and the CD spectrum shows the overall fold is maintained (Fig. 2A). Despite its inability to activate proline, the ΔINS-ProRS construct was able to misactivate alanine but was defective in pretransfer editing of Ala-AMP (Fig. 3). This observation is consistent with a previous study showing that residue Asp-350 in the INS domain is critical for pretransfer editing activity as D350A-ProRS is able to activate alanine but is completely defective in hydrolyzing Ala-AMP (18). The ΔINS deletion construct was not active in aminoacylation or in pre-transfer editing (Fig. 5) either alone or in the presence of the isolated INS domain. The lack of post-transfer editing is also consistent with a previous report showing that deletion of 65% of the bacterial A. aeolicus ProRS insertion sequence abolished post-transfer editing (33).

In contrast to the ΔINS-ProRS construct, the cloned INS domain did carry out post-transfer editing of Ala-microhelixPro (Fig. 5). The anticodon domain of tRNA^{Pro} is an important site of interaction with ProRS (27, 28, 42, 45), and the inability of the INS domain to deacetyl full-length mischarged tRNA may be due to the absence of binding determinants that facilitate correct positioning of the tRNA acceptor stem in the editing active site. To ensure that the observed microhelix deacetylation by the isolated INS domain was dependent upon the same active site residues as in the full-length protein (18), we showed that a K279A mutation abolished post-transfer editing (Fig. 5).

DISCUSSION

CP1 editing domains from two class I synthetases, E. coli IleRS and Bacillus stearothermophilus ValRS, have been cloned and expressed as isolated protein domains. The stably folded 211-amino acid CP1_Ile and 275-amino acid CP1_Val segments are capable of performing specific post-transfer editing of Val-tRNA^{Val} and Thr-tRNA^{Val}, respectively (14). In contrast, the 260-amino acid editing domain of E. coli LeuRS (CP1_{Leu}), which also appears to fold into a stable protein, was reported to lack editing function (9). Thus, in this case, it was proposed that the CP1 domain is active only in the context of the full-length LeuRS enzyme. The present work is the first study to examine the activity of an isolated editing domain from a class II synthetase.

To determine whether the editing domain of class II E. coli ProRS is functional when separated from the rest of the enzyme, we cloned and expressed the 248-amino acid INS domain as an independent protein. In addition, ΔINS-ProRS, a mutant construct with 86% of the INS domain deleted, was also examined for its aminoacylation and editing activities (Fig. 1). The ΔINS-ProRS deletion construct failed to activate cognate proline or non-cognate glycine (data not shown). The crystal structure of Thermus thermophilus ProRS complexed to the prolyl-adenylate intermediate suggests that residues that align with Glu-410, His-413, and Phe-415, which are missing in our ΔINS-ProRS construct, are critical for prolyl-adenylate formation (44). Since the deleted INS domain lies between motifs 2 and 3, which constitute the amino acid binding pocket, it is also possible that the large 169-amino acid deletion results in local misfolding of the active site. This structural change must be quite subtle since the active site is sufficiently formed to be able to misactivate alanine, and the CD spectrum shows the overall fold is maintained (Fig. 2A). Despite its inability to activate proline, the ΔINS-ProRS construct was able to misactivate alanine but was defective in pretransfer editing of Ala-AMP (Fig. 3). This observation is consistent with a previous study showing that residue Asp-350 in the INS domain is critical for pretransfer editing activity as D350A-ProRS is able to activate alanine but is completely defective in hydrolyzing Ala-AMP (18). The ΔINS deletion construct was not active in aminoacylation or in post-transfer editing (Fig. 5) either alone or in the presence of the isolated INS domain. The lack of post-transfer editing is also consistent with a previous report showing that deletion of 65% of the bacterial A. aeolicus ProRS insertion sequence abolished post-transfer editing (33).

In contrast to the ΔINS-ProRS construct, the cloned INS domain did carry out post-transfer editing of Ala-microhelixPro (Fig. 5). The anticodon domain of tRNA^{Pro} is an important site of interaction with ProRS (27, 28, 42, 45), and the inability of the INS domain to deacetyl full-length mischarged tRNA may be due to the absence of binding determinants that facilitate correct positioning of the tRNA acceptor stem in the editing active site. To ensure that the observed microhelix deacetylation by the isolated INS domain was dependent upon the same active site residues as in the full-length protein (18), we showed that a K279A mutation abolished post-transfer editing (Fig. 5).
without disrupting the protein fold (Fig. 2B). The demonstration of microhelix editing by the ProRS INS domain emphasizes the modular architecture of synthetases and tRNAs as well as the use of appended or extra domains for amino acid editing (21–23). Although class I E. coli IleRS is also capable of deacylating a mischarged minihelixIle (46), class II E. coli AlaRS does not edit Gly-minihelixAla (47). In both cases, it has been shown that the entire tRNA structure is required to activate the editing reaction (47, 48).

Based on phylogenetic and structural arguments, it has been hypothesized that the class II ThrRS and ProRS editing domains may have been added later in evolution than the class I CP1 domains or the class II AlaRS editing domain (20, 23). The latter may be one of the most ancient editing domains and appears to have co-evolved with the active site domain (20). The unique editing domain of class II ProRS is not conserved through evolution but is present in the majority of prokaryotic-like ProRSs (27–29). A similar distribution is found with the ThrRS editing domain, which is not present in Archaea but is found in bacteria and eukaryotes.

Despite the lack of a domain homologous to the INS domain, some eukaryotic-like ProRSs, such as the archaeabacterial enzymes from Methanococcus jannaschii (30) and Methanobacterium thermoautotrophicum,2 have been shown to possess pre-transfer editing activity similar to the E. coli enzyme as well as weak post-transfer activity (30). In contrast, human ProRS lacks an INS domain, and no editing activity was detected (30).

The eukaryotic-like ProRSs examined to date have an ∼10-fold higher initial specificity for their cognate amino acid than the prokaryotic enzymes (30). In the latter case, when the relative in vivo concentration of proline and alanine are taken into account, the “effective discrimination factor” is only about 1200 (17, 49), which is below the threshold value (of ∼3300) where editing is expected to be required in vivo (50). The lack of accurate data on cellular concentrations of amino acids makes it difficult to predict the need for editing in Eukaryotae and Archaea.

The weaker editing activity detected in some of the ProRSs from the eukaryotic-like group may be a functional remnant of an activity that is no longer essential. Some enzymes within this group may therefore have either lost or never needed alanine editing capability. Alternatively a separate protein may act as a cofactor to hydrolyze mischarged products in these species. Our observation that the H. influenzae YbaK protein (HI1434), which shares significant homology with the prokaryotic ProRS INS domain, can carry out post-transfer editing of full-length Ala-tRNAPro and Ala-microhelixPro but not cognate Pro-tRNAPro (Fig. 6) provides support for this proposal. Fig. 7A shows an alignment of the E. coli ProRS INS domain with the H. influenzae YbaK protein. The residues highlighted in purple indicate a-helices (upper case) and b-sheets (lower case) as shown in B.

---

2 M. Nagan and K. Musier-Forsyth, unpublished observations.
sequences of the YbaK family. As shown in the structure of the
*H. influenzae* protein (Fig. 7B), this residue (Lys-46) is located in
a crevice formed by a highly curved mixed seven-stranded
β-sheet motif (35).

The present day function of the microbial YbaK protein family
(also known as EheC) is still unknown, and the structure of
*H1434* provides few insights into this open question, although
a remote structural relationship between *H1434* and C-lectins
has been noted (35). Using the Conserved Domain Architecture
Retrieval Tool (CDART) (ncbi.nlm.nih.gov), in addition to the
bacterial ProRS INS domain, 155 proteins from bacteria, 14
from Eukaryotae, and two from Archaea are identified as being
homologous to HI1434. The thermophilic Archaea *Aeropyrum
pernix* is an example of an organism that possesses a YbaK
homolog along with a ProRS that lacks the INS domain. In this
case, it is reasonable to hypothesize that the YbaK homolog,
which shares 32% sequence identity over 105 residues with the
homolog along with a ProRS that lacks the INS domain. In this
investigation.

Acknowledgments—We are grateful to Dr. Osnat Herzberg at the
University of Maryland for the gift of plasmid pCYB2_H1434 express-
ing the YbaK fusion protein. We also thank Dr. Mahadevan Seetharaman for
preparation of Fig. 7; Dr. Gianlugi Veglia, Nathan Lockwood, and
Bethany Buck for assistance with the CD analysis; and Drs. Joseph
Chihade and Mahadevan Seetharaman for critical reading of the manu-
script and helpful discussions.

REFERENCES
1. Freist, W. (1989) *Biochemistry* **28**, 6787–6795
2. Orgel, L. E. (1963) *Proc. Natl. Acad. Sci. U. S. A.* **50**, 517–521
3. Laughrue, M. (1982) *Exp. Gerontol.* **17**, 305–317
4. Freist, W., Sternbach, H., Pardowitz, I., and Cramer, F. (1998) *J. Theor. Biol.* **193**, 19–38
5. Jakubowski, H., and Goldman, E. (1992) *Microbiol. Rev.* **56**, 412–429
6. Baldwin, A. N., and Berg, P. (1966) *J. Biol. Chem.* **241**, 839–845
7. Eldred, E. W., and Schimmel, P. R. (1972) *J. Biol. Chem.* **247**, 2961–2964
8. Engelsch, S., Engelsch, U., von der Haar, F., and Cramer, P. (1996) *Nucleic Acids Res.* **14**, 7529–7539
9. Chen, J.-F., Guo, N.-N., Li, T., Wang, E.-D., and Wang, Y.-L. (2000) *Biochemistry* **39**, 6726–6731
10. Lincecum, T. L., and Martinis, S. A. (2000) *SAAS Bull. Biochem. Biotechnol.* **13**, 23–33
11. Mursinna, R., Lincecum, T. L., and Martinis, S. A. (2001) *Biochemistry* **40**, 5376–5381
12. Starzyk, R. M., Webster, T. A., and Schimmel, P. (1987) *Science* **237**, 1614–1618
13. Schmidt, E., and Schimmel, P. (1995) *Biochemistry* **34**, 11204–11210
14. Lin, L., Hale, S. P., and Schimmel, P. (1996) *Nature* **384**, 33–34
15. Dock-Bregenon, A. C., Sankaranarayanan, R., Rompy, P., Caillet, J., Springer, M., Rees, B., Francklyn, C. S., Ehresmann, C., and Moras, D. (2000) *Cell* **103**, 877–884
16. Sankaranarayanan, R., Dock-Bregenon, A.-C., Rees, B., Bovee, M., Caillet, J., Rompy, P., Francklyn, C. S., and Moras, D. (2000) *Nat. Struct. Biol.* **7**, 461–465
17. Beuning, P. J., and Musier-Forsyth, K. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97**, 8916–8920
18. Wong, F. C., Beuning, P. J., Nagan, M. C., Shiba, K., and Musier-Forsyth, K. (2002) *Biochemistry* **41**, 7108–7115
19. Tsu, W.-C., and Fersht, A. R. (1981) *Nucleic Acids Res.* **9**, 4627–4637
20. Beebe, K., Ribas de Pouplana, L., and Schimmel, P. (2003) *EMBO J.* **22**, 668–675
21. Schimmel, P., and Ribas de Pouplana, L. (2000) *Trends Biochem. Sci.* **25**, 207–209
22. Alexander, R. W., and Schimmel, P. (2001) *Prog. Nucleic Acids Res. Mol. Biol.* **69**, 317–349
23. Schimmel, P., and Ribas de Pouplana, L. (2001) *Cold Spring Harbor Symp. Quant. Biol.* **66**, 161–166
24. Buechter, D. D., and Schimmel, P. (1995) *Biochemistry* **34**, 6014–6019
25. Augustine, J., and Francklyn, C. S. (1997) *Biochemistry* **36**, 3473–3482
26. Rho, S. B., Lincecum, T. L., Jr., and Martinis, S. A. (2002) *EMBO J.* **21**, 6874–6881
27. Cusack, S., Yaremchuk, A., Krikiliviv, I., and Tukalo, M. (1998) *Structure* **6**, 101–108
28. Stehlin, C., Burke, B., Yang, F., Liu, H., Shiba, K., and Musier-Forsyth, K. (1998) *Biochemistry* **37**, 8605–8613
29. Ribas de Pouplana, L., Brown, J. R., and Schimmel, P. (2001) *J. Mol. Evol.* **53**, 261–268
30. Beuning, P. J., and Musier-Forsyth, K. (2001) *J. Biol. Chem.* **276**, 30773–30785
31. Yaremchuk, A., Cusack, S., and Tukalo, M. (2000) *EMBO J.* **19**, 4745–4758
32. Kamtekar, S., Kennedy, W. D., Wang, J., Stathopoulos, C., Soll, D., and Steitz, T. A. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100**, 1673–1678
33. Ahel, I., Stathopoulos, C., Ambrogelly, A., Sauerwald, A., Toogood, H., Hartsch, T., and Soll, D. (2002) *J. Biol. Chem.* **277**, 34473–34478
34. Wolf, Y. I., Arvindl, L., Grishin, N. V., and Roonin, E. V. (1999) *Genome Res.* **9**, 698–710
35. Zhang, H., Huang, K., Li, Z., Banerjei, L., Fisher, K. E., Grishin, N. V., Eisenstein, E., and Herzberg, O. (2000) *Proteins Struct. Funct. Genet.* **40**, 86–97
36. Liu, H., Yap, L.-P., and Musier-Forsyth, K. (1996) *J. Am. Chem. Soc.* **118**, 2523–2524
37. Yap, L.-P., Stehlin, C., and Musier-Forsyth, K. (1995) *Chem. Biol.* **2**, 661–666
38. Burke, B., Yang, F., Chen, F., Stehlin, C., Chan, B., and Musier-Forsyth, K. (2000) *Biochemistry* **39**, 15540–15547
39. Stehlin, C., Heacock, D. H., Liu, H., and Musier-Forsyth, K. (1997) *Biochemistry* **36**, 2932–2938
40. Fersht, A. R., Ashford, J. S., Bruton, C. J., Jakes, R., Koch, G. L., and Hartley, B. S. (1975) *Biochemistry* **14**, 1–4
41. Musier-Forsyth, K., Searinge, S., Usman, N., and Schimmel, P. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88**, 209–213
42. Liu, H., Peterson, B., and Musier-Forsyth, K. (1995) *Nucleic Acids Res.* **23**, 165–169
43. Heacock, D., Forsyth, C. J., Shiba, K., and Musier-Forsyth, K. (2002) *Biochemistry* **41**, 207–209
44. Yaremchuk, A., Cusack, S., and Tukalo, M. (2001) *EMBO J.* **20**, 384–388
45. Musier-Forsyth, K., and Schimmel, P. (2003) *J. Biol. Chem.* **278**, 45056–45061
46. Nordin, B. E., and Schimmel, P. (1999) *J. Biol. Chem.* **274**, 6835–6838
47. Baum, R., and Rosenthal, H. (1970) *Acta Chem. Scand.* **24**, 2737–2744
48. Fersht, A. R. (1981) *Proc. R. Soc. Lond. Ser. B. Biol. Sci.* **212**, 351–379