

# Restoring species-specific posttransfer editing activity to a synthetase with a defunct editing domain

Julius SternJohn\*, Sanchita Hati<sup>†‡</sup>, Paul G. Siliciano\*, and Karin Musier-Forsyth\*<sup>†§</sup>

Departments of \*Biochemistry, Molecular Biology, and Biophysics and <sup>†</sup>Chemistry, University of Minnesota, Minneapolis, MN 55455

Communicated by Paul R. Schimmel, The Skaggs Institute for Chemical Biology, La Jolla, CA, December 14, 2006 (received for review November 20, 2006)

**Aminoacyl-tRNA synthetases are multidomain proteins responsible for the attachment of specific amino acids to their tRNA substrates. Prolyl-tRNA synthetases (ProRSs) are notable due to their particularly diverse architectures through evolution. For example, *Saccharomyces cerevisiae* ProRS possesses an N-terminal extension with weak homology to a bacterial-specific domain typically present as an insertion (INS) within the aminoacylation active site. The INS domain has been shown to contain a “post-transfer” editing active site responsible for cleaving the aminoacyl-ester bond of misacylated Ala-tRNA<sup>Pro</sup> species. However, wild-type *S. cerevisiae* ProRS does not perform posttransfer editing *in vitro*. Here, we show that replacement of the N-terminal domain of *S. cerevisiae* ProRS with the *Escherichia coli* INS domain confers posttransfer editing function to this chimeric enzyme, with specificity for yeast Ala-tRNA<sup>Pro</sup>. In contrast, the isolated INS domain displays only weak editing activity and lacks tRNA sequence specificity. These results emphasize the modular nature of synthetase editing active sites and demonstrate how in evolution, a weak editing activity can be converted to a more robust state through fusion to the body of a synthetase. In this manner, a single editing module can be distributed to different synthetases, and simultaneously acquire specificity and enhanced activity.**

Prolyl-tRNA synthetase | *Saccharomyces cerevisiae*

Protein synthesis is a dynamic process that occurs within the complex machinery of the ribosome. Aminoacylated or “charged” tRNAs are delivered to the ribosome following amino acid attachment by a family of multidomain enzymes known as aminoacyl-tRNA synthetases (aaRS). Each aaRS possesses an ancient catalytic core domain, which functions to activate a specific amino acid and attach it to either the 2' or 3' hydroxyl of the terminal ribose associated with its cognate tRNA molecules (1). In some cases, the inability to discriminate between noncognate amino acids that are isosteric to or smaller than the desired substrate results in misactivated amino acids and mischarged tRNAs. To ensure high fidelity of genetic code translation, some synthetases have evolved error-correcting mechanisms known as “pretransfer” editing (hydrolysis of misactivated adenylates) and “posttransfer” editing (hydrolysis of mischarged tRNAs). Whereas posttransfer editing has been shown to occur in an active site that is distinct from the site of aminoacylation (2–7), the site of pretransfer editing is less clear. In the case of some synthetases such as isoleucyl-tRNA synthetase, the pretransfer editing reaction has also been suggested to take place in the posttransfer editing domain (8). In contrast, in the case of prolyl-tRNA synthetase (ProRS), pretransfer editing appears to occur in the aminoacylation active site (9).

Posttransfer editing domains are found in both classes of aminoacyl-tRNA synthetase. The class I synthetases specific for isoleucine, leucine, and valine share a highly conserved editing motif known as connective polypeptide 1 (CP1), which is inserted into the Rossmann-fold catalytic domain (3, 10–13). CP1 is present in all extant species and is believed to be of ancient origin (14). The class II synthetases, ProRS (15, 16), alanyl-tRNA synthetase (AlaRS) (17, 18), threonyl-tRNA synthetase (ThrRS) (19, 20), and phenylalanyl-tRNA synthetase (21), contain a

variety of editing domains that are unrelated to CP1. In the case of AlaRS, phylogenetic analysis suggested that, as for class I enzymes, the editing domain is present throughout evolution and co-evolved with the aminoacylation active site (18); it is thus believed to be of ancient origin. Bacterial and eukaryotic ThrRSs, contain an editing domain with weak homology to the AlaRS editing domain, whereas the archaeal editing domain is distinct (19, 22, 23). The latter was recently shown to be structurally similar to D-amino acid deacylases found in eubacteria and eukaryotes (24), and is found as a free-standing editing module in some species (23).

The unique ProRS editing domain is found in at least three different structural contexts: as an insertion (INS) between motifs 2 and 3 of the aminoacylation active site of bacterial enzymes (16, 25, 26), as an N-terminal extension in ProRSs of lower eukaryotes (27, 28), and as a free-standing editing module (25, 28). A highly conserved lysine (K279 in Ec ProRS) found in most INS domain homologs is critical for posttransfer editing function (16, 29). Archaeal and higher eukaryotic ProRSs lack an INS-like editing domain (Fig. 1A). Based on phylogenetic and structural arguments, the bacterial ProRS editing domain has been hypothesized to be a late addition in evolution (14, 18). However, the possibility exists that the domain is also ancient and was lost in some cases.

In this work, we investigate the posttransfer editing activity of *Saccharomyces cerevisiae* (Sc) ProRS, a lower eukaryotic enzyme containing an N-terminal extension with weak homology to the bacterial editing domain. Our data support the notion that the N-terminal domain of Sc ProRS is a defunct editing module that was likely inactivated during evolution, but nevertheless retained due to its role in optimizing tRNA aminoacylation efficiency. The lack of posttransfer editing by wild type Sc ProRS allowed investigation of the effect of fusing an active posttransfer editing module derived from the *Escherichia coli* (Ec) INS domain to the N terminus of a nonediting synthetase. Taken together, this work emphasizes the modular nature of editing domains, as well as the important contribution of synthetase domains outside the editing active site to enhancing editing efficiency and tRNA specificity.

## Results

**Sc ProRS N Terminus, *Haemophilus influenzae* (Hi) YbaK Protein, and Ec INS Domain Share Strong Structural Homology.** The known posttransfer editing domain (INS) of Ec ProRS does not appear

Author contributions: J.S., S.H., P.G.S., and K.M.-F. designed research; J.S. and S.H. performed research; J.S., S.H., P.G.S., and K.M.-F. analyzed data; and J.S., S.H., P.G.S., and K.M.-F. wrote the paper.

The authors declare no conflict of interest.

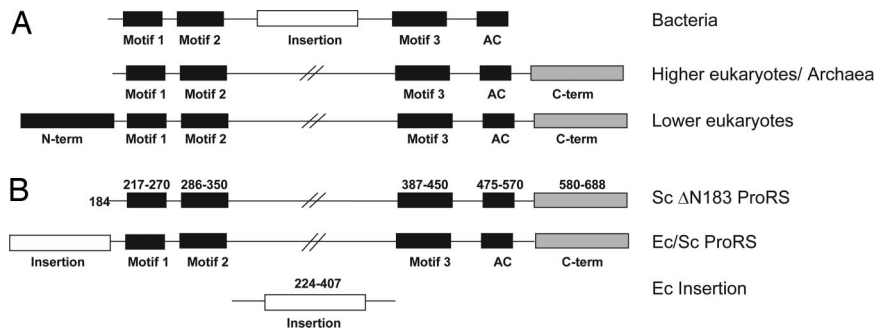
Abbreviations: ProRS, prolyl-tRNA synthetase; CP1, connective polypeptide 1; AlaRS, alanyl-tRNA synthetase; INS, insertion; Sc, *Saccharomyces cerevisiae*; Ec, *Escherichia coli*; Hi, *Haemophilus influenzae*.

<sup>†</sup>Present address: Department of Chemistry, University of Wisconsin, Eau Claire, WI 54702-4004.

<sup>§</sup>To whom correspondence should be sent at the present address: Department of Chemistry, Ohio State University, Columbus, OH 43210. E-mail: musier@chemistry.ohio-state.edu.

This article contains supporting information online at [www.pnas.org/cgi/content/full/0611110104/DC1](http://www.pnas.org/cgi/content/full/0611110104/DC1).

© 2007 by The National Academy of Sciences of the USA

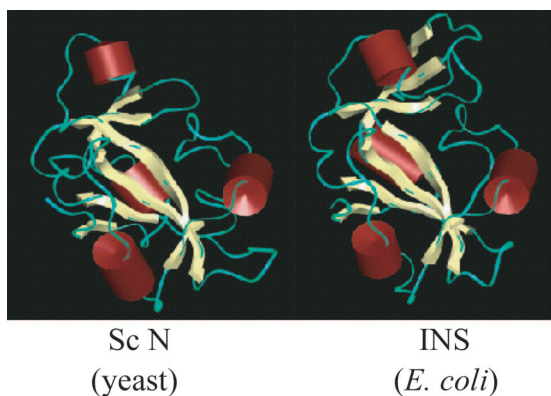


**Fig. 1.** Schematic representation of the domain architecture of ProRS enzymes. (A) ProRS enzymes from all three domains of life. The class II consensus motifs and anticodon domains are shown in black and the bacterial INS domain is indicated as a white box. The C-terminal extension domains present in eukaryotic and archaeal ProRSs, which are not homologous to the INS domain, are shown in gray. The N-terminal extension unique to lower eukaryotic ProRSs is indicated as a white box with black stripes to indicate weak homology with the bacterial INS domain. (B) ProRS constructs used in this work: Sc ProRS with a deletion of the N-terminal 183 residues (Sc  $\Delta$ N183 ProRS, *Top*); chimera with the Ec INS domain substituted for the first 183 aa of Sc ProRS (Ec/Sc ProRS, *Middle*); the Ec ProRS INS domain (Ec Insertion, *Bottom*).

to share homology to any of the other known synthetase-editing domains. Sequence alignments have revealed that yeast and other lower eukaryotic ProRSs contain an “extra” N-terminal domain with weak homology to the bacterial ProRS INS domain (Fig. 1A) (15, 27, 28). Although the extra N-terminal domain common to lower eukaryotes is absent from higher eukaryotes, both share a C-terminal extension as a common feature (in addition to the class-defining active-site core and anticodon binding domain).

The N terminus of Sc ProRS also shares weak sequence similarity to the free-standing Hi YbaK protein (27). Although sequence identity is relatively low (13%), to explore structural similarity, we prepared a homology model of the Sc ProRS N-terminal extension based on the known structure of Hi YbaK (27). A model of the Ec INS domain was generated in a similar manner. The secondary structures of the proteins were determined based on Kabsch–Sander secondary structure definitions (30). The three-dimensional models show overall similarity in secondary structure, with the same number and arrangement of  $\alpha$ -helices and  $\beta$ -sheets (Fig. 2). Based on this analysis, we conclude that, despite the relatively low sequence identity, the N terminus of Sc ProRS and the INS domain of Ec ProRS are likely to exhibit significant structural homology with the Hi YbaK protein.

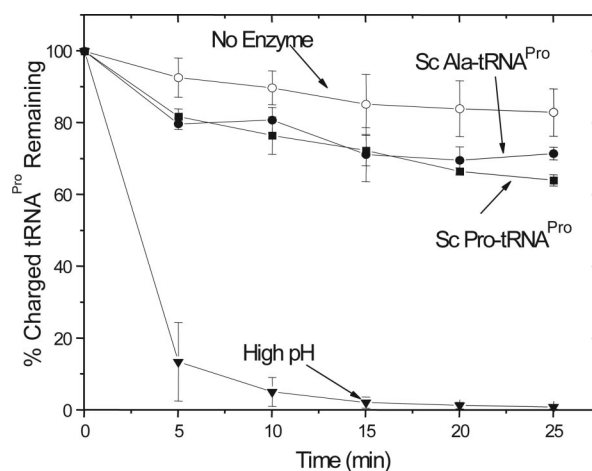
**Probing Sc ProRS Posttransfer Editing Activity.** To investigate WT Sc ProRS posttransfer editing activity, a Sc tRNA<sup>Pro</sup> U70/A73 double mutant was generated to facilitate high levels of alanine



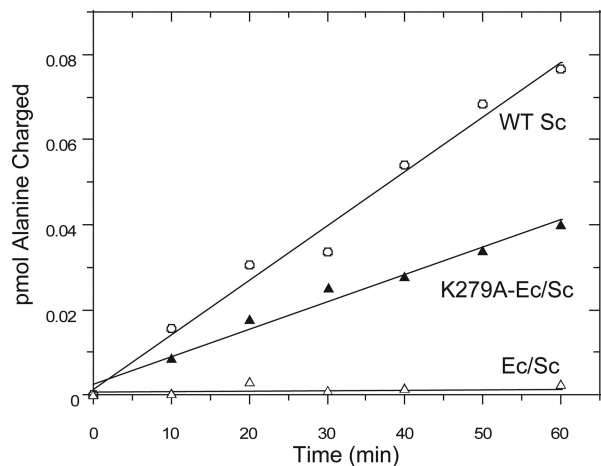
**Fig. 2.** Homology models. Models of the Ec INS domain (*Right*) and the Sc ProRS N terminus (*Left*) generated by using the known x-ray structure of the Hi YbaK protein (27).

mischarging by Ec AlaRS. As has been reported for Ec Ala-tRNA<sup>Pro</sup> (28), Sc U70/A73-Ala-tRNA<sup>Pro</sup> was not a substrate for deacylation by Sc ProRS (Fig. 3). As expected, cognate Sc Pro-tRNA<sup>Pro</sup> was also not a substrate for the yeast enzyme (Fig. 3). Thus, despite the fact that the N terminus of yeast ProRS displays significant structural homology to the known editing modules, Ec ProRS INS and Hi YbaK, WT Sc ProRS lacks the capacity to hydrolyze mischarged Ala-tRNA<sup>Pro</sup>. Consistent with this result, weak misacylation with alanine of WT Sc tRNA<sup>Pro</sup> by yeast ProRS is observed (Fig. 4). The extent of misacylation by the yeast enzyme is low, which indicates that pretransfer editing may be the predominant mode of clearing misactivated alanine by Sc ProRS (9).

**Ec/Sc Chimera Aminoacylates Sc tRNA<sup>Pro</sup> and Gains Posttransfer Editing Function.** Although Sc ProRS does not appear to possess posttransfer editing activity *in vitro*, the high structural homology between the Ec INS domain and the Sc N terminus suggests that the latter may be the remnant of a once functional editing domain. To probe this idea further, a chimeric Ec/Sc ProRS enzyme was constructed, wherein the first 183 aa of Sc ProRS



**Fig. 3.** Deacylation by WT Sc ProRS. Deacylation of Sc U70/A73-[<sup>3</sup>H]Ala-tRNA<sup>Pro</sup> (filled circle) or Sc [<sup>3</sup>H]Pro-tRNA<sup>Pro</sup> (filled square) by 4  $\mu$ M Sc ProRS. A background (no protein) reaction was performed for all tRNAs tested and all were within 11% of the representative background data shown (open circle). An additional reaction (filled inverted triangle) was performed by incubating Sc U70/A73-[<sup>3</sup>H]Ala-tRNA<sup>Pro</sup> with 0.2 M NaOH (pH 13). The plots represent the average of at least three assays.



**Fig. 4.** Aminoacylation assays. Reaction time courses showing charging of alanine onto Sc tRNA<sup>Pro</sup> by WT Sc ProRS (open circle), Ec/Sc chimera (open triangle), and K279A-Ec/Sc chimera (filled triangle). Assays were performed at 30°C using 12  $\mu$ M enzyme and 8  $\mu$ M WT Sc tRNA<sup>Pro</sup>. The plots represent the average of two assays with the values differing by <15%.

were replaced with residues 224–407 of the Ec INS domain (Fig. 1B, Ec/Sc chimera). An N-terminal deletion mutant lacking the first 183 residues of Sc ProRS was also prepared (Fig. 1B, Sc  $\Delta$ N183 ProRS). To establish whether the N terminus of yeast ProRS contributes to aminoacylation activity, the  $\Delta$ N183 variant was tested for its ability to charge a WT Sc tRNA<sup>Pro</sup> transcript *in vitro*. Interestingly, the  $k_{cat}/K_M$  for cognate proline charging was only 3-fold reduced relative to charging by WT Sc ProRS (9). Similar reductions were observed in the  $k_{cat}/K_M$  for proline (5-fold) and alanine (2-fold) activation upon deletion of the N-terminal domain (Table 1). These results suggest that the N terminus of yeast ProRS contributes modestly to amino acid activation and aminoacylation *in vitro*. The N-terminal truncation mutant was also tested for alanine mischarging activity. Very weak mischarging was detected, but no increase in this activity was observed relative to the WT enzyme (Fig. 4 and data not shown). Taken together, these results suggest that the N terminus of yeast ProRS contributes modestly to amino acid activation and aminoacylation *in vitro*, but plays no role in posttransfer editing.

The Ec/Sc chimera was able to charge Sc tRNA<sup>Pro</sup> transcripts with proline. Similar to the  $\Delta$ N183 variant, substitution of the N-terminal domain of Sc ProRS with the INS domain of Ec ProRS resulted in an  $\approx$ 3-fold reduction in aminoacylation

**Table 1. Kinetic parameters for activation of proline and alanine by WT and truncated Sc ProRS and WT Ec ProRS**

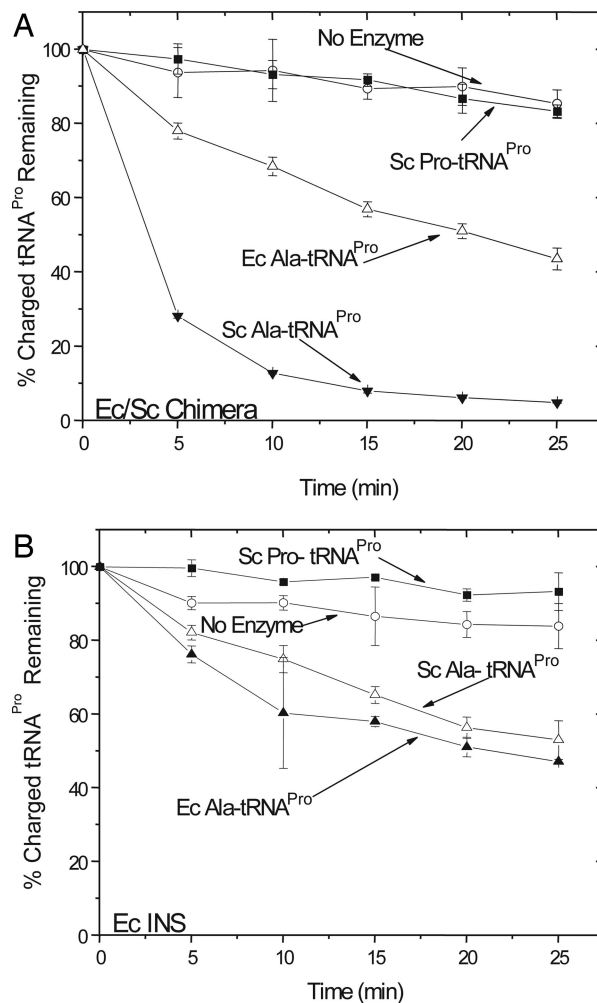
ProRS	Amino acid	$k_{cat}$ , s <sup>-1</sup>	$K_M$ , mM	$k_{cat}/K_M$ , s <sup>-1</sup> mM <sup>-1</sup>	$k_{cat}/K_M$ , relative*
Sc-WT <sup>†</sup>	Proline	8.83	2.32	3.73	1
Sc-WT <sup>†</sup>	Alanine	0.114	5.0 $\times$ 10 <sup>3‡</sup>	2.3 $\times$ 10 <sup>-5</sup>	6 $\times$ 10 <sup>-6</sup>
Sc- $\Delta$ N <sup>†</sup>	Proline	2.85	3.77	0.74	1
Sc- $\Delta$ N <sup>†</sup>	Alanine	0.071	5.9 $\times$ 10 <sup>3‡</sup>	1.2 $\times$ 10 <sup>-5</sup>	1.6 $\times$ 10 <sup>-5</sup>
Ec WT <sup>§</sup>	Proline	70	0.25	280	1
Ec WT <sup>§</sup>	Alanine	1.7	140	0.013	5 $\times$ 10 <sup>-5</sup>

\* $k_{cat}/K_M$  is relative to proline in each set, which was set to 1.0.

<sup>†</sup>Each assay was performed at least twice with the values differing by <30%.

<sup>‡</sup>Values given are estimates using 200–800 mM alanine; more accurate determinations require prohibitively high amino acid concentrations.

<sup>§</sup>As reported in ref. 17.



**Fig. 5.** Deacylation by the Ec/Sc chimera and the Ec INS domain. (A) Deacylation of Sc U70/A73-[<sup>3</sup>H]Ala-tRNA<sup>Pro</sup> (filled inverted triangle), Sc [<sup>3</sup>H]Pro-tRNA<sup>Pro</sup> (filled square), and Ec [<sup>3</sup>H]Ala-tRNA<sup>Pro</sup> (open triangle) in the presence of Ec/Sc chimeric ProRS (1  $\mu$ M). (B) Deacylation of Sc U70/A73-[<sup>3</sup>H]Ala-tRNA<sup>Pro</sup> (open triangle), Sc [<sup>3</sup>H]Pro-tRNA<sup>Pro</sup> (filled square), and Ec [<sup>3</sup>H]Ala-tRNA<sup>Pro</sup> (filled triangle) in the presence of the Ec INS domain (20  $\mu$ M). A background (no protein, open circle) reaction was performed for all tRNAs tested, and all of the plots present the average of at least three assays.

activity ( $k_{cat}/K_M = 0.0116 \pm 0.003$  s<sup>-1</sup>  $\mu$ M<sup>-1</sup>) relative to WT Sc ProRS ( $k_{cat}/K_M = 0.0308 \pm 0.005$  s<sup>-1</sup>  $\mu$ M<sup>-1</sup>).

The Ec/Sc chimera was next tested for posttransfer editing activity *in vitro*. The chimera (1  $\mu$ M) rapidly deacylated mischarged Sc U70/A73-Ala-tRNA<sup>Pro</sup>, while preserving the cognate Pro-tRNA<sup>Pro</sup> species (Fig. 5A). The chimera demonstrated species-specificity, as much weaker deacylation activity was observed in the presence of mischarged Ec Ala-tRNA<sup>Pro</sup>. We estimate that the rate of deacylation of Ec tRNA<sup>Pro</sup> was  $\approx$ 8-fold lower than that of Sc tRNA<sup>Pro</sup>. The separately cloned Ec INS domain (Fig. 1B) was previously shown to weakly deacylate a mischarged Ala-microhelix<sup>Pro</sup> variant (25). High concentrations (20  $\mu$ M) of purified Ec INS were also required to deacylate Ec Ala-tRNA<sup>Pro</sup>, and very similar levels of Sc Ala-tRNA<sup>Pro</sup> deacylation by the INS module were observed (Fig. 5B). Thus, whereas the cloned INS domain displays little, if any, sequence preference, the Ec/Sc chimera represents a gain of function mutant with species specificity for Sc tRNA<sup>Pro</sup>. Additionally, this result shows that the activity of the INS domain fused to ProRS *in cis*, is significantly greater than its activity as a *trans* editing protein.



**Deacylation by the Ec/Sc Chimera Depends on the K279 Residue.** To establish whether the observed posttransfer editing activity of the Ec/Sc chimera depends on the strictly conserved K279 residue shown to be critical for editing by Ec ProRS (16), we mutated this residue to alanine in the context of the chimera and tested the alanine mischarging efficiency. As described earlier, under conditions of high enzyme and tRNA concentration, WT Sc ProRS can mischarge alanine onto Sc-tRNA<sup>Pro</sup> transcripts. As expected, the Ec/Sc chimera prevents any detectable mischarging (Fig. 4). In contrast, in the presence of the K279A Ec/Sc chimera, significant mischarging is observed (Fig. 4). Thus, in the context of the chimera, the INS domain activity depends on K279, just as it does when present as an internal domain of Ec ProRS. This result suggests that the mechanism of posttransfer editing by the N-terminal INS domain is similar to that of the internal editing module.

**Role of the N-Terminal Domain of Sc ProRS *in Vivo*.** Although deletion of the N-terminal domain of Sc ProRS had only minor effects on the *in vitro* kinetic parameters for tRNA aminoacylation and amino acid activation, the N terminus may still contribute to an additional function *in vivo*. To probe this hypothesis, the ability of  $\Delta$ N183 Sc ProRS to complement a Sc ProRS null strain was examined. Under minimal growth conditions, the truncation mutant was fully functional [see [supporting information \(SI\) Fig. 6](#)]. Furthermore, the Ec/Sc chimera also appeared fully functional *in vivo* and maintained growth as effectively as WT and  $\Delta$ N183 Sc ProRS at 30°C ([SI Fig. 6](#)). Complementation assays were also performed in the presence of 1 M alanine in rich media. Under these conditions, cell growth was significantly reduced. However, these experiments also revealed no difference between WT Sc ProRS and the chimeric enzyme (data not shown).

## Discussion

Unlike the CP1 editing module found in class I synthetases or the class II AlaRS editing domain, the bacterial ProRS editing domain is not conserved through evolution. However, free-standing proteins with homology to the INS domain are found in the genomes of representative species from all three kingdoms of life (25, 28). Interestingly, there appear to be at least two distinct classes of ProRS-like editing modules with different specificities. Whereas the Ec ProRS INS domain (16, 25) and the *Clostridium sticklandii* PrdX protein (28) edit Ala-tRNA<sup>Pro</sup>, the Hi YbaK protein functions primarily as a Cys-tRNA<sup>Pro</sup> deacylase (29), with only weak activity toward Ala-tRNA<sup>Pro</sup> (25, 31).

In addition to these free-standing ProRS-like editing proteins, a module with homology to the INS domain is appended to the N terminus of yeast and other lower eukaryotic ProRSs (27, 28). Previously, *Plasmodium falciparum* ProRS, a lower eukaryotic enzyme with an N-terminal INS-like domain, was shown to catalyze editing of Ec Ala-tRNA<sup>Pro</sup>, whereas Sc ProRS lacked this activity (28). The latter result is not surprising, given the low cross-species aminoacylation demonstrated by Sc ProRS (J.S., P.S., and K.M.-F., unpublished observations). Here, we use mischarged Sc Ala-tRNA<sup>Pro</sup> to confirm that Sc ProRS appears to lack posttransfer editing function. We also show that the presence of the N-terminal domain plays a minor role in enhancing *in vitro* adenylate formation (Table 1) and tRNA aminoacylation (9). The specificity of amino acid activation (i.e., relative  $k_{\text{cat}}/K_M$  for cognate proline vs. alanine) is also enhanced  $\approx$ 3-fold *in vitro* (Table 1). However, because the truncated Sc  $\Delta$ N183 ProRS complements yeast ProRS null strain as effectively as the WT enzyme, the role of the N-terminal domain *in vivo* is still unclear. The possibility exists that the N-terminal domain plays a nonessential or redundant function that is not detected in our *in vivo* complementation assay, such as facilitating interactions with other synthetases. We have shown that

this domain is not associated with Sc ProRS editing function and based on a blast (blastp and CDART) search, there does not appear to be a free-standing INS-like editing domain homolog in *S. cerevisiae*. Thus, posttransfer editing does not appear to be required in this organism, and the N-terminal domain may have been retained based on its role in optimizing the synthetic step.

The structure of a prokaryotic ProRS possessing an INS domain has recently been reported (32). The posttransfer editing active site is poised over the aminoacylation catalytic domain, with  $\approx$ 35-Å separation between the two active sites. The results reported here suggest that the N-terminal domain of Sc ProRS, an enzyme of unknown structure, may be similarly oriented even though its function has diverged.

Recently, it was shown that human mitochondrial leucyl-tRNA synthetase appears to possess a defunct CP1 editing domain (33). In addition, it was reported that a “single sieve” editing mechanism efficiently discriminates against noncognate isoleucine based on the enzyme’s enhanced specificity for activation of cognate leucine relative to Ec leucyl-tRNA synthetase (33). Similarly, in the case of Sc ProRS, we estimate that the  $K_M$  for alanine is  $\approx$ 2,000-fold greater than that of proline (Table 1). This, together with an  $\approx$ 80-fold more favorable  $k_{\text{cat}}$  value for proline activation, likely precludes the need for posttransfer editing in this system. Moreover, the overall amino acid discrimination factor (i.e.,  $k_{\text{cat}}/K_M^{\text{Pro}}/k_{\text{cat}}/K_M^{\text{Ala}}$ ) measured for Sc ProRS is enhanced at least 10-fold relative to the Ec enzyme (Table 1), and is comparable to the level of discrimination achieved by *M. jannaschii* and human ProRS, which lack an INS-like domain (34).

A sequence alignment of the N-terminal domain of lower eukaryotic ProRSs was performed along with the Ec INS domain (see [SI Fig. 7](#)). The highly conserved nature of the lysine residue that aligns with K279 of Ec ProRS is striking. This residue was previously shown to be critical for posttransfer editing of both Ec ProRS (16) and Hi YbaK (29), and we confirmed in this work that it also plays a critical role in editing by the EC/SC chimera. Two other residues are highly conserved among these domains: K284 and D350 (Ec ProRS numbering). Although the function of K284 has not yet been probed, D350 was also shown to be a key determinant for editing by the Ec enzyme (16). Although Sc ProRS contains these three residues, a comparison of the sequences shows that it also lacks several residues previously shown to be important for posttransfer editing by the bacterial INS domain, including H369 (16) and G336 (C. Silvers and K.M.-F., unpublished data). The latter residue, which is part of a GXXXP motif, was shown to be part of an oxyanion hole proposed to form a putative ligand binding crevice in Hi YbaK (27). The lack of H369 and the GXXXP motif may explain the lack of detectable posttransfer editing activity in the case of yeast ProRS. In support of this hypothesis, *P. falciparum* ProRS, which possesses editing activity (28), contains both of these elements. The fact that substitution of the N-terminal domain of Sc ProRS with the Ec INS domain restores posttransfer editing function of the Ec/Sc chimera also supports the requirement for an intact GXXXP motif.

As more genome sequence information becomes available, the widespread presence of synthetase-like proteins is becoming more evident (35). In particular, the functional and evolutionary importance of these free standing synthetase-like modules is of great interest. Although some of these single-domain homologs have been shown to carry out functions unrelated to tRNA charging (35, 36) others, such as YbaK and PrdX, have clearly been shown to possess aminoacylation or editing activity (22, 23, 25, 28, 29, 31, 37). Free-standing editing proteins found in nature appear to have evolved alternate mechanisms to function as efficient and specific editing modules *in trans* (23, 28, 29, 31). Precisely how they accomplish this *in vivo* is an open question, but plausible mechanisms include noncovalent interactions with synthetases (38) or other factors.

In contrast, cloned editing modules derived from both classes of synthetases and expressed as free-standing proteins show significantly lower activity (3, 25) or lack editing activity altogether (11). It has been speculated that some editing domains have been incorporated into synthetases to increase editing efficiency and specificity (28). In support of this hypothesis, the free-standing INS domain displays very weak tRNA binding affinity ( $K_d > 5 \mu\text{M}$ , data not shown) and lacks tRNA specificity. However, in the case of the chimera, the domain is at least operationally acting in *cis*, through association with a core synthetase domain that serves to enhance the local concentration of tRNA around the editing module. The fact that this domain can be readily transplanted into the heterologous Sc ProRS framework and confer efficient and specific editing function on a non-posttransfer editing synthetase is also a remarkable demonstration of the modular nature of editing domains. These results suggest how a generalized editing module can be distributed to different synthetases, and simultaneously acquire specificity and activity. The enhanced activity and newly acquired tRNA specificity is a strong evolutionary driving force for *trans* to *cis* conversion of the editing domain, with adaptation to different amino acids occurring through accumulation of mutations.

### Materials and Methods

All chemicals and amino acids were purchased from Sigma-Aldrich (St. Louis, MO) unless otherwise noted. [ $^3\text{H}$ ]Alanine (54 Ci/mmol) was from Amersham Biosciences (Piscataway, NJ). All restriction enzymes were purchased from New England Biolabs (Ipswich, MA) unless otherwise noted.

**Plasmid Construction for *in Vitro* Expression.** WT yeast ProRS and an N-terminal deletion allele lacking 183 residues ( $\Delta\text{N183}$  Sc ProRS) were cloned into protein expression vector pQE30 (Qiagen, Valencia, CA). The Sc ProRS PCR product was amplified from yeast genomic DNA using primers yPRS SphI/P1 (5'-ACACATG-CAGCGCCTGTTTCGGAAGCGTTTGCC-3') and MBP/P2 (5'-CCCAAGCTTCTAATAAGAACGACCGCAACAT-3'). Both the pQE30 vector and the Sc ProRS PCR product were digested with SphI and HindIII. After digestion, the products were ligated together to generate plasmid pJS550. The  $\Delta\text{N183}$  yeast ProRS gene was PCR amplified via primers  $\Delta\text{N-term183}$  SphI (5'-ACACATG-CATGCCTTATT GGTATCACCGTAGAC-3') and MBP/P2, and cloned into pQE30 to construct pJS371. All constructs were confirmed through sequencing.

The recombinant Ec/Sc chimera construct for protein purification was generated via a stepwise process. This construct contains an N-terminal INS domain derived from Ec ProRS (residues 224–407), appended to residues 184–688 of Sc ProRS. PCR was first performed using plasmid pCS-M1S as a template (39) and primers cc-oJS1-INS 670 (5'-CGCGGATCCTCCGACACCTCTGACTATGCA-3') and cc-oJS2-INS 1221 (5'-CATGCATGCACGTTTGAT CAGCTTACC-3') to generate the Ec INS domain bearing BamHI and SphI restriction sites. Plasmid pJS371 and the PCR product were digested with BamHI and SphI and ligated together to create plasmid pJS1058. All constructs were confirmed through sequencing. The K279A mutant of the Ec/Sc chimera was created using the QuikChange site-directed mutagenesis kit (Stratagene, La Jolla, CA). The mutant plasmid was generated via PCR using WT plasmid pJS1058 as template and primers 5'-GAAAACGGTTGC-GACTCTGCTGG-3' and 5'-CCAGCAGAGTCGCAAC-CGTTTTTC-3'.

**Plasmid Construction for *in Vivo* Expression.** The gene encoding the Ec/Sc was PCR amplified using plasmid pJS1058 as a template and primers oJS-pG-1 chimera PRS (5'-CGCGGATCCGC-CACCATGTCCGACACCTCTGACTATGCA-3') and oJS3-

pG-1 yPRS C-terminal (5'-ACGGTCGACCTAATAAGAAC-GACCGAACAT-3'). The PCR products were digested with BamHI and SalI, and cloned into the high copy yeast expression vector, pG-1 (TRP1 2  $\mu\text{M}$ ) (40), to construct plasmid pJS1085.

The  $\Delta\text{N183}$  Sc PROS gene was PCR amplified using template pJS371 and primers  $\Delta\text{N183}$  pG-1 (5'-CGCGGATCCGCCAC-CATGCTTATTGGTATCACCGTAGAC-3') and oJS3-pG-1 yPRS C-terminal. The PCR product was digested with BamHI and SalI and cloned into pG-1 to construct plasmid pJS380.

The WT Sc PROS gene was amplified via PCR using plasmid pJS550 and primers oJS10-pG-1 N-terminal (5'-CGCGTC-GACGCCACCATGCCTGTTTCGGAAGCGTTTG-3') and oJS11-Sc-pG-1 C-terminal (5'-CGCGTCGACCTAATAA-GAACGACCGAACAT-3'). The PCR product and the yeast expression vector pG-1 were digested with SalI and ligated together to create plasmid pJS505. All constructs were confirmed through sequencing.

**Yeast Strains and Growth Conditions.** All yeast strains in this study (see SI Table 2), with the exception of SSL212, are isogenic with TR2 (41). Yeast transformations were performed either by electroporation or using the lithium acetate transformation method (42). Media, mating, sporulation, and tetrad dissection were carried out according to standard methods (40). The deletion of the yeast PROS chromosomal copy is described in SI Text.

**Protein Preparation.** WT Sc ProRS,  $\Delta\text{N183}$  Sc ProRS, WT *Deinococcus radiodurans* ProRS, WT Ec AlaRS, and the Ec ProRS INS domain were purified using a Talon cobalt affinity resin (Clontech, Mountain View, CA) as described (9, 43). Plasmids encoding the Ec/Sc chimera and the K279A Ec/Sc chimera were transformed into Ec SG13009 [pREP4] (Qiagen) competent cells and protein expression was induced with 0.1 mM isopropyl  $\beta$ -D-thiogalactoside for 22 h at room temperature. Histidine-tagged proteins were purified using a Talon resin according to the manufacturer's protocol. The desired proteins were eluted with 50–200 mM imidazole, concentrated, and stored as described (30). Total protein concentrations were determined by the Bradford assay (Bio-Rad, Hercules, CA) (44). The concentrations of synthetases were determined using the active site titration assay (45). In the case of the INS protein, the concentration was obtained by using the Bradford assay.

**Preparation of tRNA<sup>Pro</sup>.** All tRNAs used in this study were prepared by *in vitro* transcription as described (15, 46, 47). Point mutations in Sc tRNA<sup>Pro</sup> to generate U70/A73-tRNA<sup>Pro</sup> were introduced by using the QuikChange Site-Directed Mutagenesis kit (Stratagene).

**Preparation of Aminoacyl-tRNAs.** Aminoacyl-tRNAs for use in deacylation assays were prepared as described (15) using WT Ec AlaRS (5.0  $\mu\text{M}$ ) to aminoacylate Sc U70/A73-tRNA<sup>Pro</sup> (10  $\mu\text{M}$ ), and WT *D. radiodurans* ProRS (10  $\mu\text{M}$ ) to aminoacylate WT Ec tRNA<sup>Pro</sup> (8  $\mu\text{M}$ ) with [ $^3\text{H}$ ]alanine in buffer containing 50 mM Hepes (pH 7.5), 4 mM ATP, 20 mM KCl, 20 mM 2-mercaptoethanol, 25 mM MgCl<sub>2</sub>, and 0.1 mg/ml BSA. The misacylated tRNAs were purified by repeated phenol extractions, followed by ethanol precipitation. The purified aminoacyl-tRNAs were quantified by scintillation counting and stored at  $-80^\circ\text{C}$  in 50 mM KPO<sub>4</sub> (pH 5.0). Pro-tRNA<sup>Pro</sup> species were prepared by using WT Sc ProRS (4  $\mu\text{M}$ ) to aminoacylate both Sc WT tRNA<sup>Pro</sup> (12  $\mu\text{M}$ ) and U70/A73-tRNA<sup>Pro</sup> (12  $\mu\text{M}$ ) with [ $^3\text{H}$ ]proline as described above. Preparation of Ec Pro-tRNA<sup>Pro</sup> was performed similarly using Ec ProRS (4  $\mu\text{M}$ ).

**ATP-PP<sub>i</sub> Exchange Assays.** To test for activation of proline and alanine, the ATP-PP<sub>i</sub> exchange assay was performed as de-

scribed (43) at 37°C in 144 mM Tris·HCl (pH 7.5), 10 mM MgCl<sub>2</sub>, 10 mM KF, 10 mM 2-mercaptoethanol, 1 mM PP<sub>1</sub>, and 0.2 mg/ml BSA. The following concentrations were used: proline (2–25 mM), alanine (200–800 mM), WT Sc ProRS (10 nM), and ΔN183 Sc ProRS (500 nM).

**Cognate Aminoacylation and Mischarging Assays.** Aminoacylation assays were performed at 30°C as described (18, 47, 48) using purified cognate Sc tRNA<sup>Pro</sup> transcripts (0.5 μM), 100 nM Sc WT ProRS, Ec/Sc chimera, or K279A chimera, and 22.7 μM [<sup>3</sup>H]proline. Mischarging assays in the presence of alanine were also carried out at 30°C using standard conditions (15) with 8 μM Sc tRNA<sup>Pro</sup>, 12 μM enzyme (Sc WT ProRS, Ec/Sc chimera, or K279A chimera), and 22.7 μM [<sup>3</sup>H]alanine.

**Deacylation Assays.** Deacylation assays were carried out at room temperature according to published conditions (15). Reactions contained ≈0.2 μM Ec [<sup>3</sup>H]Ala-tRNA<sup>Pro</sup> or 0.6 μM Sc U70/A73 [<sup>3</sup>H]Ala-tRNA<sup>Pro</sup>, and were initiated with either 20 μM Ec ProRS INS domain, 1 μM Ec/Sc chimera, or 4 μM WT Sc ProRS. A background reaction was carried out in each case in which

buffer (0.15 M KPO<sub>4</sub>, pH 7.0) was used to initiate the reaction. Results are the average of at least three trials.

**Sc ProRS Structure Modeling.** A homology model of the Sc N-terminal domain (residues 9–147) was generated based on the known x-ray structure (residues 31–158) of the Hi YbaK protein (27). The sequence alignment of the target and template proteins was first achieved using CLUSTALW (49), followed by manual editing. Using the MODELLER-4 program (50), 40 models were generated, and the model with the lowest value of the objective function (derived from template protein spatial restraints and the CHARMM force field) was taken as the final comparative model. The stereo-chemical quality of the models was further checked using internal routines in the MODELLER and by PROCHECK (51, 52). A model of the insertion domain (residues 248–390) of Ec ProRS was generated similarly; however, in this case, residues 14–158 of the template YbaK protein were used in the alignment.

We thank Dr. Kathryn Splan and Ms. Kelly Hoadley for assistance with some of the experiments related to this work. This work was supported by National Institutes of Health Grant GM49928.

1. Berg P (1961) *Annu Rev Biochem* 30:293–324.
2. Fersht AR, Dingwall C (1979) *Biochemistry* 18:2627–2631.
3. Lin L, Hale SP, Schimmel P (1996) *Nature* 384:33–34.
4. Nureki O, Vassilyev DG, Tateno M, Shimada A, Nakama T, Fukai S, Konno M, Hendrickson TL, Schimmel P, Yokoyama S (1998) *Science* 280:578–582.
5. Sankaranarayanan R, Dock-Bregeon AC, Romby P, Caillet J, Springer M, Rees B, Ehresmann C, Ehresmann B, Moras D (1999) *Cell* 97:371–381.
6. Schmidt E, Schimmel P (1994) *Science* 264:265–267.
7. Silvan LF, Wang J, Steitz TA (1999) *Science* 285:1074–1077.
8. Nomanbhoy TK, Hendrickson TL, Schimmel P (1999) *Mol Cell* 4:519–528.
9. Hati S, Ziervogel B, SternJohn J, Wong F-C, Nagan MC, Rosen AE, Siliciano PG, Chihade JW, Musier-Forsyth K (2006) *J Biol Chem* 281:27862–27872.
10. Starzyk RM, Webster TA, Schimmel P (1987) *Science* 237:1614–1618.
11. Chen JF, Guo NN, Li T, Wang ED, Wang YL (2000) *Biochemistry* 39:6726–6731.
12. Mursinna RS, Lincecum TL, Jr, Martinis SA (2001) *Biochemistry* 40:5376–5381.
13. Schmidt E, Schimmel P (1995) *Biochemistry* 34:11204–11210.
14. Ribas de Pouplana L, Schimmel P (2001) *Cold Spring Harbor Symp Quant Biol* 66:131–166.
15. Beuning PJ, Musier-Forsyth K (2000) *Proc Natl Acad Sci USA* 97:8916–8920.
16. Wong FC, Beuning PJ, Nagan M, Shiba K, Musier-Forsyth K (2002) *Biochemistry* 41:7108–7115.
17. Tsui WC, Fersht AR (1981) *Nucleic Acids Res* 9:4627–4637.
18. Beebe K, Ribas De Pouplana L, Schimmel P (2003) *EMBO J* 22:668–675.
19. Dock-Bregeon A, Sankaranarayanan R, Romby P, Caillet J, Springer M, Rees B, Francklyn CS, Ehresmann C, Moras D (2000) *Cell* 103:877–884.
20. Sankaranarayanan R, Dock-Bregeon AC, Rees B, Bovee M, Caillet J, Romby P, Francklyn CS, Moras D (2000) *Nat Struct Biol* 7:461–465.
21. Roy H, Ling J, Irnov M, Ibba M (2004) *EMBO J* 23:4639–4648.
22. Beebe K, Merriman E, De Pouplana LR, Schimmel P (2004) *Proc Natl Acad Sci USA* 101:5958–5963.
23. Korencic D, Ahel I, Schelert J, Sacher M, Ruan B, Stathopoulos C, Blum P, Ibba M, Söll D (2004) *Proc Natl Acad Sci USA* 101:10260–10265.
24. Dwivedi S, Kruparani SP, Sankaranarayanan R (2005) *Nat Struct Mol Biol* 12:556–557.
25. Wong FC, Beuning PJ, Silvers C, Musier-Forsyth K (2003) *J Biol Chem* 278:52857–52864.
26. Ahel I, Stathopoulos C, Ambrogelly A, Sauerwald A, Toogood H, Hartsch T, Söll D (2002) *J Biol Chem* 277:34743–34748.
27. Zhang H, Huang K, Li Z, Banerjee L, Fisher KE, Grishin NV, Eisenstein E, Herzberg O (2000) *Proteins* 40:86–97.
28. Ahel I, Korencic D, Ibba M, Söll D (2003) *Proc Natl Acad Sci USA* 100:15422–15427.
29. An S, Musier-Forsyth K (2004) *J Biol Chem* 279:42359–42362.
30. Kabsch W, Sander C (1983) *Biopolymers* 22:2577–2637.
31. Ruan B, Söll D (2005) *J Biol Chem* 280:25887–25891.
32. Crepin T, Yaremchuk A, Tkalco M, Cusack S (2006) *Structure (London)* 14:1511–1525.
33. Lue SW, Kelley SO (2005) *Biochemistry* 44:3010–3016.
34. Beuning PJ, Musier-Forsyth K (2001) *J Biol Chem* 276:30779–30785.
35. Geslain R, Ribas de Pouplana L (2004) *Trends Genet* 20:604–610.
36. Schimmel P, Ribas De Pouplana L (2000) *Trends Biochem Sci* 25:207–209.
37. Campanacci V, Dubois DY, Becker HD, Kern D, Spinelli S, Valencia C, Pagot F, Salomoni A, Grisel S, Vincentelli R, et al. (2004) *J Mol Biol* 337:273–283.
38. An S, Musier-Forsyth K (2005) *J Biol Chem* 280:34465–34472.
39. Stehlin C, Heacock DH, Jr, Liu H, Musier-Forsyth K (1997) *Biochemistry* 36:2932–2938.
40. Guthrie C, Fink GR (1991) *Guide to Yeast Genetics and Molecular Biology* (Academic, San Diego).
41. Parker R, Simmons T, Shuster EO, Siliciano PG, Guthrie C (1988) *Mol Cell Biol* 8:3150–3159.
42. Schiestl RH, Gietz RD (1989) *Curr Genet* 16:339–346.
43. Heacock D, Forsyth CJ, Shiba K, Musier-Forsyth K (1996) *Bioorg Chem* 24:273–289.
44. Bradford MM (1976) *Anal Biochem* 72:248–254.
45. Fersht AR, Ashford JS, Bruton CJ, Jakes R, Koch GL, Hartley BS (1975) *Biochemistry* 14:1–4.
46. Hou YM, Westhof E, Giegé R (1993) *Proc Natl Acad Sci USA* 90:6776–6780.
47. Liu H, Peterson R, Kessler J, Musier-Forsyth K (1995) *Nucleic Acids Res* 23:165–169.
48. Musier-Forsyth K, Scaringe S, Usman N, Schimmel P (1991) *Proc Natl Acad Sci USA* 88:209–213.
49. Thompson JD, Higgins DG, Gibson TJ (1994) *Nucleic Acids Res* 22:4673–4680.
50. Sali A, Blundell TL (1993) *J Mol Biol* 234:779–815.
51. Sali A, Potterton L, Yuan F, van Vlijmen H, Karplus M (1995) *Proteins* 23:318–326.
52. Laskowski RA, MacArthur MW, Moss DW, Thornton JM (1993) *J Appl Crystallogr* 26:283–291.