Pyrrolysine is not hardwired for cotranslational insertion at UAG codons

Alexandre Ambrogelly*, Sarath Gundllapalli*, Stephanie Herring*, Carla Polycarpo*†, Carina Frauer*‡, and Dieter Söll*§¶

Departments of *Molecular Biophysics and Biochemistry and §Chemistry, Yale University, New Haven, CT 06520-8114

Contributed by Dieter Söll, December 28, 2006 (sent for review December 23, 2006)

Pyrrolysine (Pyl), the 22nd naturally encoded amino acid, gets acylated to its distinctive UAG suppressor tRNA^{Pyl} by the cognate pyrrolysyl-tRNA synthetase (PylRS). Here we determine the RNA elements required for recognition and aminoacylation of tRNA^{Pyl} *in vivo* by using the Pyl analog *N*-ε-cyclopentyloxycarbonyl-L-lysine. Forty-two *Methanosarcina barkeri* tRNA^{Pyl} variants were tested in *Escherichia coli* for suppression of the *lac* amber A24 mutation; then relevant tRNA^{Pyl} mutants were selected to determine *in vivo* binding to *M. barkeri* PylRS in a yeast three-hybrid system and to measure *in vitro* tRNA^{Pyl} aminoacylation. tRNA^{Pyl} identity elements include the discriminator base, the first base pair of the acceptor stem, the T-stem base pair G51:C63, and the anticodon flanking nucleotides U33 and A37. Transplantation of the tRNA^{Pyl} identity elements into the mitochondrial bovine tRNA^{Ser} scaffold yielded chimeric tRNAs active both *in vitro* and *in vivo*. Because the anticodon is not important for PylRS recognition, a tRNA^{Pyl} variant could be constructed that efficiently suppressed the *lac* opal U4 mutation in *E. coli*. These data suggest that tRNA^{Pyl} variants may decode numerous codons and that tRNA^{Pyl}:PylRS is a fine orthogonal tRNA:synthetase pair that facilitated the late addition of Pyl to the genetic code.

orthogonal tRNA | suppression | tRNA identity | pyrrolysyl-tRNA synthetase | aminoacyl-tRNA synthetase

Incorporation of noncanonical amino acids into proteins is an exciting and active research field. To date, >30 unnatural amino acids have been placed into proteins with high fidelity mostly directed by the amber codon UAG (1, 2). The other two termination codons as well as enlarged codons (with 4–6 bases) have also been used (e.g., refs. 3 and 4). The key step in this process is the introduction of an orthogonal tRNA:aminoacyl-tRNA synthetase pair into the host protein synthesizing system. Such an orthogonal tRNA should not be recognized by any endogenous aminoacyl-tRNA synthetase, whereas the orthogonal synthetase should acylate solely the orthogonal tRNA with the unusual amino acid.

Less attention was devoted to incorporation of the noncanonical amino acids selenocysteine (Sec) and pyrrolysine (Pyl) that arose from natural expansion of the genetic code (5). Sec, the 21st cotranslationally inserted amino acid (6), is not suitable, because many organisms need this amino acid for viability and because its UGA-directed insertion requires additional RNA and protein components. Pyl, the 22nd cotranslationally inserted amino acid, appears more suited for this purpose, because it is restricted to a small number of organisms, where it accomplishes a special function (7). The *Methanosarcinaceae* contain a devoted UAG-recognizing suppressor tRNA^{Pyl} (8) and a pyrrolysyl-tRNA synthetase (PylRS) dedicated to forming Pyl-tRNA^{Pyl} (9, 10). Initial studies indicated that Lys-tRNA^{Pyl} can be recognized by bacterial EF-Tu (11) and that in *Escherichia coli* tRNA^{Pyl} acts like an amber suppressor (10, 12).

A thorough investigation in *E. coli* of archaeal tRNA^{Pyl} should uncover the structural determinants that may make tRNA^{Pyl} and PylRS an ideal orthogonal pair when used in bacterial protein synthesis. Here we investigate the interaction of *Methanosarcina barkeri* PylRS and *M. barkeri* tRNA^{Pyl} and explore the fitness and coding response of this orthogonal tRNA for translation in *E. coli*.

Results

Nucleotides that Determine Fitness of tRNA^{Pyl} for Translation in *E. coli*. To screen a large number of *M. barkeri* Fusaro tRNA^{Pyl} variants generated by mutagenesis, we made use of this tRNA's ability to suppress the *lac* amber mutation A24 in a *lacI–lacZ* fusion system (12, 13). Therefore, we transformed *E. coli* strain XAC/A24 with plasmid-borne copies of *M. barkeri pylS* and 42 mutant *pylT* genes and grew the transformants in the presence of the Pyl analog *N*-ε-cyclopentyloxycarbonyl-L-lysine (Cyc), because Pyl is not commercially available (12). Suppression was quantitated by measuring β-galactosidase activity (Table 1). The mutations covered all regions of the tRNA^{Pyl}; however, the anticodon was not altered because the suppression assay depended on the integrity of amber codon recognition.

Discriminator Base and Acceptor Helix. Systematic mutation of the nucleotides in the acceptor stem revealed that the discriminator base G73 and the first base pair of the acceptor stem are major tRNA^{Pyl} identity elements. The G73A and G73U mutations decreased suppression efficiency markedly. The base pair G1:C72, conserved in all known tRNA^{Pyl} species, could be flipped with some loss of activity, yet replacement with a weaker pair (G1:U72) or loss of the base pair (A1:C72) resulted in severe loss of suppression efficiency. These results suggest that the primary role of the first base pair is to ensure the proper productive placement of recognition elements, such as the discriminator base and the terminal adenosine. Mutation of the G2:C71 base pair to A2:U71 reduced suppression efficiency by ≈50%. Conversion of the A3:U70 base pair into a G3:U70 base pair resulted in a drop of suppression efficiency. This was unexpected, as this G3 is found in wild-type *M. barkeri* MS tRNA^{Pyl}. However, the MS tRNA also contains a C44U mutation in the first nucleotide of the variable loop. Because both independent mutations result in a decrease in suppression efficiency (Table 1) and because the wild-type tRNA^{Pyl} species from both *M. barkeri* strains are equally good substrates for suppression, the two variations compensate their respective negative impacts on the suppression efficiency.

Author contributions: A.A., S.G., and D.S. designed research; A.A., S.G., S.H., C.P., and C.F. performed research; A.A., S.G., S.H., C.P., and D.S. analyzed data; and A.A. and D.S. wrote the paper.

The authors declare no conflict of interest.

Abbreviations: Pyl, pyrrolysine; PylRS, pyrrolysyl-tRNA synthetase; DHFR, dihydrofolate reductase.

†Present address: Instituto de Bioquimica Medica, Universidade Federal do Rio de Janeiro, CEP 21941-901, Rio de Janeiro, Brazil.

‡Present address: Institute for Pharmacy and Molecular Biotechnology, University of Heidelberg, D-69117 Heidelberg, Germany.

¶To whom correspondence should be addressed at: Department of Molecular Biophysics and Biochemistry, Yale University, P.O. Box 208114, 266 Whitney Avenue, New Haven, CT 06520-8114. E-mail: dieter.soll@yale.edu.

Table 1. Suppression efficiency of *M. barkeri* tRNA^{Pyl} variants

Mutations	Suppression efficiency, %
Wild type	
Fusaro	100
MS (C44U/A3G)	95
Discriminator base	
G73→C	88
G73→A	33
G73→U	43
Acceptor stem	
G1→A	26
C72→U	30
G1:C72→C:G	70
G2:C71→A:U	52
A3→G	58
A5:U68→C:G	100
D-stem and loop	
U15→A	39
G18→U	87
U20→A	43
A21→U	95
G10:C25→U:A	50
A11:U24→C:G	88
U12:A23→G:C	112
C13:G22→A:U	25
Insert at G18	47
TψC-stem and loop	
U54→A	80
A56→U	33
A58→U	40
U60→A	108
C50:G64→A:U	79
G51:C63→U:A	29
G52:C62→U:A	63
G53:C61→U:A	63
Variable loop	
C44→U	56
A45→G	23
G48→U	17
Insert A at C44	17
Anticodon stem/loop	
A27:U43→G:A	56
U29a→C	40
Δ(U29aG41b)	12
G29b:C41a→U:A	80
G30:C40→ U:A	74
A31:U39→C:G	46
C32→A	31
U33→G	16
A37→C	13
A38→C	49

Suppression efficiency was measured by β-galactosidase activity in *E. coli* strain XAC/A24. A percentage of 100 corresponds to 3,800 Miller units (13). In the absence of Cyc, background suppression in all cases was <5%.

D- and T-Stem/Loop Regions. Losses in suppression observed upon mutation of nucleotides in these regions of the tRNA molecule are probably related to their role in maintaining D-loop/T-loop interactions, crucial for the ability of the tRNA to maintain its L-shape tertiary structure. Mutations of U20A in the D-loop and A56U in the T-loop resulted in loss of suppression efficiency. Similarly, the negative effect of the D-loop mutation U15A supports a role of this nucleotide in D-loop/T-loop interaction via nucleotide U59 (14). One of the most severe decreases in suppression efficiency was the T-loop mutation A58U, because it would disrupt the putative T-loop upon interaction with U54 (14). However, the U54A mutation resulted only in a minor reduction of suppression efficiency, suggesting that the integrity of the T-loop interaction is not critical for tRNA^{Pyl} activity and that A58 could be directly involved in PylRS binding and recognition. This idea is supported by the fact that mutation of A58 in *Desulfitobacterium hafniense* tRNA^{Pyl} resulted in a >1,000-fold loss of *in vitro* aminoacylation efficiency (15).

Base pair mutations G10:C25 to U:A, C13:G22 to A:U in the D-arm, and G51:C63 to U:A in the T-arm caused significant reduction of suppression efficiency. Although the observed effect of mutation of G10:C25 can be attributed to a role in maintaining the core structure of tRNA^{Pyl} via possible interactions with a nucleotide in the variable loop, base pairs C13:G22 and G51:C63 do not have obvious structural roles (14), suggesting a direct contribution of these nucleotides to tRNA^{Pyl} fitness in the translation machinery.

Variable Loop. The short variable loop (only three nucleotides instead of the normal five) is one of the distinctive features of tRNA^{Pyl}. Mutation of any of these nucleotides resulted in strong reduction of suppression efficiency; the most dramatic effects were observed in mutants A45G and G48U, and when an additional A was inserted to make a 4-nt variable loop. Such dramatic effects are consistent with the function of the variable loop nucleotides in ensuring the proper relative positioning of the two stacked helices that make up the tRNA L-shape. The effect of the A45G mutation may also be considered in light of the reduction observed upon mutating the first base pair of the D-arm (G10:C25) because they make a potential tertiary interaction (14).

Anticodon Stem and Loop. The elongated anticodon stem is another striking feature of tRNA^{Pyl}. Although the mutations of the anticodon stem base pairs resulted in only moderate suppression loss, three mutations are nevertheless worth noticing. Disruption of the first base pair of the anticodon stem (A27:U43 to G:A) lowered the suppression efficiency. Conversion of the wobble pair U29a:G41b to a C:G pair or of A31:U39 to C31:G39 also caused significant reductions of tRNA^{Pyl} fitness as a UAG suppressor. Deletion of the U29a:G41b pair, resulting in a canonical 5-bp anticodon stem was particularly detrimental to suppression efficiency. Although mutations in any of the four anticodon loop bases that flank the anticodon led to loss of suppression efficiency, the major effects were noticed upon mutation of the conserved nucleotides U33 and A37, directly adjacent to the anticodon triplet. From these results, we can infer that the length of the anticodon stem is critical for tRNA^{Pyl} activity because it places the important anticodon loop nucleotides at an appropriate distance from the 3′ terminal adenosine.

Effect of Mutations in tRNA^{Pyl} on *in Vitro* Charging by and *in Vivo* Binding to *M. barkeri* PylRS. Because suppression efficiency of tRNA reflects the sum of this molecule's properties from aminoacylation to codon recognition and through the later steps of protein synthesis, we wanted to determine the aminoacylation and PylRS binding properties of the mutant tRNAs most affected in suppression. Therefore, we determined the *in vitro* kinetic parameters for acylation of the corresponding tRNA^{Pyl} molecules with Cyc by the purified recombinant *M. barkeri* PylRS, and we measured the effect of point mutations in tRNA^{Pyl} on binding to PylRS by using the *in vivo* yeast three-hybrid method (Table 2). Both methods unambiguously demonstrated the lack of recognition of the tRNA^{Pyl} anticodon by PylRS. Generally, the anticodon is one of the major identity elements in tRNA synthetase recognition. Seryl-tRNA synthetase (SerRS), leucyl-tRNA synthetase (LeuRS), histidyl-tRNA synthetase (HisRS), and alanyl-tRNA synthetase (AlaRS) are the four other synthetases that do not, at least in *E. coli*, rely

Table 2. Aminoacylation and *in vivo* binding of *M. barkeri* tRNA^{Pyl} variants by *M. barkeri* PylRS

tRNA	Mutations	K_M, μM	k_{cat}, s^{-1}	k_{cat}/K_M	*L*	Binding, %
tRNA^{Pyl}	Wild type	0.23 ± 0.04	(4.0 ± 1.2) × 10^{-2}	1.74 × 10^{-1}	1	100
tRNA^{Pyl}	G73A	0.48 ± 0.14	(1.5 ± 0.2) × 10^{-3}	3.00 × 10^{-3}	58.0	55
tRNA^{Pyl}	C72U	0.10 ± 0.03	(1.6 ± 0.3) × 10^{-4}	1.60 × 10^{-3}	108.7	76
tRNA^{Pyl}	C13:G22→A:U	0.16 ± 0.04	(1.9 ± 0.3) × 10^{-2}	1.16 × 10^{-1}	1.5	95
tRNA^{Pyl}	A45G	0.51 ± 0.08	(3.2 ± 0.9) × 10^{-2}	6.31 × 10^{-2}	2.8	92
tRNA^{Pyl}	G48U	0.39 ± 0.04	(9.2 ± 3.1) × 10^{-3}	2.36 × 10^{-2}	7.4	79
tRNA^{Pyl}	G51:C63→U:A	0.97 ± 0.23	(8.2 ± 1.4) × 10^{-3}	8.49 × 10^{-3}	20.5	69
tRNA^{Pyl}	U33G	0.12 ± 0.02	(1.2 ± 0.2) × 10^{-3}	1.00 × 10^{-2}	17.4	67
tRNA^{Pyl}	A37C	0.20 ± 0.04	(1.3 ± 0.3) × 10^{-3}	6.63 × 10^{-3}	26.2	55
tRNA^{Pyl}	Anticodon CUA→CAA	0.09 ± 0.02	(2.0 ± 0.4) × 10^{-2}	2.19 × 10^{-1}	0.8	82
tRNA^{Pyl}	Anticodon CUA→AUA	0.67 ± 0.10	(1.7 ± 0.4) × 10^{-2}	2.57 × 10^{-2}	6.8	94
tRNA^{Pyl}	Anticodon CUA→CUU	0.17 ± 0.02	(3.2 ± 0.6) × 10^{-2}	1.88 × 10^{-1}	0.9	86
tRNA^{Ser}	Wild type	ND	ND	ND	ND	ND
tRNA$^{Ser/Pyl}$	Chimera 1	0.26 ± 0.04	(2.1 ± 0.4) × 10^{-2}	8.15 × 10^{-3}	2.13	78.3
tRNA$^{Ser/Pyl}$	Chimera 2	0.29 ± 0.05	(2.3 ± 0.5) × 10^{-2}	7.93 × 10^{-3}	2.19	82.6

In vivo binding of tRNA^{Pyl} to PylRS was determined by the yeast three-hybrid method as described in *Materials and Methods*. *L* represents loss of aminoacylation efficiency and is calculated as (k_{cat}/K_M mutant)/(k_{cat}/K_M wild type). ND, the activity could not be detected.

on anticodon recognition for aminoacylation of their cognate tRNA. For SerRS and LeuRS, this is explained by the number of tRNA isoacceptors, because Ser and Leu are the only two amino acids that can be incorporated in response to six different codons. For HisRS and AlaRS, anticodon recognition has been replaced by unique decisive features in their cognate tRNAs, such as G-1 for tRNA^{His} and the acceptor stem pair G3:U70 for tRNA^{Ala} (16).

The discriminator base G73, the nucleotides flanking the anticodon (U33 and A37), and the T-stem base pair G51:C63 are significant identity elements in tRNA^{Pyl}. Furthermore, destabilization of the first acceptor stem pair (G1:C72) by conversion to G1:U72 resulted in significant decrease of aminoacylation efficiency. The A45G, G48U, and C13:G22→A:U mutations had no impact on PylRS binding or charging, which suggests that the marked loss in translational fitness of these mutant tRNA^{Pyl} species (Table 1) is due to their inability to perform steps in protein synthesis that are downstream of aminoacylation. Additionally, the suppression data (Table 1) show the existence of tRNA^{Pyl} mutations (U60A and U12:A23→G:C) that increase translational fitness of this tRNA.

Transplantation of Cyc Acceptor Activity into Bovine Mitochondrial tRNA^{Ser}. As noted earlier (9, 11, 15), tRNA^{Pyl} has an unusual structure shared with only one *Bos taurus* mitochondrial tRNA^{Ser} isoacceptor (17). Neither tRNA$^{Ser}_{CUA}$ nor a stabilized version of this RNA (see *Materials and Methods*) were substrates for any endogenous *E. coli* synthetase *in vivo* or for *M. barkeri* PylRS *in vitro*. Therefore, we attempted to transplant the tRNA^{Pyl} identity elements into this bovine tRNA^{Ser} scaffold. We constructed two slightly different tRNA$^{Ser/Pyl}$ chimeric molecules (Fig. 1) that included the tRNA^{Pyl} discriminator base G73, base pairs G1:C72, C13:G22 in the D-arm, G51:C63 in the T-stem, and A31:U39 in the anticodon stem. However, addition of some nucleotides from the tRNA^{Pyl} core structure (U8, G10:C25, G26, C44, A45, and G48) were needed to generate tRNA$^{Ser/Pyl}$ chimeras with efficient Cyc charging properties (Table 2). The yeast three-hybrid data also showed acquisition of *in vivo* tRNA binding by PylRS when wild-type tRNA^{Ser} is compared with the two chimeric tRNAs (Table 2).

A more complex pattern is seen when translational fitness of the chimeric tRNAs was tested by *lacZ* amber mutant suppression in the *E. coli* XAC/A24 strain. Some chimeras (data not shown) allowed efficient suppression in the absence of Cyc, suggesting that the resulting tRNAs were charged by endogenous *E. coli* tRNA synthetases. Edman degradation (data not shown) of a reporter protein based on the *E. coli* dihydrofolate reductase (DHFR) system containing a UAG codon in position 3 (18) revealed the presence of glutamine (80%), threonine (12%), and

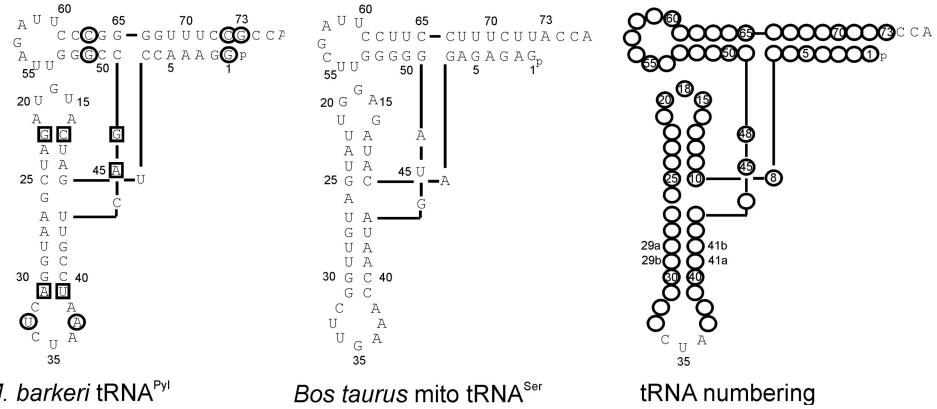

Fig. 1. *M. barkeri* tRNA^{Pyl} and bovine mitochondrial tRNA^{Ser} L-shape structures and scheme of tRNA numbering. The circled nucleotides indicate tRNA positions that, upon mutation, resulted in a significant decrease in *in vivo* suppression activity, *in vivo* binding, and *in vitro* aminoacylation. The boxed nucleotides indicate tRNA positions that, upon mutation, only affected *in vivo* suppression activity.

Table 3. Suppression efficiency of $tRNA^{Ser/Pyl}$ chimeras

tRNA	Suppression efficiency, %
$tRNA^{Ser}_{CUA}$	<5
Stabilized $tRNA^{Ser}_{CUA}$	<5
$tRNA^{Ser/Pyl}$ chimera 1	62
$tRNA^{Ser/Pyl}$ chimera 2	69
$tRNA^{Pyl}$	100

Suppression efficiency was measured by β-galactosidase activity in *E. coli* strain XAC/A24. In the absence of Cyc, background suppression in all cases was <5%.

serine (8%). The selectivity of the tRNAs toward PylRS could be reestablished by keeping the $tRNA^{Ser}$ nucleotides A8 (Fig. 1, chimera 1) or A26 and G44 (Fig. 1, chimera 2) in the chimeric tRNAs, even though their fitness in protein synthesis was somewhat lowered (Table 3). Maintaining all three $tRNA^{Ser}$ nucleotides at these positions yielded an inactive tRNA (data not shown). Taken together, these data show that Pyl identity elements could be successfully transplanted into the $tRNA^{Ser}$ scaffold.

$tRNA^{Pyl}$ Can Be Converted into an Efficient UGA Suppressor. Because the aminoacylation data (Table 2) clearly showed that the nature of the anticodon sequence is not critical for $tRNA^{Pyl}$ charging, we wanted to test whether introduction of a UCA anticodon would generate an efficient UGA suppressor $tRNA^{Pyl}$. Again, we used the strategy of suppressing *lac* nonsense mutations, either the U4 opal or the A24 amber mutant, in the *lacI–lacZ* fusion system (13). Upon transformation of the *E. coli* strains XAC/U4 and XAC/A24 with plasmids carrying *M. barkeri pylS* and either *pylT* (with a UCA anticodon) or wild-type *pylT*, the cells were grown in the presence of Cyc, and suppression efficiency was assessed by β-galactosidase activity (Table 4). The UGA suppressor $tRNA^{Pyl}_{UCA}$ is 81% as efficient as the native $tRNA^{Pyl}_{CUA}$ amber suppressor. Thus, Cyc can also be inserted into proteins in response to a UGA stop codon *in vivo*. A similar change in coding response was reported for $tRNA^{Sec}$ (19); anticodon changes led to efficient selenocysteine incorporation in response to all three termination codons (Table 4), even though the requirements for codon context are very different for Sec (UGA) and Pyl (UAG). Compared with the synthesis of wild-type LacZ protein from a $lacZ^+$ construct, the absolute suppression efficiency of $tRNA^{Pyl}_{UCA}$ is 20%, which is the same level as we observed (data not shown) and had been reported for the *E. coli* $tRNA^{Arg}_{UCA}$ suppressor (20).

Table 4. Decoding properties for UAG and UGA of *M. barkeri* $tRNA^{Pyl}$ and *E. coli* $tRNA^{Sec}$

tRNA	Anticodon	*lacZ* codon (mutation)	Suppression efficiency, %
$tRNA^{Pyl}$	CUA	UAG (A24)	100
$tRNA^{Pyl}$	CUA	UGA (U4)	<1
$tRNA^{Pyl}$	UCA	UAG (A24)	<1
$tRNA^{Pyl}$	UCA	UGA (U4)	81
$tRNA^{Sec}$	CUA	UAG	68*
$tRNA^{Sec}$	CUA	UGA	<1*
$tRNA^{Sec}$	UCA	UAU	<1*
$tRNA^{Sec}$	UCA	UGA	100*

Suppression efficiency was measured by β-galactosidase activity in *E. coli* strain XAC/U4 (for opal) and in *E. coli* strain XAC/A24 (for amber). In the absence of Cyc, background suppression in all cases was <5%.
*Suppression efficiencies for $tRNA^{Sec}$ are from ref. 19.

Discussion

Sec and Pyl are the 21st and 22nd genetically encoded amino acids. Unlike any of the 20 canonical amino acids in protein synthesis, these two unusual amino acids share the property of being cotranslationally inserted in-frame at UGA and UAG, respectively; these codons normally specify termination. However, the mechanism of reassigning these codons to sense appears to differ between Sec and Pyl (21). Although the insertion of Sec at UGA requires a specific mRNA context provided by the presence of the SECIS stem/loop structure as well the selenocysteine-specific elongation factor SelB (6), Pyl can be efficiently inserted into proteins in an anonymous context and appears not to depend on the presence of additional proteins. This characteristic is well demonstrated by the fact that Pyl or Cyc could be efficiently incorporated into *E. coli* reporter proteins lacking any specific RNA structures in the vicinity of the UAG codon when tested in *Methanosarcina acetivorans* (22) or in *E. coli* (12). A stem/loop structure located just downstream of the UAG codon and therefore termed PYLIS by analogy to the established SECIS element, was proposed to stimulate Pyl incorporation (22, 23). However, the lack of sequence and structure conservation between the predicted PYLIS structures from the three Pyl-containing methylamine methyltransferase genes is a clear argument against a functional role for such an RNA motif (21).

The difference in recoding strategies used for Pyl and Sec incorporation may reflect different evolutionary histories. A compelling body of data indicates that Sec was already present at the time of the last common ancestor and, therefore, already was using contextualized UGA as a sense codon for Sec insertion. Although Sec is present in organisms from all three domains of life, it is absent from many bacteria and most archaea and present in higher eukaryotes with the exception of plants (5, 6, 24, 25). Although the widespread presence of this unusual amino acid in organisms from all three domains is a clear sign of vertical inheritance from the last common ancestor, its absence from many organisms is attributed to the loss of Sec coding capability due to environmental factors (25). Based on the known genomes, the organismal distribution of PylRS suggests that this enzyme is a late archaeal invention designed to meet the specific physiological needs of a particular archaeal lineage. PylRS would then be another example of genetic code evolution after the era of the last common universal ancestor (5). The late apparition of PylRS implies that the *Methanosarcinaceae* progenitor added Pyl to its genetic code. The fact that the $tRNA^{Pyl}$ anticodon is not recognized by PylRS may then have conferred a significant advantage because it allowed testing of different codons, eventually selecting the UAG codon as the least disruptive for protein synthesis. In *M. barkeri*, for instance, the UAG codon is used at a much lower frequency (0.05%) than UAA (1.8%) or UGA (1.2%). Furthermore, when UAG codons are used as genuine stop codons, they are almost immediately followed by either UAA or UGA codons, thus reducing the negative impact that the unintended read-through of the UAG termination signal may have on the integrity of the *Methanosarcinaceae* proteomes (25). Our data suggest that $tRNA^{Pyl}$ anticodon variants may respond to any codon, a possibility that may be tested in an organism that does not use all sense codons [e.g., *Micrococcus luteus* (26)]. The nonessentiality of the $tRNA^{Pyl}$ anticodon sequence, combined with the absence of competition by any other aminoacyl-tRNA, and the low usage of UAG codons may have been determining factors for the successful insertion of Pyl into the *Methanosarcinaceae* genetic code.

Considerable efforts have been made in devising orthogonal aminoacyl-tRNA synthetase:tRNA pairs solely specific for recognition of nonnatural amino acids and in assuring their subsequent incorporation at termination codons into *E. coli* and

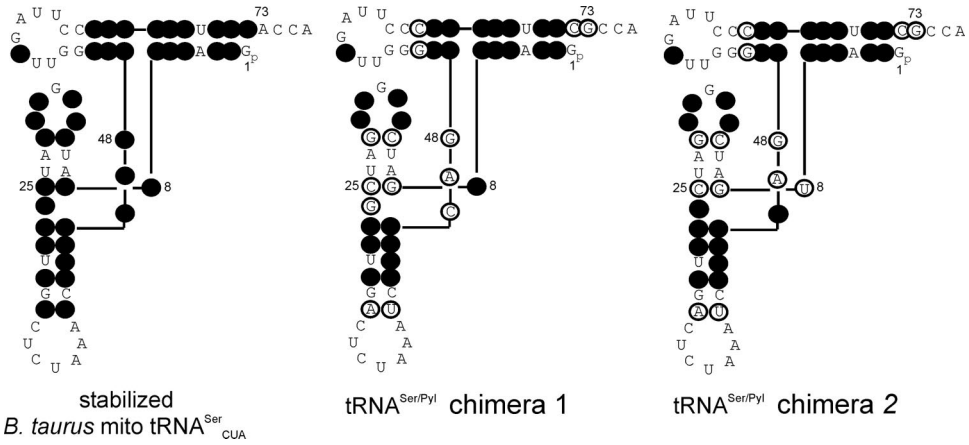

Fig. 2. Transplantation of *M. barkeri* tRNA^{Pyl} identity elements into the bovine mitochondrial tRNA^{Ser} scaffold. L-shape structures of stabilized tRNA^{Ser} (see *Materials and Methods*) and of two tRNA$^{Ser/Pyl}$ chimeric molecules (see *Results*). Filled circles refer to nucleotides different in the stabilized tRNA^{Ser} and tRNA^{Pyl}; circled nucleotides refer to positions mutated in the stabilized tRNA^{Ser} required to obtain Cyc accepting activity. The other indicated nucleotides are common to the stabilized tRNA^{Ser} and tRNA^{Pyl} species.

eukaryotic proteins (1, 3). Because Pyl-tRNA^{Pyl} formation is independent of the tRNA anticodon sequence, because the PylRS structure is being solved (27), because tRNA^{Pyl} is a strong suppressor (at least as tested for UAG and UAA) in *E. coli* (10, 12), and because additional RNA or protein factors are not required for Pyl or Cyc incorporation, the potential use of the PylRS:tRNA^{Pyl} orthogonal pair, optimized through natural or man-made evolutionary processes, is particularly appealing.

Materials and Methods

General. Oligonucleotide synthesis, DNA sequencing, and Edman degradation were performed by the Keck Foundation Biotechnology Resource Laboratory at Yale University. [γ-^{32}P]ATP (3,000 Ci/mmol) and [α-^{32}P]ATP (3,000 Ci/mmol) (1 Ci = 37 GBq) were from Amersham Biosciences (Uppsala, Sweden). Cyc was purchased from Sigma (St. Louis, MO). The yeast strain L40 was a gift from Marvin Wickens (University of Wisconsin, Madison, WI).

tRNA Mutants. Mutant tRNA genes were constructed by PCR (QuikChange kit; Qiagen, Hilden, Germany) using as templates the wild-type *M. barkeri pylT*, cloned into pTECH (12) for *in vivo* expression and pUC18 (28) for production of *in vitro* T7 RNA polymerase transcripts. The sequence of each *pylT* mutant construct was confirmed by dideoxy-automated sequencing. The *M. barkeri* tRNA^{Pyl} sequence has two variants: the *M. barkeri* MS tRNA^{Pyl} contains a G in position 3 and a U in position 44, whereas the tRNA of the Fusaro strain has an A and a C at the respective positions (8). The bovine mitochondrial tRNA^{Ser} was shown to be rather unstable (14). Therefore, a stabilized tRNA^{Ser} was obtained by replacing the four wobble pairs present in the molecule (Fig. 2) by standard Watson–Crick pairs: G5:U69 was replaced with G5:C69; G50:U64 and G51:U63 were replaced with the corresponding A:U pairs, and finally the G28:U42 pair was replaced with a G:C pair. Given the unusual structure of tRNA^{Pyl}, the nucleotide numbering is defined in Fig. 2.

Suppression of *E. coli* XAC/A24 and XAC/U4 Nonsense Mutations. *E. coli* strains XAC/A24 and XAC/U4 carry an inactivating mutation in the *lacI–lacZ* fusion system: Trp codon position 658 to UAG nonsense and leucine codon position 565 to UGA, respectively (13). Both strains have been used in the past for insertions of a number of different amino acids. For both strains, the presence of an in-frame stop codon results in the inability of this strain to degrade the chromogenic 2-nitrophenyl β-D-galactopyranoside. Cells were then cotransformed by electroporation with plasmids carrying the *M. barkeri* Fusaro *pylS* (pCBS) and *M. barkeri* Fusaro *pylT* (pTECH) wild-type and mutant genes. The transformants were grown in Luria broth containing 10 mM Cyc, and β-galactosidase activity was measured as previously mentioned (12). In cases of low suppression, Northern blots revealed that tRNA^{Pyl} was not limiting in the experiment (data not shown).

Suppression of a UAG (Codon 3) in an *E. coli folA* Reporter Gene. The *E. coli folA* gene (encoding DHFR) containing a UAG triplet in place of codon 3 was amplified from genomic DNA, cloned into pCR 2.1-TOPO plasmid vector (Invitrogen, Carlsbad, CA), sequenced, and subcloned into pRSFDuet-1. *E. coli* BL21(DE3)-competent cells were cotransformed with a tRNA$^{Ser/Pyl}$ chimera cloned into pTECH and the *folA* reporter construct and grown in LB at 37°C, in the presence of ampicillin and kanamycin. Production of recombinant DHFR was induced with 1 mM isopropyl β-D-thiogalactoside when cells reached an A_{600} of 1.0. The cells were harvested after 17 h; DHFR was purified by using standard Ni-NTA technology and then blotted onto a PVDF membrane. The first eight amino acids were identified by Edman degradation.

***In Vivo* Binding Affinity of tRNA^{Pyl} for PylRS by Using the Yeast Three-Hybrid System.** The yeast three-hybrid experiments were performed as described (29). The effect of point mutations in tRNA^{Pyl} on PylRS binding was measured by using the *in vivo* yeast three-hybrid method. Yeast strain YBZ-1, auxotrophic for histidine and carrying a genomic copy of the first hybrid molecule (MS2 coat protein fused to the GAL4-DB DNA binding protein), was transformed with two vectors encoding the RNA hybrid and protein hybrid molecules. The shuttle plasmid pIIIA/MS2–1 with a URA3 marker was used to express the RNA hybrid, an RNA hairpin with two MS2 binding sites for the coat protein fused to the 5′ end of tRNA^{Pyl}, as well as yeast tRNA^{Lys} and tRNA^{Gln} for specificity controls. The third hybrid is a fusion between the *M. barkeri* PylRS gene and the GAL4 activation domain (768–881) and was expressed in a 2-μm LEU2 vector pGADT7 under ADH1 promoter. When tRNA^{Pyl} interacts with its cognate PylRS through the chain-interaction of the three-hybrid protein and RNA molecules, GAL4-dependent transcription of a histidine biosynthetic gene and *lacZ* reporter genes is activated. Qualitatively, weak and strong interactions can be discriminated phenotypically by using the growth rate on minimal media lacking histidine and a quantitative assessment by

measurement of β-galactosidase activity in the yeast cell extract. $tRNA^{Lys}$ and $tRNA^{Gln}$ did not result in any binding.

Purification of Recombinant *M. barkeri* PylRS and *in Vitro* Transcription of tRNA Genes. The recombinant *M. barkeri* Fusaro PylRS was overexpressed and purified as reported (9). The *M. barkeri* Fusaro wild-type, mutants $tRNA^{Pyl}$, and $tRNA^{Ser/Pyl}$ chimeric molecules were prepared as reported (28).

tRNA Aminoacylation Assays. Aminoacylation assays were adapted from a recently described procedure (30). Aminoacylation reactions (10 μl) were carried out at 37°C (unless otherwise noted) in 100 mM Na-Hepes (pH 7.2)/25 mM $MgCl_2$/60 mM NaCl/5 mM ATP/1 mM DTT/10 mM Cyc/tRNA 3′-labeled with [α-^{32}P]ATP ranging in concentration from 0.25 to 5 times K_M concentrations. PylRS concentrations ranged from 5 to 90 nM. Reactions were stopped by removing 1 μl of the reaction and adding it to 3 μl of 2.5 units/μl nuclease P1 (where 1 unit liberates 4 μmoles of orthophosphate from 3′AMP per minute at 37°C; American Bioanalytical, Natick, MA) and incubated at 25°C for 30 min. Nuclease P1 digests were spotted on polyethyleneimine-cellulose plates (Baker, Phillipsburg, NJ), and [α-^{32}P]AMP (from uncharged tRNA) and aminoacyl-[α-^{32}P]AMP (from charged tRNA) were separated in a running buffer of 0.1 M sodium acetate and 5% acetic acid (31) and visualized by using a phosphoimager system. Kinetic parameters represent the results of at least three independent experiments.

Acid Urea Gel Electrophoresis of Aminoacyl-tRNA and Northern Hybridization. This method (32) allows the separation of charged and uncharged tRNAs based on the electrophoretic mobility difference seen between the two species. Hybridization to a ^{32}P-labeled sequence-specific probe permits the determination of the identity of the aminoacylated tRNA. The aminoacylation level of $tRNA^{Pyl}$ mutants isolated from *E. coli* strain XAC/A24 grown in the presence of 10 mM Cyc was verified. Unfractionated tRNA was isolated from each strain under acidic conditions [0.3 M sodium acetate (pH 4.5)/10 mM EDTA], dissolved in 2× loading buffer [7 M urea/0.3 M sodium acetate (pH 4.5)/10 mM EDTA/0.1% bromophenol blue/0.1% xylene cyanol] and were loaded (20 μg) on a 6.5% polyacrylamide gel containing 7 M urea and 0.1 M sodium acetate (pH 5.0). Northern blot was performed as described earlier (28).

C.F. held a fellowship of the Stiftung der Deutschen Wirtschaft. This work was supported by National Institute of General Medical Sciences Grant GM22854, Department of Energy Grant DE-FG02-98ER20311, and National Science Foundation Grant MCB-0645283.

1. Wang L, Xie J, Schultz PG (2006) *Annu Rev Biophys Biomol Struct* 35:225–249.
2. Köhrer C, RajBhandary UL (2005) in *The Aminoacyl-tRNA Synthetases*, eds Ibba M, Francklyn CS, Cusack S (Landes Bioscience, Georgetown, TX), pp 353–363.
3. Köhrer C, Sullivan EL, RajBhandary UL (2004) *Nucleic Acids Res* 32:6200–6211.
4. Anderson JC, Magliery TJ, Schultz PG (2002) *Chem Biol* 9:237–244.
5. Ambrogelly A, Palioura S, Söll D (2007) *Nat Chem Biol* 3:29–35.
6. Böck A, Thanbichler M, Rother M, Resch A (2005) in *Aminoacyl-tRNA Synthetases*, eds Ibba M, Francklyn CS, Cusack S (Landes Bioscience, Georgetown, TX), pp 320–327.
7. Krzycki JA (2005) *Curr Opin Microbiol* 8:706–712.
8. Srinivasan G, James CM, Krzycki JA (2002) *Science* 296:1459–1462.
9. Polycarpo C, Ambrogelly A, Berube A, Winbush SM, McCloskey JA, Crain PF, Wood JL, Söll D (2004) *Proc Natl Acad Sci USA* 101:12450–12454.
10. Blight SK, Larue RC, Mahapatra A, Longstaff DG, Chang E, Zhao G, Kang PT, Green-Church KB, Chan MK, Krzycki JA (2004) *Nature* 431:333–335.
11. Theobald-Dietrich A, Frugier M, Giegé R, Rudinger-Thirion J (2004) *Nucleic Acids Res* 32:1091–1096.
12. Polycarpo CR, Herring S, Bérubé A, Wood JL, Söll D, Ambrogelly A (2006) *FEBS Lett* 580:6695–6700.
13. Coulondre C, Miller JH (1977) *J Mol Biol* 117:525–567.
14. Hayashi I, Kawai G, Watanabe K (1998) *J Mol Biol* 284:57–69.
15. Herring S, Ambrogelly A, Polycarpo CR, Söll D (2007) *Nucleic Acids Res*, doi:10.1093/nar/gkl1151.
16. Giegé R, Sissler M, Florentz C (1998) *Nucleic Acids Res* 26:5017–5035.
17. Watanabe Y, Kawai G, Yokogawa T, Hayashi N, Kumazawa Y, Ueda T, Nishikawa K, Hirao I, Miura K, Watanabe K (1994) *Nucleic Acids Res* 22:5378–5384.
18. McClain WH (2006) *Proc Natl Acad Sci USA* 103:4570–4575.
19. Heider J, Baron C, Böck A (1992) *EMBO J* 11:3759–3766.
20. McClain WH, Foss K, Jenkins RA, Schneider J (1990) *Proc Natl Acad Sci USA* 87:9260–9264.
21. Zhang Y, Baranov PV, Atkins JF, Gladyshev VN (2005) *J Biol Chem* 280:20740–20751.
22. Longstaff DG, Blight SK, Zhang L, Green-Church KB, Krzycki JA (2007) *Mol Microbiol* 63:229–241.
23. Namy O, Rousset JP, Napthine S, Brierley I (2004) *Mol Cell* 13:157–168.
24. Gladyshev VN, Kryukov GV, Fomenko DE, Hatfield DL (2004) *Annu Rev Nutr* 24:579–596.
25. Zhang Y, Romero H, Salinas G, Gladyshev VN (2006) *Genome Biol* 7:R94.
26. Kano A, Ohama T, Abe R, Osawa S (1993) *J Mol Biol* 230:51–56.
27. Yanagisawa T, Ishii R, Fukunaga R, Nureki O, Yokoyama S (2006) *Acta Crystallogr F* 62:1031–1033.
28. Polycarpo C, Ambrogelly A, Ruan B, Tumbula-Hansen D, Ataide SF, Ishitani R, Yokoyama S, Nureki O, Ibba M, Söll D (2003) *Mol Cell* 12:287–294.
29. Hook B, Bernstein D, Zhang B, Wickens M (2005) *RNA* 11:227–233.
30. Wolfson AD, Pleiss JA, Uhlenbeck OC (1998) *RNA* 4:1019–1023.
31. Bullock TL, Uter N, Nissan TA, Perona JJ (2003) *J Mol Biol* 328:395–408.
32. Varshney U, Lee CP, RajBhandary UL (1991) *J Biol Chem* 266:24712–24718.