

Evolution of multiple, mutually orthogonal prolyl-tRNA synthetase/tRNA pairs for unnatural amino acid mutagenesis in *Escherichia coli*

Abhishek Chatterjee^{a,b,1}, Han Xiao^{a,b,1}, and Peter G. Schultz^{a,b,2}

^aDepartment of Chemistry and ^bThe Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, CA 92037

Contributed by Peter G. Schultz, July 20, 2012 (sent for review June 12, 2012)

The site-specific incorporation of unnatural amino acids (UAAs) into proteins in living cells relies on an engineered tRNA/aminoacyl-tRNA synthetase (tRNA/aaRS) pair, orthogonal to the host cell, to deliver the UAA of choice in response to a unique nonsense or frameshift codon. Here we report the generation of mutually orthogonal prolyl-tRNA/prolyl-tRNA synthase (ProRS) pairs derived from an archaeobacterial ancestor for use in *Escherichia coli*. By reprogramming the anticodon-binding pocket of *Pyrococcus horikoshii* ProRS (PhProRS), we were able to identify synthetase variants that recognize engineered *Archaeoglobus fulgidus* prolyl-tRNAs (Af-tRNA^{Pro}) with three different anticodons: CUA, AGGG, and CUAG. Several of these evolved PhProRSs show specificity toward a particular anticodon variant of Af-tRNA^{Pro}, whereas others are promiscuous. Further evolution of the Af-tRNA^{Pro} led to a variant exhibiting significantly improved amber suppression efficiency. Availability of a prolyl-tRNA/aaRS pair should enable site-specific incorporation of proline analogs and other N-modified UAAs into proteins in *E. coli*. The evolution of mutually orthogonal prolyl-tRNA/ProRS pairs demonstrates the plasticity of the tRNA-aaRS interface and should facilitate the incorporation of multiple, distinct UAAs into proteins.

anticodon recognition

The development of methods to encode unnatural amino acids (UAAs) with novel properties genetically in living cells has significantly enhanced our ability to manipulate protein structure and function (1). Central to this system is a unique tRNA/aminoacyl-tRNA synthetase (aaRS) pair, orthogonal to the host cell (i.e., not cross-reactive with endogenous host tRNAs and aaRSs), which is engineered to deliver the UAA of interest in response to a noncoding termination or four-base frameshift codon. A number of orthogonal tRNA/aaRS pairs have been generated (1–5) for use in *Escherichia coli*, among which two archaeobacteria-derived pairs (tyrosyl and pyrrolysyl) have been engineered successfully to encode a large number of UAAs with unique functionalities (1). Nonetheless, the generation of new orthogonal tRNA/aaRS pairs will further expand our ability to genetically encode UAAs with novel structural features. Moreover, mutually orthogonal pairs that suppress distinct nonsense or frameshift codons should facilitate the simultaneous incorporation of multiple, distinct UAAs in the same polypeptide (2, 6–8).

Among the 20 canonical amino acids, proline is the only one with a secondary amine. N-modified amino acids are known to impart unique properties to peptides, including distinctive conformational propensities to the peptide backbone, increased cellular permeability, and resistance to protease mediated degradation (9–13). The generation of an orthogonal suppressor prolyl-tRNA (tRNA^{Pro})/prolyl-tRNA synthase (ProRS) pair may facilitate the site-specific incorporation of UAAs with backbone modifications into proteins, which has been difficult to achieve using existing tRNA/aaRS pairs. To use a heterologous tRNA/aaRS pair to encode a noncanonical amino acid, the pair must be recoded to suppress a nonsense or frameshift codon, and the recoding necessitates mutating the anticodon of the corresponding

tRNA. The amber nonsense codon TAG often is used for this purpose because of its low frequency in *E. coli* and the observation that natural amber suppressors do not affect growth rates significantly. The other two nonsense codons, TAA and TGA, and several frameshift codons also have been used successfully (2, 3, 7, 14, 15). In the case of the tRNA^{Pro}/ProRS pair, the anticodon is a major identity element for tRNA recognition by its cognate synthetase (16–19). The consensus sequence of the proline anticodons (NGG) in tRNA^{Pro} interacts extensively with an array of conserved amino acid residues in the anticodon-binding pocket of ProRS, as evidenced in the tRNA^{Pro}/ProRS cocystal structure from *Thermus thermophilus* (18, 19). As a consequence, alteration of the anticodon negatively affects the tRNA^{Pro}-ProRS interaction, leading to low or undetectable activity in *E. coli*. Thus, to adapt this pair for the incorporation of UAAs in response to nonsense or frameshift codons, it is first necessary to engineer ProRS to accept variants of tRNA^{Pro} bearing nonnative anticodons.

Here we report the development of a heterologous archaeal tRNA^{Pro}/ProRS pair, *Archaeoglobus fulgidus* tRNA^{Pro} (Af-tRNA^{Pro})/*Pyrococcus horikoshii* ProRS (PhProRS), for UAA mutagenesis in *E. coli*. Modification of the anticodon-binding pocket of PhProRS reestablished its functional binding interaction with multiple anticodon-variants of tRNA^{Pro}. Additionally, we show that the extensive anticodon-ProRS interface can be manipulated to render these interactions specific, allowing the generation of three mutually non-cross-reactive tRNA^{Pro}/ProRS pairs suppressing distinct nonsense and frame-shift codons.

Results and Discussion

Identifying a Functional, Orthogonal tRNA^{Pro}/ProRS Pair. For the expression host *E. coli*, tRNA^{Pro}/ProRS pairs of archaeal origin are likely to be orthogonal because the substantial sequence divergence from their bacterial counterparts has been shown to result in low cross-reactivity (17). Moreover, genome sequences for a large number of archaeal species are now available. It has been shown previously that a heterologous combination of tRNA and aaRS can provide maximal activity in the expression host (3). Thus, six different ProRS and tRNA^{Pro} from different archaeobacteria were chosen to represent the broad natural sequence diversity (Fig. S1). The ProRS genes were inserted into the pBK plasmid (20) and expressed constitutively from the *glnS* promoter. The anticodons of the tRNAs were mutated to 5'-CUA (Fig. 1A) to allow amber suppression, and 5'-CCA was added to the 3'-terminus, if necessary. The tRNAs were incorporated

Author contributions: A.C. and P.G.S. designed research; A.C. and H.X. performed research; A.C. and H.X. contributed new reagents/analytic tools; A.C., H.X., and P.G.S. analyzed data; and A.C. and P.G.S. wrote the paper.

The authors declare no conflict of interest.

¹A.C. and H.X. contributed equally to this work.

²To whom correspondence should be addressed. E-mail: schultz@scripps.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1212454109/-DCSupplemental.

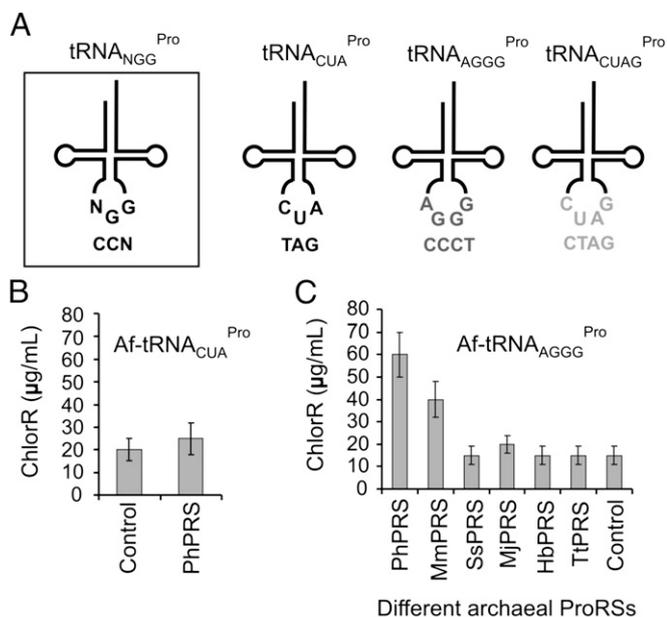


Fig. 1. Suppression efficiency of tRNA^{Pro}/ProRS pairs is affected by the anticodon. (A) Different anticodon variants of the tRNA^{Pro} used in this study. In the box at the far left is shown the wild-type tRNA^{Pro}_{NGG}, which has an anticodon sequence similar to tRNA^{Pro}_{AGGG} but not to the other two variants. (B) Prolyl-tRNA^{Pro}_{CUA} is not charged efficiently by wild-type ProRS. The efficiency of Af-tRNA^{Pro}_{CUA} in TAG suppression was measured as ChlorR using the pRepCM3b reporter plasmid with an empty pBK vector (control) or with a pBK vector expressing PhProRS. (C) Af-tRNA^{Pro}_{AGGG} is charged efficiently by wild-type ProRS. The efficiency of Af-tRNA^{Pro}_{AGGG} in suppressing CCCT was measured as ChlorR using the pRepCM3b reporter plasmid with an empty pBK vector (control) or with a pBK vector expressing six different wild-type ProRSs from the following archaeal species: *P. horikoshii*, *M. mazei*, *Sulfolobus solfataricus*, *Methanocaldococcus jannaschii*, *Halo bacterium* sp. (NRC-1), and *T. thermophilus*.

into the pRepCM3b plasmid (21) and expressed constitutively using the *lpp* promoter. A chloramphenicol acetyltransferase (CAT) gene encoded in pRepCM3b and containing a permissive amber mutation (Gln98TAG) was used to evaluate the suppression efficiency of tRNA^{Pro}_{CUA} in the presence and the absence of different ProRSs. When transformed with pRepCM3b, which lacks an amber suppressor tRNA, *E. coli* does not survive chloramphenicol concentrations greater than 5 μg/mL. An increase in chloramphenicol resistance (ChlorR) conferred by a suppressor tRNA^{Pro} alone or in the presence of a ProRS is a measure of the host cross-reactivity of the corresponding tRNA and the suppression activity of the pair, respectively. When all 36 combinations of these tRNA^{Pro}/ProRS pairs were tested, only one combination, Af-tRNA^{Pro}_{CUA}/PhProRS, showed a small enhancement in ChlorR over the background activity of the tRNA alone (Fig. 1B). Although this effect was weak, it was reproducible. The inability of these archaeal prolyl pairs to confer reasonable levels of amber suppression can be explained by the disruption of the tRNA^{Pro}-ProRS binding interaction as a result of mutations introduced in the anticodon.

Among the 64 triplet codons, 61 represent sense codons and are not suitable for UAA incorporation methodology unless engineered to be absent from the host genome. Quadruplet codons, on the other hand, are noncoding and have been used for UAA incorporation in a number of instances (2, 6, 14, 15). Therefore we exploited the sequence diversity of available quadruplet codons to probe the anticodon tolerance of the ProRSs, using Af-tRNA^{Pro} variants equipped with different four-base anticodons (Fig. 1A). To this end, we constructed an Af-tRNA^{Pro}

with a four-base anticodon (AGGG) that retains the conserved elements (NGG) of the wild-type tRNA^{Pro} (Fig. 1A). To evaluate its suppression activity, the in-frame TAG codon in the CAT reporter of pRepCM3b was replaced with the corresponding four-base codon CCCT. If a ProRS is able to charge tRNA^{Pro}_{AGGG}, the tRNA should suppress this CCCT codon in the CAT gene, resulting in the production of active CAT protein and enhanced ChlorR. The inability to do so would result in exclusive suppression of the CCCT codon by the host tRNA^{Pro}_{AGGG}, resulting in a frame shift and no enhancement in ChlorR. When this construct was tested with six different ProRSs, significant enhancement in ChlorR over the activity of the tRNA alone was observed with ProRS from *P. horikoshii* and *Methanosarcina mazei* (Fig. 1C). We also constructed two other four-base suppressor tRNAs that lack the consensus sequence of wild-type tRNA^{Pro}, Af-tRNA^{Pro}_{UCCU} (AGGA) and Af-tRNA^{Pro}_{CUAG} (CTAG), which had no observable suppression activity with PhProRS. These results strongly support the importance of anticodon recognition in tRNA^{Pro}/ProRS suppression activity, and identified the Af-tRNA^{Pro}/PhProRS pair as the most promising candidate for further development.

Evolution of PhProRS to Restore Recognition of Af-tRNA^{Pro}_{CUA}. Mutations in the anticodon-binding pocket of aaRSs previously have been shown to improve interactions with nonsense or frameshift suppressor tRNAs (22–24). These examples suggest that it may be possible to evolve the anticodon-binding pocket of PhProRS to recognize and charge Af-tRNA^{Pro}_{CUA} efficiently. The cocrystal structure of the tRNA^{Pro}/ProRS pair from *T. thermophilus*, which is homologous to the pair identified above, reveals the molecular basis of anticodon recognition by ProRS (18, 19). We found that the consensus sequence of the proline anticodon (GG) interacts extensively with several conserved amino acid residues in the anticodon-binding pocket of ProRS (Fig. 2A). Six of these residues (Tyr335, Glu338, Arg345, Glu347, Asp352, and Lys368) were randomized to all possible amino acids (NNK randomization) to generate a library (PhPac1) with a theoretical diversity of 10⁹, which was covered by >10¹⁰ individual clones (>99.9% coverage). To identify desired PhProRS variants from this library, a selection was devised based on the previously reported dual-reporter plasmid pREP, expressing the suppressor Af-tRNA^{Pro}_{CUA} (25). This vector harbors the CAT (Asp111TAG) gene and a T7 promoter-driven GFP as reporters. A T7 RNA polymerase gene containing two amber mutations at permissible sites also is encoded in this plasmid. Active aaRS variants able to charge the suppressor tRNA^{Pro}_{CUA} can be identified based on a selective growth advantage in the presence of chloramphenicol and/or expression of GFP.

The reporter plasmid (pREP-AfPwt) was cotransformed with the PhPac1 library and subjected to selection in the presence of increasing chloramphenicol concentrations. GFP expression was used as a secondary screen to avoid survivors cheating the chloramphenicol selection scheme. After three successive rounds of selection, several PhProRS variants (PhPRS_{CUA}) were identified conferring ChlorR up to 180 μg/mL. DNA sequence analysis of 96 such clones revealed several unique sequences (Fig. 2B and Fig. S2). Although these active variants represent considerable sequence diversity, some similarity could be observed at certain positions. Tyr335 was either conserved or replaced by Phe, His, or Asn, representing the natural diversity of the archaeal ProRS family. Glu347 was altered, preferably to a small amino acid (Gly, Ser, Thr, or Ala), and Asp352 often was changed to Gln or Glu. Interestingly, additional, serendipitous mutations were identified in some cases. Reversal of these mutations in PhPRS_{CUA}-h3 (Gly206Trp and Glu98Gly) to the original sequence significantly reduced activity, indicating the beneficial role of the mutations (Fig. S3). In the crystal structure of *Thermus thermophilus* ProRS, these mutations localize near

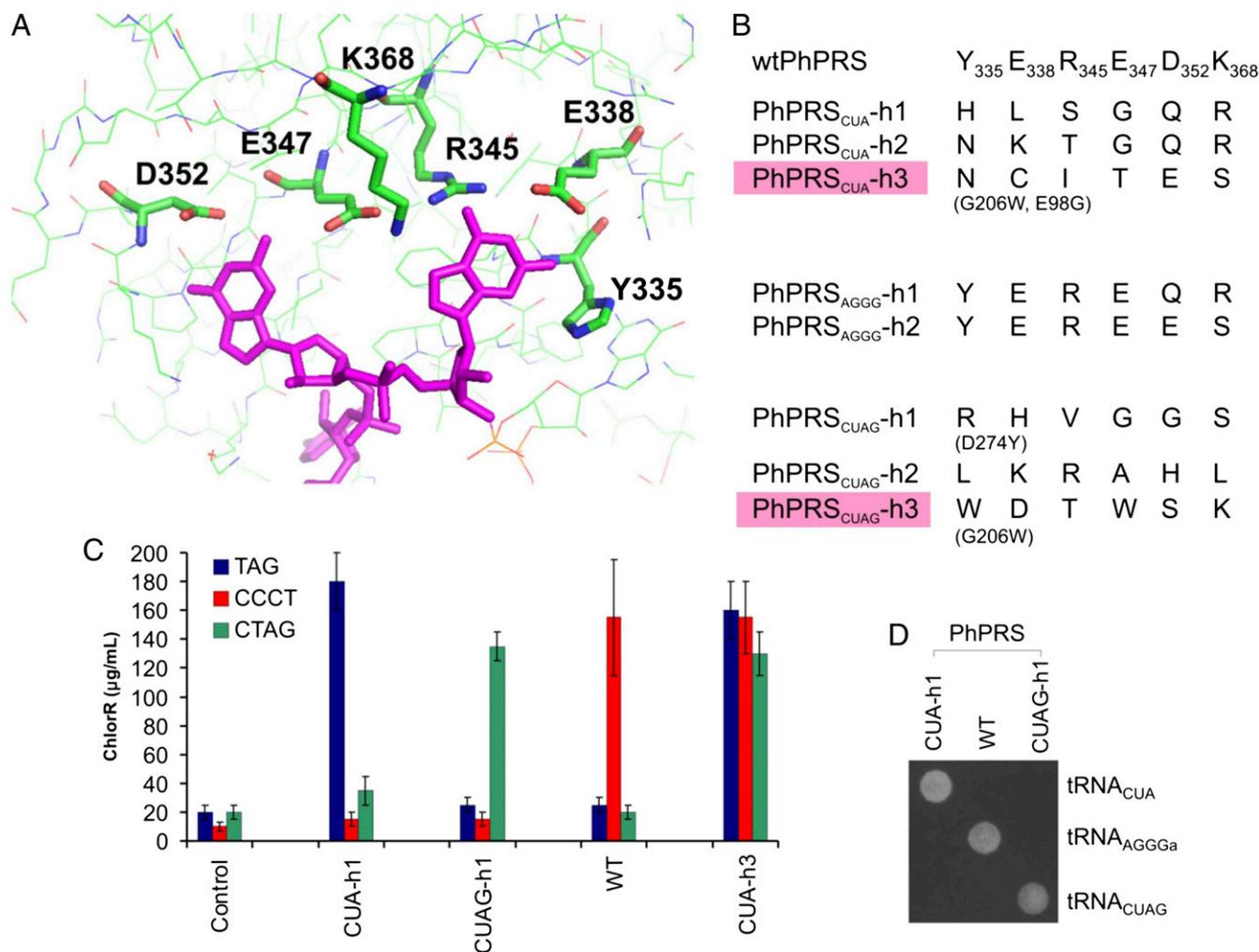


Fig. 2. Evolution of the anticodon-binding pocket of PhProRS to charge different Af-tRNA^{Pro} anticodon variants. (A) Anticodon-binding site of ProRS in the crystal structure from *T. thermophilus*. The consensus sequence of the proline anticodon (GG), shown in magenta, interacts strongly with the amino acid residues in the binding pocket. Highlighted amino acid residues, involved in the recognition of this GG motif, were randomized, and the corresponding residues in PhProRS are shown. (B) Sequences of evolved PhProRS mutants that efficiently charge different anticodon variants of Af-tRNA^{Pro}. PhProRS variants highlighted in magenta are anticodon permissive and efficiently charge Af-tRNA^{Pro} bearing any of the three anticodons, CUA, AGGG, or CUAG. (C) PhProRS_{CUA}-h1/Af-tRNA_{CUA}^{Pro}, PhProRS_{CUAG}-h1/Af-tRNA_{CUAG}^{Pro}, and wtPhProRS/Af-tRNA_{AGGGa}^{Pro} are mutually orthogonal tRNA/aaRS pairs suppressing TAG, CTAG, and CCCT, respectively: Af-tRNA_{CUA}^{Pro}, Af-tRNA_{AGGGa}^{Pro}, and Af-tRNA_{CUAG}^{Pro} in the pREP, pREP4b-ccct, and pREP4b-ctag reporter plasmids, respectively, were used to determine the TAG, CCCT, and CTAG suppression activity measured as ChlorR. Using these reporter plasmids the TAG (blue), CCCT (red), and CTAG (green) suppression activities of two anticodon-specific PhProRS mutants (CUA-h1, CUAG-h1), wild-type PhProRS, and one anticodon-permissive PhProRS mutant (CUA-h3) were measured. The control experiment measured ChlorR elicited by the reporter plasmid in the absence of any PhProRS variants (empty pBK vector) (D) Specific PhProRS variants (identified at the top) allow *E. coli* to survive in the presence of 120 µg/mL chloramphenicol only when paired with the Af-tRNA^{Pro} bearing the cognate anticodon (identified at the right).

the active site and may engage in beneficial interactions with the tRNA acceptor stem (Fig. S3). In an attempt to improve the suppression efficiency of this pair further, a second library was generated (PhPac2). Results from the last selection were used as a guide to restrict the randomization of the previously investigated amino acid residues, and two additional positions (Phe334 and Val359) in the PhProRS anticodon-binding pocket were fully randomized (Fig. S4A and B). Although screening of this library revealed many active variants (Fig. S4C), none exhibited significantly enhanced activity over their previously identified counterparts.

Evaluation of Anticodon Tolerance of the Evolved PhProRS_{CUA} Variants. We next determined whether the PhProRS_{CUA} variants, selected for efficient charging of tRNA_{CUA}^{Pro}, are specific in their interactions for the CUA anticodon. Reporter plasmids (pRepCM3b)

were constructed expressing Af-tRNA^{Pro} with three different four-base anticodon (AGGG, UCCU, and CUAG) substitutions and a CAT gene carrying the corresponding in-frame four-base codons (CCCT, AGGA, and CTAG). These four-base codons were chosen because of their structural diversity and superior suppression efficiency, as demonstrated previously (14, 15). Interestingly, most of these evolved PhProRS_{CUA} exhibited no AGGA or CCCT suppression activity over background. A single exception to this trend was PhProRS_{CUA}-h3, which robustly suppressed CCCT; it also weakly suppressed AGGA. The presence of the previously described additional, favorable mutations in this variant, presumably interacting with the acceptor stem of the tRNA (Fig. S3A), may explain its atypical tolerance for other anticodons. Indeed, when the beneficial mutations in PhProRS_{CUA}-h3 were reverted back to the wild-type sequence, CCCT suppression activity was reduced to a negligible level

(Fig. S3B). Many of the PhPRS_{CUA} variants were able to charge Af-tRNA_{CUAG}^{Pro} efficiently. This result is reasonable, because the Af-tRNA_{CUAG}^{Pro} is identical to AfP-tRNA_{CUA}^{Pro} except for the addition of a single nucleotide in the anticodon loop. The high degree of selectivity exhibited by PhPRS_{CUA} variants toward the Af-tRNA^{Pro} with a CUA anticodon suggests that it might be possible to evolve PhPRS mutants that selectively recognize other anticodons. Such mutants would allow the creation of mutually orthogonal tRNA^{Pro}/ProRS pairs able to suppress distinct codons.

Evolution of Efficient PhProRS/Af-tRNA^{Pro} Pairs for CCCT Suppression.

Although the wild-type PhProRS interacts poorly with Af-tRNA_{CUA}^{Pro}, it efficiently charges Af-tRNA_{AGGG}^{Pro} because of its structural similarity to the native tRNA_{NGG}^{Pro}. Thus, wild-type PhPRS (wtPhPRS)/Af-tRNA_{AGGG}^{Pro} and PhPRS_{CUA}/Af-tRNA_{CUA}^{Pro} represent mutually orthogonal pairs, suppressing CCCT and TAG, respectively. However, the superior suppression efficiency of TAG (up to 180 μg/mL ChlorR) relative to CCCT (60 μg/mL ChlorR) suggests that the latter can be improved. Mutation of the nucleotides adjacent to the anticodon has been reported to improve the activity of suppressor tRNAs (2, 3, 26). Therefore, we randomized all four nucleotides in the anticodon loop of Af-tRNA_{AGGG}^{Pro} flanking the anticodon to generate a small library (Fig. S5A), which was selected for the highest level of CCCT suppression in the presence of wild-type PhProRS (wtPhProRS). A single-nucleotide variant of the original tRNA (G39A; Af-tRNA_{AGGGa}^{Pro}) was identified that significantly enhanced CCCT suppression activity in the presence of wtPhPRS, while retaining non-cross-reactivity against the PhPRS_{CUA} variants (Fig. S5). The beneficial nature of the G39A mutation may be attributed to the disruption of the guanine-rich sequence of the anticodon loop in Af-tRNA_{AGGG}^{Pro}, because such sequences tend to form potentially unfavorable secondary structures.

Next, we attempted to improve the efficiency of CCCT suppression further by mutagenizing the anticodon-binding domain of PhProRS. A dual-reporter plasmid for CCCT suppression activity (pREP4b-ccct) was generated by replacing the in-frame TAG codons in CAT and T7RP reporters with CCCT in the pREP plasmid, followed by the introduction of Af-tRNA_{AGGG}^{Pro}. This vector was used to isolate PhProRS variants from the PhPac1 anticodon-binding site library that efficiently charge Af-tRNA_{AGGG}^{Pro}. After three rounds of selection, the most promising clones were analyzed by DNA sequencing to reveal a high abundance of the wtPhPRS. The previously described anticodon-permissive PhPRS_{CUA}-h3 mutant and two other unique variants were also identified (Fig. S6). When compared for CCCT suppression activity, PhPRS_{AGGG}-h1 showed minor improvement over wtPhPRS, although at the expense of enhanced cross-reactivity with Af-tRNA_{CUA}^{Pro} (Fig. S6). The inability to identify a significantly improved PhProRS variant for CCCT suppression indicates preexisting, favorable interactions between wtPhPRS and the anticodon of Af-tRNA_{AGGG}^{Pro}. Also, these results identify wtPhPRS/Af-tRNA_{AGGGa}^{Pro} as an efficient CCCT-suppressing pair that exhibits minimal cross-reactivity with the PhPRS_{CUA}-h1/Af-tRNA_{CUA}^{Pro} nonsense suppressor pair.

Evolution of Efficient PhPRS/Af-tRNA^{Pro} Pairs for CTAG Suppression.

The significant structural differences between the anticodons of Af-tRNA_{AGGG}^{Pro} and Af-tRNA_{CUA}^{Pro} allowed us to evolve PhPRS variants able to differentiate the two effectively. Although discrimination between the nearly identical Af-tRNA_{CUAG}^{Pro} and Af-tRNA_{CUA}^{Pro} appears to be challenging, the selectivity of certain evolved PhPRS_{CUA} variants (e.g., PhPRS_{CUA}-h1) (Fig. S2) for the latter tRNA suggests that it should be possible to evolve analogs (PhPRS_{CUAG}) specific for tRNA_{CUAG}^{Pro}. Moreover, some of these PhPRS_{CUAG} variants may not cross react with the structurally distant CCCT-suppressor tRNA, resulting in

a CTAG-suppressor ProRS/tRNA^{Pro} pair mutually orthogonal to both its TAG- and CCCT-suppressing counterparts.

A dual-reporter plasmid for CTAG-suppression activity (pREP4b-ctag) was constructed, similarly to pREP4b-ccct. Using this vector, the PhPac1 library was subjected to two rounds of selection to isolate effective CTAG-suppressing variants. To eliminate members within this surviving pool that can charge Af-tRNA_{CUA}^{Pro}, a negative round of selection was performed based on the expression of the toxic barnase gene containing permissible TAG mutations (20). DNA-sequencing analysis of surviving clones identified several unique variants, some of which were able to charge Af-tRNA_{CUAG}^{Pro} efficiently but not Af-tRNA_{CUA}^{Pro} or Af-tRNA_{AGGG}^{Pro} (Fig. S7). An anticodon-permissive variant, PhPRS_{CUAG}-h4, also was identified that, like the previously described PhPRS_{CUA}-h3, contains the serendipitous G206W mutation. The major differences between the PhPRS_{CUA} and PhPRS_{CUAG} variants are as follows: Tyr335 substitutions in PhPRS_{CUA} variants reflect the conservation of natural diversity (Tyr/Asn/His/Phe), whereas mutation to hydrophobic residues is found in PhPRS_{CUAG}; Arg345 is conserved in PhPRS_{CUAG} (except for variants carrying additional mutations) but not in PhPRS_{CUA} variants.

To demonstrate the mutual orthogonality of these three prolyl suppressor pairs, the following combinations were selected: PhPRS_{CUA}-h1/Af-tRNA_{CUA}^{Pro}, wtPhPRS/Af-tRNA_{AGGGa}^{Pro}, and PhPRS_{CUAG}-h1/Af-tRNA_{CUAG}^{Pro}. When their suppression efficiency was compared in a ChlorR assay, each tRNA variant exhibited high levels of activity only when paired with its cognate synthetase (Fig. 2 C and D). The evolution of PhProRS variants able to distinguish even the subtle variations in the anticodon loop (e.g., TAG- and CTAG-suppressor tRNAs) suggests that it should be possible to create other mutually orthogonal tRNA^{Pro}/ProRS pairs based solely on different anticodons. Furthermore, the modular nature of ProRS, composed of distinct charging and anticodon-binding domains, should allow the facile combination of variant charging domains (specific for UAAs) with the evolved anticodon-binding domains described here. It is likely that the prolyl pairs described here also are orthogonal to the other tRNA/aaRS pairs adapted in *E. coli*, thus facilitating their co-use in incorporating multiple, distinct UAAs into the same protein.

Evolution of a Highly Efficient TAG Suppressor Af-tRNA^{Pro}. To evaluate the efficiency of the PhPRS_{CUA}-h1/Af-tRNA_{CUA}^{Pro} nonsense suppressor pair for the recombinant expression of mutant proteins, it was incorporated in the pEvol plasmid and was used to express GFP (Tyr151TAG) driven by a T7 promoter (27). Although we were able to purify and characterize the expected Tyr151Pro mutant by mass spectrometry, the suppression efficiency was <10% of that of wild-type GFP (wtGFP) (isolated yield 1 mg/L compared with 11 mg/L for wtGFP) (Fig. 3C). To improve the amber suppression activity further, we mutagenized the suppressor tRNA acceptor stem. Mutations in the acceptor stem have been used previously with considerable success to improve tRNA activity (2–4). Therefore, we randomized the first five base pairs in the acceptor stem and C34 and G39 in the anticodon loop of Af-tRNA_{CUA}^{Pro} to all possible combinations (Fig. 3A). The theoretical diversity of this tRNA library (~1.7 × 10⁷) was covered by >10⁹ individual clones. This library first was subjected to two rounds of selection to enrich efficient TAG-suppressor tRNA variants in the presence of the cognate PhPRS_{CUA}-h1 and CAT (Gln98TAG) based on chloramphenicol survival. To eliminate surviving variants that cross react with *E. coli* aaRSs, a negative selection was performed in the absence of the cognate PhPRS, using a toxic barnase gene containing two permissible TAG mutations. From the surviving library members, one Af-tRNA_{CUA}^{Pro} variant was identified that exhibited significantly enhanced suppression activity while remaining non-cross-reactive to *E. coli* (Fig. 3 A and B). When introduced into

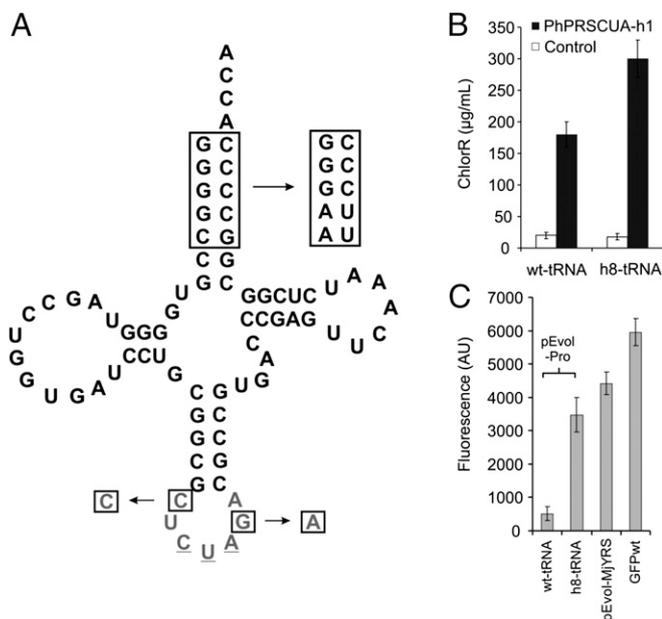


Fig. 3. Evolution of a highly efficient amber suppressor, Af-tRNA_{CUA}^{Pro}. (A) The highlighted segments of Af-tRNA_{CUA}^{Pro} were randomized. A highly efficient variant incorporating the indicated mutations, Af-tRNA_{CUA}^{Pro}-h8, was identified. (B) Suppression activity of the wild type and h8 Af-tRNA_{CUA}^{Pro}, measured as ChlorR, in the presence (black bar) and absence (white bar) of the cognate PhPRSCUA-h1. The h8-tRNA variant has enhanced activity relative to wild-type tRNA and remains orthogonal to *E. coli*. (C) Expression levels of GFP-Tyr151TAG using pEvol-Pro encoding the original (wild type) or the evolved h8-tRNA. Expression levels of the same GFP-Tyr151TAG construct using pEvol-MjYRS (wild type) and of the corresponding wild-type GFP construct (GFPwt), not carrying a nonsense codon, are provided for comparison.

pEvol in conjunction with PhPRSCUA-h1, this tRNA showed a high level of amber suppression (~60% of wtGFP) (Fig. 3C). The suppression efficiency of this system was assessed further using a C-terminal hexahistidine-tagged *Pseudomonas putida* ketosteroid isomerase (KSI) amber mutant (KSI-His78TAG); the corresponding expression level was compared with that of wild-type KSI. A pET28 plasmid expressing KSI (His78TAG) under the T7 promoter was cotransformed with the aforementioned pEvol-Pro plasmid harboring PhPRSCUA-h1 and the evolved Af-tRNA_{CUA}^{Pro} variant (h8) into the BL21 (De3) strain, and the expression level was assessed in LB medium. A high level of expression (15 mg/L) for KSI (His78TAG) (Fig. S8), ~50% of the yield obtained with wtKSI, was observed.

Conclusion

We have evolved three efficient ProRS/tRNA^{Pro} pairs that suppress three distinct codons in *E. coli* from a single ancestral pair. Most of the orthogonal tRNA/aaRS pairs adopted in *E. coli* for UAA mutagenesis do not exhibit significant anticodon recognition (2–5, 22, 28). In contrast, the proline system described here represents a family of tRNA/aaRS pairs in which strong recognition of the anticodon causes severely attenuated activity when the anticodon is altered. Here we demonstrate that for such systems it is possible to evolve the anticodon-binding domain of the synthetase to recognize tRNAs carrying nonnative anticodons and that specific aaRS partners can be generated for multiple anticodon variants of the same tRNA by taking advantage of the extensive anticodon–aaRS interface. Interestingly, several evolved active PhProRS variants carried additional, advantageous mutations. The Gly206Trp mutation located near the active site in two different variants renders them highly

anticodon permissive and also accounts for most of their activity. In the absence of structural information, the molecular basis of the altered interaction patterns between the evolved tRNA^{Pro} and ProRS variants remains unclear. However, the results described here attest to the plasticity of the tRNA–aaRS interface and its potential utility for evolving variants with unique properties. Crystal structures of ProRS from *T. thermophilus* (18, 19) and *Meihanothermobacter thermautotrophicus* (29) bound to prolyl-adenylate analogs are available, providing excellent structural bases for the evolution of the analogous PhProRS to encode genetically novel proline analogs and other N-modified UAAs. In addition, the development of mutually orthogonal prolyl pairs further expands our ability to incorporate multiple, distinct UAAs into the same protein, eventually leading to genetically encoded unnatural biopolymers.

Materials and Methods

Genomic DNA from different archaeal species was obtained from American Type Culture Collection. PCR reactions were carried out with Platinum Pfx polymerase (Invitrogen), following the manufacturer's protocol. DNA oligomers for PCR reactions were purchased from Integrated DNA Technologies. Restriction enzymes and T4 DNA ligase from New England BioLabs were used. Plasmid DNA was purified using Qiagen Spin Miniprep kits. Macherey-Nagel NucleoSpin columns were used to purify DNA following digestion or gel electrophoresis. Unless otherwise mentioned, phage-resistant DH10b cells (Invitrogen) were used for all *in vivo* experiments. BL21 (De3) cells were used for GFP overexpression from the pET101 plasmid (Invitrogen). Ni-NTA-agarose resin from Qiagen was used to purify C-terminally His₆-tagged GFP. Bacterial cells were grown on LB/agar plates or 2× YT (Yeast extract Tryptone) liquid medium with the following antibiotic concentrations unless otherwise mentioned: 50 µg/mL kanamycin, 10 µg/mL tetracycline, 50 µg/mL chloramphenicol, and 100 µg/mL ampicillin. Media components and different antibiotics were obtained from Difco and Sigma-Aldrich, respectively.

Cloning of tRNA^{Pro}-ProRS Pairs. Different ProRSs were introduced to the pBK plasmid using polymerase incomplete primer extension (PIPE) cloning (30). The vector backbone was amplified using primers pBK-PIPE-F and pBK-PIPE-R. ProRSs from different archaeal species were amplified from the genomic DNA. Inserts for the tRNA expression cassette were formed by overlap extension PCR of the two terminal fragments and were inserted into pRepCM3b using KpnI and EagI restriction sites. The 5' and 3' terminal fragments were generated using the primers pRepC_Kpn-F+tRNA-R and pRepC-Eag-R+tRNA-F, respectively. For cloning tRNAs into pREP, a similar strategy was used with different restriction sites (ScaI and EagI) and the vector-specific primers (pRep_Sca-F and pRep_Eag-R). The pRepCM3b-Af-tRNA^{Pro} plasmid with CTAG, AGGA, and CCCT suppressor/reporter was generated by site-directed mutagenesis of the anticodon and the TAG codon of the CAT using the QuikChange Multi Site-Directed Mutagenesis Kit (Agilent Technologies) with primers AfP-nnnn-R and CATcm-nnnn-F (nnnn represents ctag, agga, or ccct), following the manufacturer's directions.

Determination of Suppression Efficiency. An overnight culture of DH10b cells cotransformed with the aaRS and the tRNA-reporter plasmids was diluted to an OD₆₀₀ of 0.01, and a 3-µL aliquot was spotted on an LB/agar plate containing 30 µg/mL kanamycin, 5 µg/mL tetracycline, and a varying concentration of chloramphenicol. After 36–48 h of incubation at 37 °C, growth patterns were analyzed.

Construction of pREP Reporter Plasmids. To construct pREP4b (ctag or ccct) plasmid from pREP, the TAG codons in T7RP were mutated to the respective four-base codons by generating the corresponding insert using overlap extension PCR and inserting it between the KpnI and NsiI restriction sites. Terminal primers pRep_T7m-tF and -tR and mutagenic primers pRep-T7mF1 and R1-nnnn and pRep-T7mF2 and R1-nnnn (nnnn represents ctag or ccct) were used to amplify the individual fragments. The TAG codon in the CAT gene was mutated similarly using the BsaAI and ScaI restriction sites and the primers CAT-tF, CAT-nnnn-F (nnnn represents ctag or ccct) and CAT-nnnn-R, CAT-tR.

Construction of Libraries. To construct the anticodon binding site library of PhProRS, the relevant region was amplified with mutagenic primers as two fragments using pBK-PhProRS as the template. These two fragments were joined together by overlap extension PCR, amplified using pBK-F and ProRS-

Ph-R, and were inserted into pBK-PhPRS using the BamHI and NcoI sites. For PhPacb1, pBK-F/PhPacLb-R and PhPRS-R/PhPacLb-F were used to amplify the 5' and the 3' fragments, respectively. For PhPacLb2, PhPacLb-R and PhPacLb-F were replaced by PhPfacLb-R and PhPfacb-F, respectively. The tRNA library was constructed into pBKCM10b (31) using overlap extension PCR between the NcoI/XhoI sites as described previously. The following primers with randomized positions were used to construct the library: AfP_tRNALb-F and AfP_tRNALb-R.

General Selection Strategies. For the selection of aaRS variants, the library was cotransformed with the appropriate pREP reporter plasmid and plated on different concentrations of chloramphenicol. Cells were harvested from the plates with the highest chloramphenicol concentration sustaining robust bacterial growth. The library plasmid was isolated and purified from these cells and subjected to further rounds of selection as described above. After the desired rounds of selection, individual colonies able to grow at higher chloramphenicol concentrations were picked, and the associated library members were characterized by DNA sequencing. For PhPRS_{CUAG} selection, a negative-selection step was included after two rounds of positive selection, in which the surviving library members were cotransformed with a pNeg plasmid carrying an Af-tRNA_{CUA}^{Pro} expression cassette and were plated on selective solid medium containing 0.02% L-arabinose. After the negative selection, the surviving pool was subjected to one more round of positive selection before individual clones were characterized. Active but orthogonal tRNA variants were selected as described previously (31).

Construction of pEvol-Pro and Expression of GFP. To construct pEvol-Pro, the tRNA cassette under the *proK* promoter first was constructed by overlap extension PCR of two fragments and was amplified using the following primer combinations using a variant of the reported pEvol-MjY plasmid as the template: pEvol-Not-F/pEvol-AfP-wttRNA-R and pEvol-Nco-R/pEvol-AfP-wttRNA-F. To this plasmid, two copies of PhPRS_{CUA}-h1 were integrated under *araBAD* and *glnS* promoters using the SbfI and the NotI sites, respectively. To express the GFP (Tyr151TAG) mutant using the prolyl-suppressor pair, the pEvol-Pro plasmid was cotransformed with pET101-GFP (Tyr151TAG) into the BL21(D_{e3}) strain. The resultant host was grown in 2× YT medium at 37 °C to an OD₆₀₀ of 0.6, at which point the expressions of GFP and the aaRS were coinduced using 1 mM isopropylthio-β-galactoside and 0.02% arabinose, respectively. After overnight expression at 30 °C, the cells were harvested and lysed with BugBuster (EMD Millipore), and the protein was purified using Ni-NTA agarose resin. Alternatively, to compare expression levels, the overnight culture was resuspended in PBS, and the fluorescence (excitation, 395 nm; emission, 510 nm) and OD₆₀₀ were measured using a SpectraMax 96-well plate reader (Molecular Devices). Normalized fluorescence (fluorescence/OD₆₀₀) was used for comparison.

ACKNOWLEDGMENTS. We thank Virginia Seely for her assistance in manuscript preparation. This work was funded by Grant DE-FG03-00ER46051 from the Division of Materials Sciences, Department of Energy (to P.G.S.). This is paper no. 21818 from The Scripps Research Institute.

- Liu CC, Schultz PG (2010) Adding new chemistries to the genetic code. *Annu Rev Biochem* 79:413–444.
- Anderson JC, et al. (2004) An expanded genetic code with a functional quadruplet codon. *Proc Natl Acad Sci USA* 101:7566–7571.
- Anderson JC, Schultz PG (2003) Adaptation of an orthogonal archaeal leucyl-tRNA and synthetase pair for four-base, amber, and opal suppression. *Biochemistry* 42: 9598–9608.
- Santoro SW, Anderson JC, Lakshman V, Schultz PG (2003) An archaeobacteria-derived glutamyl-tRNA synthetase and tRNA pair for unnatural amino acid mutagenesis of proteins in *Escherichia coli*. *Nucleic Acids Res* 31:6700–6709.
- Hughes RA, Ellington AD (2010) Rational design of an orthogonal tryptophanyl nonsense suppressor tRNA. *Nucleic Acids Res* 38:6813–6830.
- Neumann H, Wang K, Davis L, Garcia-Alai M, Chin JW (2010) Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature* 464: 441–444.
- Wan W, et al. (2010) A facile system for genetic incorporation of two different noncanonical amino acids into one protein in *Escherichia coli*. *Angew Chem Int Ed Engl* 49:3211–3214.
- Hoesl MG, Budisa N (2011) In vivo incorporation of multiple noncanonical amino acids into proteins. *Angew Chem Int Ed Engl* 50:2896–2902.
- Chatterjee J, Gilon C, Hoffman A, Kessler H (2008) N-methylation of peptides: A new perspective in medicinal chemistry. *Acc Chem Res* 41:1331–1342.
- Hanessian S, Auzzas L (2008) The practice of ring constraint in peptidomimetics using bicyclic and polycyclic amino acids. *Acc Chem Res* 41:1241–1251.
- Haviv F, et al. (1993) Effect of N-methyl substitution of the peptide bonds in luteinizing hormone-releasing hormone agonists. *J Med Chem* 36:363–369.
- Mazur RH, et al. (1980) Bradykinin analogues containing N alpha-methyl amino acids. *J Med Chem* 23:758–763.
- Tal-Gan Y, Freeman NS, Klein S, Levitzki A, Gilon C (2010) Synthesis and structure-activity relationship studies of peptidomimetic PKB/Akt inhibitors: The significance of backbone interactions. *Bioorg Med Chem* 18:2976–2985.
- Anderson JC, Magliery TJ, Schultz PG (2002) Exploring the limits of codon and anticodon size. *Chem Biol* 9:237–244.
- Magliery TJ, Anderson JC, Schultz PG (2001) Expanding the genetic code: Selection of efficient suppressors of four-base codons and identification of “shifty” four-base codons with a library approach in *Escherichia coli*. *J Mol Biol* 307:755–769.
- Yokozawa J, Okamoto K, Kawarabayashi Y, Kuno A, Hasegawa T (2003) Molecular recognition of proline tRNA by prolyl-tRNA synthetase from hyperthermophilic archaeon, *Aeropyrum pernix* K1. *Nucleic Acids Res Suppl* (3):247–248.
- Burke B, Lipman RS, Shiba K, Musier-Forsyth K, Hou YM (2001) Divergent adaptation of tRNA recognition by *Methanococcus jannaschii* prolyl-tRNA synthetase. *J Biol Chem* 276:20286–20291.
- Yaremchuk A, Cusack S, Tukalo M (2000) Crystal structure of a eukaryote/archaeon-like prolyl-tRNA synthetase and its complex with tRNA^{Pro}(CGG). *EMBO J* 19: 4745–4758.
- Cusack S, Yaremchuk A, Krikliiviy I, Tukalo M (1998) tRNA(Pro) anticodon recognition by *Thermus thermophilus* prolyl-tRNA synthetase. *Structure* 6:101–108.
- Wang L, Brock A, Herberich B, Schultz PG (2001) Expanding the genetic code of *Escherichia coli*. *Science* 292:498–500.
- Melançon CE, 3rd, Schultz PG (2009) One plasmid selection system for the rapid evolution of aminoacyl-tRNA synthetases. *Bioorg Med Chem Lett* 19:3845–3847.
- Kobayashi T, et al. (2003) Structural basis for orthogonal tRNA specificities of tyrosyl-tRNA synthetases for genetic code expansion. *Nat Struct Biol* 10:425–432.
- Neumann H, Slusarczyk AL, Chin JW (2010) De novo generation of mutually orthogonal aminoacyl-tRNA synthetase/tRNA pairs. *J Am Chem Soc* 132:2142–2144.
- Liu DR, Magliery TJ, Pastrnak M, Schultz PG (1997) Engineering a tRNA and aminoacyl-tRNA synthetase for the site-specific incorporation of unnatural amino acids into proteins in vivo. *Proc Natl Acad Sci USA* 94:10092–10097.
- Santoro SW, Wang L, Herberich B, King DS, Schultz PG (2002) An efficient system for the evolution of aminoacyl-tRNA synthetase specificity. *Nat Biotechnol* 20:1044–1048.
- Wang L, Schultz PG (2001) A general approach for the generation of orthogonal tRNAs. *Chem Biol* 8:883–890.
- Young TS, Ahmad I, Yin JA, Schultz PG (2010) An enhanced system for unnatural amino acid mutagenesis in *E. coli*. *J Mol Biol* 395:361–374.
- Ambrogely A, et al. (2007) Pyrrolysine is not hardwired for cotranslational insertion at UAG codons. *Proc Natl Acad Sci USA* 104:3141–3146.
- Kamtekar S, et al. (2003) The structural basis of cysteine aminoacylation of tRNA^{Pro} by prolyl-tRNA synthetases. *Proc Natl Acad Sci USA* 100:1673–1678.
- Klock HE, Lesley SA (2009) The Polymerase Incomplete Primer Extension (PIPE) method applied to high-throughput cloning and site-directed mutagenesis. *Methods Mol Biol* 498:91–103.
- Guo J, Melançon CE, 3rd, Lee HS, Groff D, Schultz PG (2009) Evolution of amber suppressor tRNAs for efficient bacterial production of proteins containing nonnatural amino acids. *Angew Chem Int Ed Engl* 48:9148–9151.