Chemistry & Biology
Article

Multistep Engineering of Pyrrolysyl-tRNA Synthetase to Genetically Encode N^{ε}-(*o*-Azidobenzyloxycarbonyl) lysine for Site-Specific Protein Modification

Tatsuo Yanagisawa,[1] Ryohei Ishii,[1,2,3] Ryuya Fukunaga,[1,2,4] Takatsugu Kobayashi,[1,2] Kensaku Sakamoto,[1,2] and Shigeyuki Yokoyama[1,2,*]
[1]Systems and Structural Biology Center, Yokohama Institute, RIKEN, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan
[2]Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
[3]Present address: Department of Pathology and Laboratory Medicine, Weill Medical College of Cornell University, 1300 York Avenue, New York, NY 10065, USA
[4]Present address: Department of Molecular and Cell Biology, University of California, Berkeley, 731 Stanley Hall 3220, Berkeley, CA 94720, USA
*Correspondence: yokoyama@biochem.s.u-tokyo.ac.jp
DOI 10.1016/j.chembiol.2008.10.004

SUMMARY

Pyrrolysyl-tRNA synthetase (PylRS) esterifies pyrrolysine to $tRNA^{Pyl}$. In this study, N^{ε}-(*tert*-butyloxycarbonyl)-L-lysine (BocLys) and N^{ε}-allyloxycarbonyl-L-lysine (AlocLys) were esterified to $tRNA^{Pyl}$ by PylRS. Crystal structures of a PylRS catalytic fragment complexed with BocLys and an ATP analog and with AlocLys-AMP revealed that PylRS requires an N^{ε}-carbonyl group bearing a substituent with a certain size. A PylRS(Y384F) mutant obtained by random screening exhibited higher in vitro aminoacylation and in vivo amber suppression activities with BocLys, AlocLys, and pyrrolysine than those of the wild-type PylRS. Furthermore, the structure-based Y306A mutation of PylRS drastically increased the in vitro aminoacylation activity for N^{ε}-benzyloxycarbonyl-L-lysine (ZLys). A PylRS with both the Y306A and Y384F mutations enabled the large-scale preparation (>10 mg per liter medium) of proteins site-specifically containing N^{ε}-(*o*-azidobenzyloxycarbonyl)-L-lysine (AzZLys). The AzZLys-containing protein was labeled with a fluorescent probe, by Staudinger ligation.

INTRODUCTION

In protein synthesis based on the universal genetic code, the 20 canonical amino acids are esterified to their cognate tRNAs by specific aminoacyl-tRNA synthetases (aaRSs) (Schimmel, 1987; Carter, 1993; Ibba and Söll, 2000), except that some aminoacyl-tRNAs are formed by conversion of their precursor aminoacyl-tRNAs (Gladyshev et al., 1996; Commans and Böck, 1999; Atkins and Gesteland, 2002; Ibba and Söll, 2002). The genetic code could be expanded to include the termination codon-directed incorporation of a non-canonical or non-natural amino acid into proteins. Genetic encoding of non-natural amino acids with useful physical, chemical, or biological properties (e.g. photocrosslinkers, fluorescent probes, spin labels, redox active groups, heavy atoms, and reactive groups) provides powerful tools to investigate and manipulate protein structure and function. More than 30 non-natural amino acids have been incorporated into proteins via a unique codon and an "orthogonal" tRNA•aaRS pair, which do not crossreact with the host counterparts (reviewed in Hendrickson et al., 2004; Wang et al., 2006). For example, *p*-azido-L-phenylalanine (AzPhe) was genetically encoded in *Escherichia coli*, with a mutant tyrosyl-tRNA synthetase (TyrRS) and an amber suppressor tRNA from the archaeon *Methanocaldococcus jannaschii* (Chin et al., 2002), and the azido group of the protein containing AzPhe site-specifically was selectively modified with phosphine derivatives by means of Staudinger ligation (Saxon and Bertozzi, 2000; Kiick et al., 2002; Tsao et al., 2005). In addition, the azido group was utilized for the selective labeling of a protein with a fluorescent probe (Wang et al., 2003; Speers et al., 2003; Deiters et al., 2003) and biotin (Link and Tirrell, 2003) by the Huisgen [3+2]-cycloaddition.

Pyrrolysine (Pyl, Figure 1A), a non-canonical lysine derivative with a methyl-pyrroline moiety, was first identified in the active site of monomethylamine methyltransferase from *Methanosarcina barkeri* (Atkins and Gesteland, 2002; Ibba and Söll, 2002; Srinivasan et al., 2002; Hao et al., 2002). Pyrrolysine is directly esterified to its specific tRNA ($tRNA^{Pyl}$), which has the anticodon (CUA) complementary to the UAG codon, by pyrrolysyl-tRNA synthetase (PylRS) (Blight et al., 2004; Polycarpo et al., 2004). Bacterial elongation factor Tu (EF-Tu) binds aminoacyl-$tRNA^{Pyl}$ and delivers it to the ribosome (Blight et al., 2004; Theobald-Dietrich et al., 2004). Thus, pyrrolysine is translationally incorporated, in response to the amber stop codon UAG, into the *M. barkeri* methyltransferase in *E. coli* cells expressing *M. barkeri* PylRS and $tRNA^{Pyl}$ (Blight et al., 2004). Because the primary and secondary structures of $tRNA^{Pyl}$ have a number of unusual characteristics, including the CUA anticodon, the extended acceptor stem, and the extremely small D-loop (Srinivasan et al., 2002), $tRNA^{Pyl}$ is not recognized by the *E. coli* aaRSs. Therefore, PylRS and the amber suppressor $tRNA^{Pyl}$ from *Methanosarcina* function as an orthogonal aaRS•tRNA pair in *E. coli*, and thus are suitable for the incorporation of unusual non-natural and non-canonical amino acids into proteins (Polycarpo et al., 2006; Neumann et al., 2008).

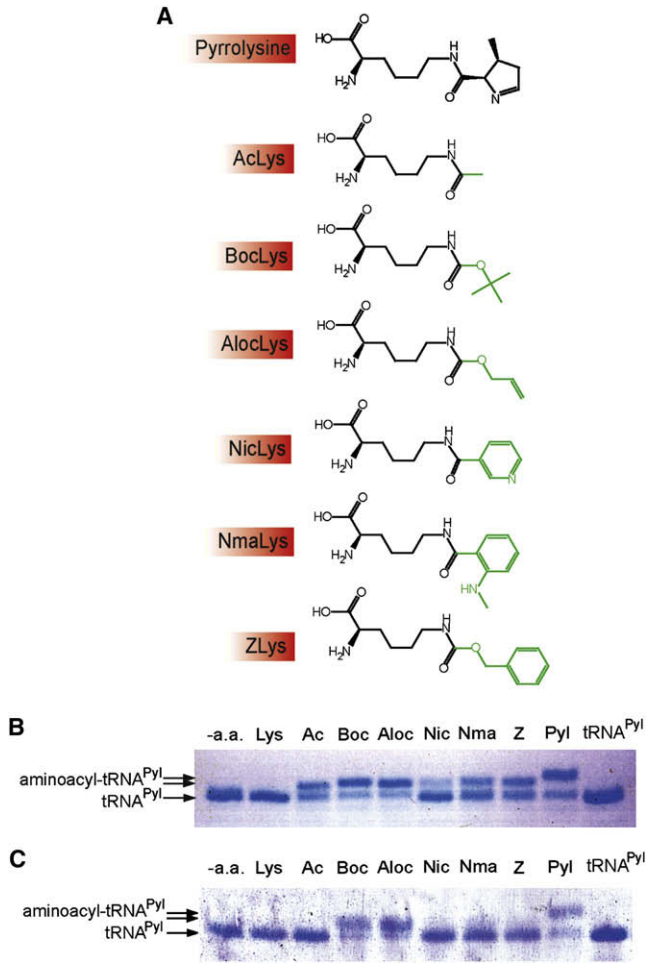

Figure 1. PylRS Activates a Variety of Non-natural and Non-canonical Amino Acids

(A) Chemical structures of pyrrolysine and lysine derivatives. Each of the unique groups is colored green. (B and C) Detection of aminoacylated tRNA by acidic urea PAGE analysis. The aminoacylation assay conditions are described in Experimental Procedures.

(B) Each lane shows a reaction in the presence of PylRS (2.8 μM) and with the following: no amino acid; 0.5 M Lys; 100 mM AcLys; 1 mM BocLys; 1 mM AlocLys; 10 mM NicLys; 7 mM NmaLys; 3.5 mM ZLys; 1 mM pyrrolysine; control $tRNA^{Pyl}$.

(C) Each lane shows a reaction in the presence of PylRS (2.8 μM) and with the following: no amino acid; Lys; AcLys; BocLys; AlocLys; NicLys; NmaLys; ZLys; pyrrolysine; control $tRNA^{Pyl}$. The concentration of each amino acid in the assay solutions was 1 mM.

Besides pyrrolysine, several non-natural pyrrolysine/lysine derivatives are esterified to $tRNA^{Pyl}$ by PylRS (Polycarpo et al., 2006). By using PylRS and $tRNA^{Pyl}$ in *E. coli*, some of these non-natural pyrrolysine/lysine derivatives are incorporated into proteins in response to the amber codon (Polycarpo et al., 2006). In order to expand the range of non-natural amino acids that can be incorporated into proteins, some active mutants of PylRS have been selected from randomly mutated libraries. First, six mutations in *M. barkeri* PylRS and four mutations in *Methanosarcina mazei* PylRS enabled the in vivo suppression with a minimal amino-acid substrate, N^{ε}-acetyl-L-lysine (AcLys) (Neumann et al., 2008; Mukai et al., 2008). Second, another set of five mutations in *M. mazei* PylRS facilitated the in vivo suppression with N^{ε}-benzyloxycarbonyl-L-lysine (ZLys), which is bulkier than pyrrolysine (Mukai et al., 2008). The *M. mazei* PylRS and $tRNA^{Pyl}$ are orthogonal not only to the bacterial systems but also to the eukaryotic systems, and thus these lysine derivatives have been successfully incorporated into a protein in yeast and mammalian cells (Mukai et al., 2008).

PylRS belongs to subclass c (Cusack, 1995) of the class-II aaRSs (Eriani et al., 1990; Cusack et al., 1990; Ruff et al., 1991). Crystal structures of a catalytic-domain fragment, including the C-terminal 270 residues of *M. mazei* PylRS [PylRS(c270), residues 185–454], have been reported by Kavran et al. (2007) and by our group (Yanagisawa et al., 2006, 2008). The pyrrolysyl moiety of the enzyme-bound pyrrolysine and pyrrolysyladenylate (pyrrolysyl-AMP) are recognized within the deep catalytic pocket of PylRS (Kavran et al., 2007; Yanagisawa et al., 2008).

In the present study, we show that *M. mazei* PylRS esterifies $tRNA^{Pyl}$ efficiently with N^{ε}-(*tert*-butyloxycarbonyl)-L-lysine (BocLys) and N^{ε}-allyloxycarbonyl-L-lysine (AlocLys). The crystal structures of *M. mazei* PylRS(c270) complexed with BocLys and adenosine 5′-(β,γ-imido)triphosphate (AMPPNP), and with AlocLys-AMP and imido-diphosphate (PNP) revealed the common features of these amino acids for the efficient recognition and aminoacylation by PylRS. A single mutation, Y384F, in *M. mazei* PylRS enhances the in vitro aminoacylation and in vivo suppression activities for BocLys and pyrrolysine. However, a single structure-based mutation, Y306A, which expands the amino acid binding pocket, successfully enabled PylRS to esterify ZLys to $tRNA^{Pyl}$. Furthermore, N^{ε}-(*o*-azidobenzyloxycarbonyl)-L-lysine (AzZLys) was also esterified to $tRNA^{Pyl}$ by the Y306A mutant, which the previous PylRS mutant could not accomplish. By combining the two PylRS mutations, Y384F and Y306A, AzZLys was successfully encoded in *E. coli*, and a milligram-quantity of a protein containing AzZLys at the amber-codon site was produced. The protein containing AzZLys site-specifically was selectively modified with a fluorescent probe. As compared with AzPhe, AzZLys has a longer linker between the azido group and the main chain; therefore, the efficiency of the site-specific modification of AzZLys is significantly higher.

RESULTS

PylRS Esterifies $tRNA^{Pyl}$ with a Variety of Non-natural Amino Acids

M. barkeri PylRS reportedly has the activity to esterify $tRNA^{Pyl}$ with several pyrrolysine/lysine derivatives, such as N^{ε}-cyclopentyloxycarbonyl-L-lysine (CycLys), N^{ε}-(D-prolyl)-L-lysine, and N^{ε}-(3-me-2,3-dihydro-pyrrol-2-yl)carbonyl-L-lysine (Polycarpo et al., 2006). All of these lysine derivatives have the N^{ε}-carbonyl group. In the present study, we examined whether the full-length *M. mazei* PylRS can esterify $tRNA^{Pyl}$ with other types of lysine derivatives with the N^{ε}-carbonyl group (Figure 1A). In order to detect aminoacylation even with poor-substrate amino acids, acidic polyacrylamide gel electrophoresis (PAGE) assays were performed at an amino-acid concentration (1–100 mM) much higher than the K_m for pyrrolysine, at a high concentration of the enzyme (2.8 μM), and for a long duration (1 hr). First, 1 mM concentrations of N^{ε}-(*tert*-butyloxycarbonyl)-L-lysine (BocLys)

and N^{ε}-allyloxycarbonyl-L-lysine (AlocLys) (Figure 1A) were esterified to $tRNA^{Pyl}$, as well as pyrrolysine at the same concentration (Figures 1B and 1C). Second, N^{ε}-nicotinoyl-L-lysine (NicLys), the fluorescent amino acid N^{ε}-(*N*-me-anthraniloyl)-L-lysine (NmaLys), and N^{ε}-benzyloxycarbonyl-L-lysine (ZLys) were esterified at higher concentrations (10 mM, 7 mM, and 3.5 mM, respectively) (Figures 1A and 1B). Note that the substituents at the N^{ε}-carbonyl group in NicLys, NmaLys, and ZLys are larger than those in pyrrolysine, BocLys, and AlocLys. Third, N^{ε}-(*p*-azidobenzoyl)-L-lysine, N^{ε}-biotinyl-L-lysine, and N^{ε}-(9-fluorenylmethoxycarbonyl)-L-lysine, which have even larger substituents at the N^{ε}-carbonyl group, were not detectably esterified to $tRNA^{Pyl}$ by *M. mazei* PylRS (data not shown). However, N^{ε}-acetyl-L-lysine (AcLys) at 100 mM was esterified to $tRNA^{Pyl}$ less efficiently than either pyrrolysine, BocLys, or AlocLys at 1 mM by *M. mazei* PylRS (Figures 1A and 1B), as in the case of *M. barkeri* PylRS (Polycarpo et al., 2006). Indeed, AcLys, NmaLys, NicLys, and ZLys at 1 mM were not esterified to $tRNA^{Pyl}$, which indicates that these four lysine derivatives are much worse substrates than pyrrolysine, BocLys, and AlocLys (Figure 1C). Taken together, these results clearly show that PylRS has a preferred size for the substituent at the N^{ε}-carbonyl group. In contrast, *M. mazei* PylRS exhibits no ligation activity for lysine derivatives without the N^{ε}-carbonyl group, such as N^{ε}-methyl-L-lysine, N^{ε}-dimethyl-L-lysine, N^{ε}-trimethyl-L-lysine, N^{ε}-isopropyl-L-lysine, N^{ε}-dansyl-L-lysine, N^{ε}-(*o*,*p*-dinitrophenyl)-L-lysine, N^{ε}-(*p*-toluenesulfonyl)-L-lysine, and N^{ε}-(DL-2-amino-2-carboxyethyl)-L-lysine, regardless of the size of the N^{ε}-substituent (data not shown).

Structural Basis for Non-natural Amino Acid Recognition by PylRS

The crystal structures of the BocLys•AMPPNP-bound PylRS(c270) and the AlocLys-AMP•PNP-bound PylRS(c270) were determined at 1.79 Å and 2.06 Å resolutions, respectively (Table 1; see Figures S1, S2A, and S2B available online). In the PylRS(c270)•BocLys•AMPPNP structure, the electron density of the enzyme-bound BocLys is clearly visible in the active site (Figure S2C). The N^{ε}-Boc group is buried in a hydrophobic environment, in a similar manner to those observed in the pyrrolysine•AMPPNP-bound (Figure 2A and 2B) (Yanagisawa et al., 2008) and the N^{ε}-cyclopentyloxycarbonyl-D-lysine (CycDLys)•ATP-bound forms (Kavran et al., 2007); the side-chain amide group of the PylRS-invariant Asn346 hydrogen bonds with the N^{ε}-carbonyl group of BocLys, and the electron density of the lysyl moiety is slightly weaker than that of the N^{ε}-substituent. The BocLys α-carboxyl group points away from the active site (Figures S2C and S2B), but it could rotate around the Cα-Cβ bond to the position of the pyrrolysine α-carboxylate (Figure 2A) to react with the α-phosphate group of ATP. Furthermore, the α-carbonyl group of BocLys makes a water-mediated hydrogen bonding interaction with the Asn346 side-chain carbonyl group, whereas the α-amino group interacts with the α-phosphate group of AMPPNP (Figure 2B, right). In addition, the enzyme-bound BocLys does not hydrogen bond with Arg330.

In the PylRS(c270)•AlocLys-AMP•PNP structure, the electron densities of the enzyme-bound AlocLys-AMP and the PNP being released are visible in the active site (Figure 2C). This indicates that the structure represents the state after the amino acid activation step. The N^{ε}-allyloxy moiety is buried in the hydrophobic environment. We initially tested a model of AlocLys-AMP in which the N^{ε}-carbonyl group points inward in the active site pocket (Figure 2C), as do those of pyrrolysine (Figure 2A) (Yanagisawa et al., 2008), BocLys (Figure 2B), pyrrolysyl-AMP (Kavran et al., 2007), and CycDLys (Kavran et al., 2007). However, the remaining electron density was still observed in the *Fo*-*Fc* omit map (red circle in Figure S2D, lower panel). Next, the AlocLys-AMP

Table 1. Data Collection and Refinement Statistics

	M. mazei PylRS(c270) •BocLys •AMPPNP Native	*M. mazei* PylRS(c270) •AlocLys-AMP •PNP Native
Data collection		
X-ray source	SPring-8 BL41XU	SPring-8 BL41XU
Wavelength	1.0000	1.0000
Space group	$P6_4$	$P6_4$
Cell dimensions		
a (Å)	104.71	104.15
b (Å)	104.71	104.15
c (Å)	70.55	71.06
α, β, γ (°)	90, 90, 120	90, 90, 120
Resolution (Å)	50–1.79 (1.82–1.79)	50–2.06 (2.10–2.06)
I/σ (I)	38.39 (1.54)	36.66 (2.67)
Completeness (%)	99.4 (96.0)	98.9 (82.4)
No. reflections	41,338	26,977
Redundancy (%)	8.2	8.5
R_{sym} [a]	6.3 (48.6)	5.9 (33.9)
Refinement		
R_{work} [b] / R_{free} [c] (%)	19.1 / 22.2	18.3/21.3
Resolution (Å)	34.28–1.79	42.0–2.06
No. atoms		
Protein	2095	2074
Others	54	47
Water	217	152
No. reflections (work / test)	37,169 / 4,162	24,284/2,681
Average B-factors		
Protein	45.1	50.2
Ligands	45.5	101.0
Water	56.4	55.5
Rmsd		
Bond length (Å)	0.01	0. 009
Bond angles (°)	1.6	1.5
Ramachandran plot		
Most favored (%)	94.6	91.6
Allowed (%)	5.4	8.4
Disallowed (%)	0.0	0.0

The numbers in parentheses are for the last shell.
[a] $R_{sym} = \Sigma\, |I_{avg}\text{-}I_i| \,/ \Sigma I_i$.
[b] $R_{work} = \Sigma\, |F_o\text{-}F_c| \,/\, \Sigma\, F_o$ for reflections of work set.
[c] $R_{free} = \Sigma\, |F_o\text{-}F_c| \,/\, \Sigma\, F_o$ for reflections of test set (10% of total reflections for *M. mazei* PylRS[c270]).

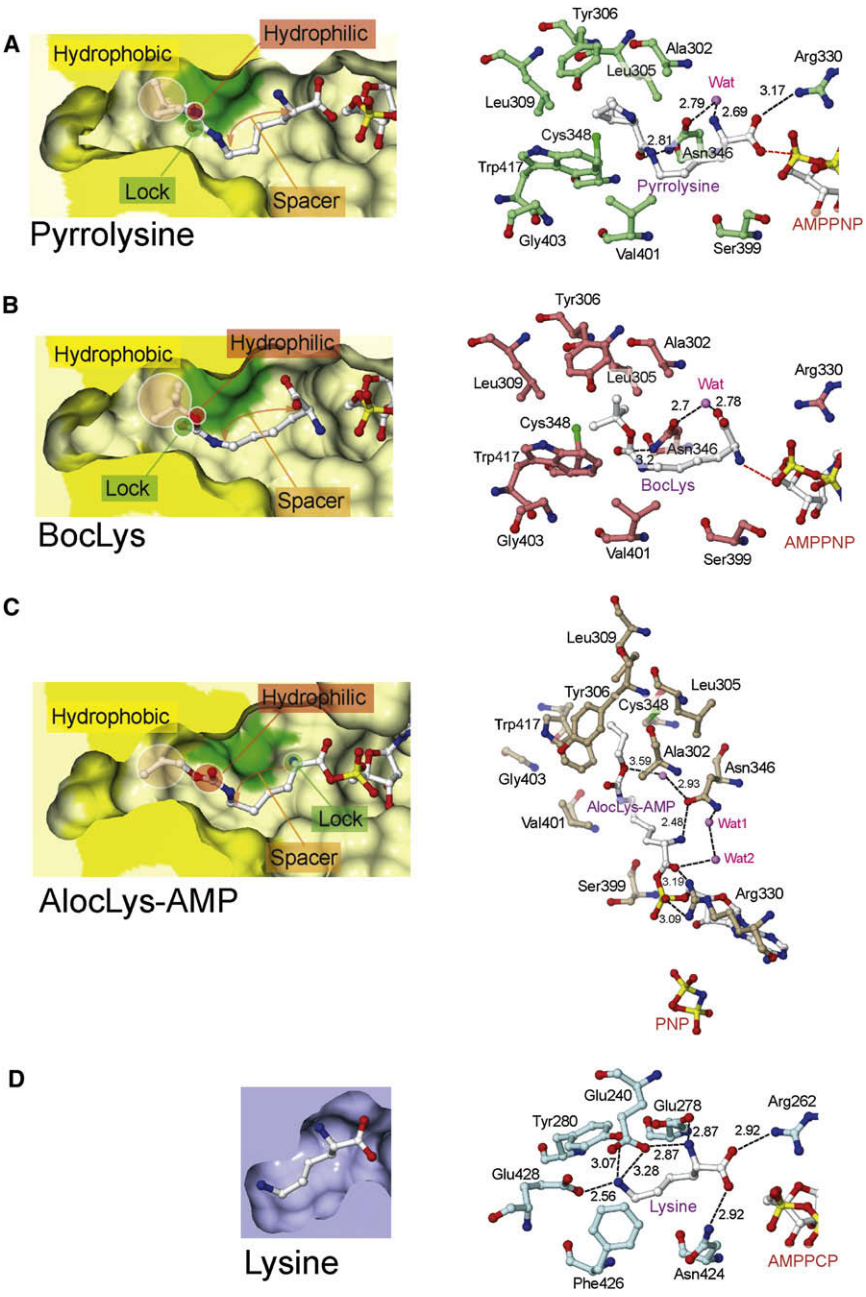

Figure 2. Comparison of the Amino Acid Binding Pocket of PylRS with that of LysRS

(A–D) The structures (Protein Data Bank codes): PylRS in complex with pyrrolysine (2ZCE) (A) (Yanagisawa et al., 2008), in complex with BocLys (2ZIN) (B), in complex with AlocLys-AMP (2ZIO) (C), and LysRS in complex with lysine (1E22) (D) (Desogus et al., 2000). The active site residues of PylRS are represented as surface models (left) and ball-and-stick models (right). The amino acid binding pocket of PylRS is much larger than those of any of the class-II aaRSs, and the amino acid recognition modes are distinct between PylRS and the other aaRSs. The lysine moiety is loosely recognized by PylRS, and simply acts as a "spacer." The large, hydrophobic pocket (Ala302, Leu305, Leu309, Tyr306, Val401, and Trp417 are shown as a yellow region), which is far from the ATP-binding site, traps the bulky, hydrophobic moiety. The side-chain amide group of Asn346 (shown as a green region) specifically interacts with and locks the side-chain carbonyl of the substrate (pyrrolysine and BocLys), or the side-chain carbonyl group of Asn346 (shown as a green region) specifically interacts with the main-chain α-amino group of the substrate (AlocLys-AMP). In addition, non-natural lysine derivatives should have a hydrophilic moiety, which points outside the active site cleft.

N^{ε}-carbonyl group was modeled to point outward from the pocket, and no residual electron density was observed (Figure S2D, upper panel). Therefore, we concluded that the N^{ε}-carbonyl group points outward in the case of AlocLys-AMP (Figure 2C, right panel). A superposition of the enzyme-bound pyrrolysine, BocLys, and AlocLys-AMP structures is shown in Figure S2E. The N^{ε}-carbonyl groups of pyrrolysine and BocLys point inward and directly hydrogen bond with the side-chain amide of Asn346. In contrast, the oxygen atom of the allyloxy moiety, rather than the N^{ε}-carbonyl group, of AlocLys-AMP interacts by a water-mediated hydrogen bond with the side-chain carbonyl group of Asn346 (Figure 2C, right panel). Furthermore, the α-amino group of the AlocLys moiety directly hydrogen bonds with the side-chain carbonyl group of Asn346 (Figure 2C, right panel). The α-carboxyl and α-phosphate groups of the enzyme-bound AlocLys-AMP interact with Arg330. In addition, the α-carboxyl group of the AlocLys moiety might make a two-water-mediated hydrogen bond bridge with the side-chain amide of Asn346. Accordingly, PylRS is likely to recognize the substrate amino acids broadly by their characteristics, such as a hydrophobic moiety with an appropriate size and bulkiness accommodated in the hydrophobic pocket, hydrogen-bond acceptors/donors located adjacent to Asn346, and the suitable length of the side-chain spacer (Figures 2A, 2B, and 2C).

Mutational Studies of PylRS for Enhanced Amber Suppression, and Site-Specific Incorporation of Non-natural Amino Acids into a Protein

The lysine side-chain moiety, from Cβ to Nε, simply acts as the spacer between the main-chain Cα and the hydrophobic (e.g., methyl-pyrroline, *tert*-butyl, and allyl) moiety (see Figures 2A, 2B, and 2C, left). The deep amino acid binding "tunnel" is twice as deep as the L-lysine-binding pocket in lysyl-tRNA synthetase (LysRS) (Figure 2D). The Ala mutations of PylRS(c270) at the highly conserved residues Leu305, Tyr306, Asn346, Val401,

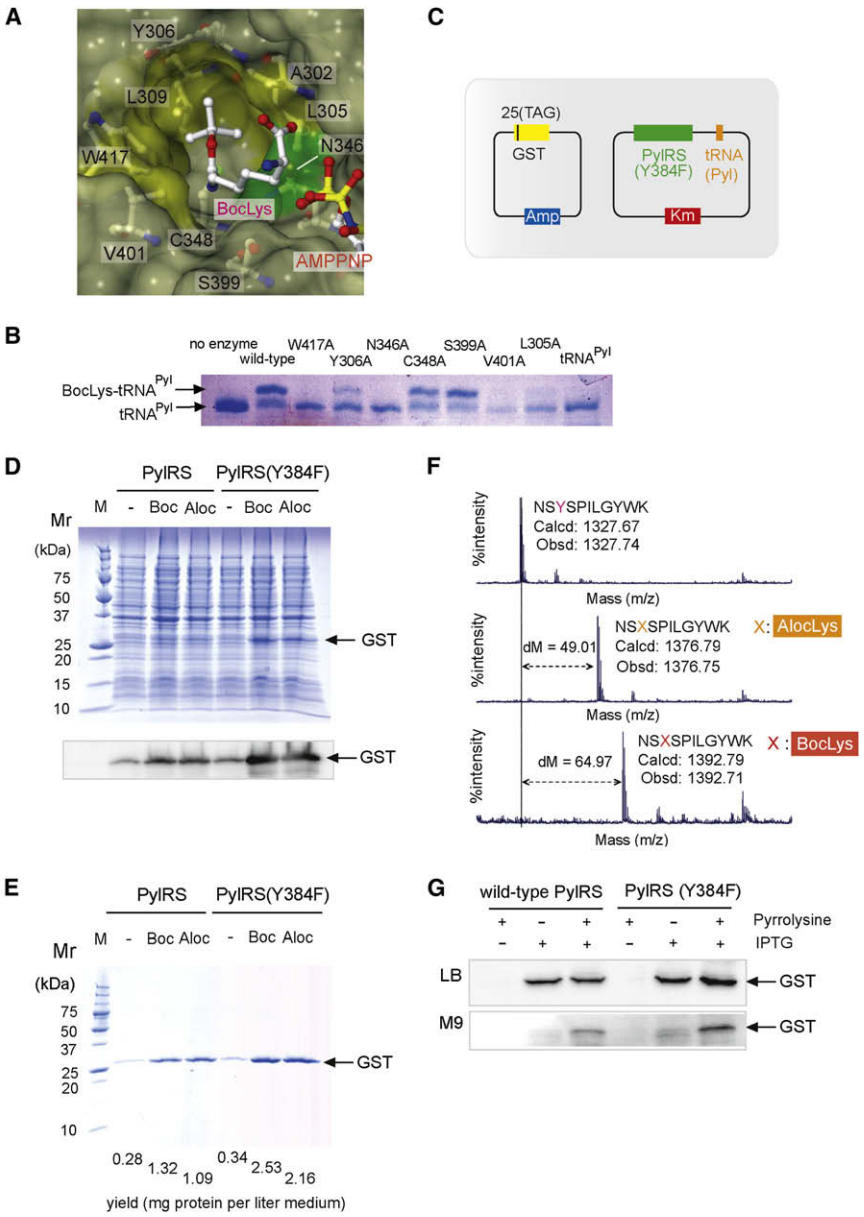

Figure 3. Active-Site Mutations Affect In Vitro PylRS Activity for Non-natural Amino Acids, and Non-natural Amino Acid Dependent Amber Suppression in *E. Coli*

(A) Ball-and-stick model of BocLys and AMPPNP bound within the PylRS(c270) active site. PylRS(c270) is represented as a surface model. Hydrophobic residues in the deep catalytic tunnel (Ala302, Leu305, Tyr306, Leu309, Val401, and Trp417) and the highly conserved Asn346, which hydrogen bonds with BocLys, are colored yellow and green, respectively.

(B) Active-site mutations of PylRS(c270) affect the aminoacylation activity for BocLys (1 mM). A 9 μM concentration of the wild-type or mutant PylRS(c270) protein was used. Starting from the left, each lane shows a reaction with the following: no enzyme; wild-type PylRS(c270); W417A; Y306A; N346A; C348A; S399A; V401A; L305A; control $tRNA^{Pyl}$.

(C) Amber suppression system using PylRS and $tRNA^{Pyl}$. The Tyr25 codon of the GST gene was mutated to TAG.

(D) When the PylRS Y384F mutant was co-expressed with $tRNA^{Pyl}$ in the presence of non-natural amino acids, the amounts of the full-length GST proteins obtained were increased by 2-fold, as compared with those produced with the wild-type PylRS. The increased readthrough products were detected by using the PylRS(Y384F)-$tRNA^{Pyl}$ system (for BocLys and AlocLys) (upper panel). These products are the GST protein, as confirmed by western blotting with an anti-GST antibody (lower panel) (Supplemental Experimental Procedures).

(E) The purified GST proteins, from in vivo amber suppression with wild-type PylRS and PylRS(Y384F), in the absence (–) or presence of BocLys (Boc) and AlocLys (Aloc).

(F) The purified GST proteins were subjected to in-gel trypsin digestion, and were analyzed by MALDI-TOF mass spectrometry to confirm the predominant incorporation of non-natural Lys derivatives into the protein.

(G) Pyrrolysine-dependent amber suppression in *E. coli*. The full-length GST proteins were detected by western blot analysis. When PylRS(Y384F) was coexpressed with $tRNA^{Pyl}$ in LB (upper panel) and M9 medium (lower panel) in the presence of pyrrolysine, the amount of the expressed GST protein was increased, as compared with the amount produced by wild-type PylRS. The readthrough products were increased in M9 medium with pyrrolysine, as compared with those without pyrrolysine.

and Trp417 impaired the activity for BocLys (Figures 3A and 3B), as in the case of the natural substrate, pyrrolysine (Yanagisawa et al., 2008).

To test the in vivo incorporation of BocLys and AlocLys into a protein site-specifically by amber suppression, cells harboring the glutathione S-transferase (GST) gene with an amber codon at position 25 [GST(25TAG)], and the PylRS and $tRNA^{Pyl}$ genes were induced with 1 mM isopropyl β-D-thiogalactopyranoside in the absence or presence of BocLys/AlocLys (Figure 3C). When the wild-type PylRS was coexpressed with $tRNA^{Pyl}$ in the presence of 4 mM BocLys and AlocLys, small amounts of the 28 kDa proteins were produced (Figure 3D). Western blot analysis with an anti-GST antibody confirmed that the 28 kDa proteins are the full-length GST (Figure 3D, lower panel).

Furthermore, to explore mutations for increasing the in vivo suppression activity for BocLys, we screened a PylRS random mutation library in the presence of BocLys, and successfully obtained the Y384F mutant of PylRS, which efficiently incorporated BocLys site-specifically into the GST protein in *E. coli* (Supplemental Experimental Procedures). The PylRS(Y384F) mutant, coexpressed with $tRNA^{Pyl}$ in the presence of 4 mM BocLys or AlocLys, produced a much larger amount of the GST protein than the wild-type PylRS (Figure 3D). The GST proteins containing BocLys/AlocLys produced by PylRS(Y384F) were purified on

a glutathione-affinity column, with yields of 2.53 and 2.16 mg protein per liter medium for BocLys and AlocLys, respectively. These yields are more than 1.91-fold and 1.98-fold of those for BocLys and AlocLys (1.32 and 1.09 mg per liter medium, respectively), respectively, with the wild-type PylRS (Figure 3E). The GST proteins obtained with the wild-type PylRS and PylRS(Y384F) in the absence of any added amino acids were 0.28 and 0.34 mg per liter medium, respectively (Figure 3E). The identity of the amino acid inserted in response to the in-frame UAG codon was confirmed by matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry analyses of the tryptic fragments from the readthrough products (Figure 3F, Figure S3, and Table S1). The peptide map of the calibrated digests revealed major peaks (obsd: m/z 1,392.71 $[M+H]^+$, m/z 1,376.75 $[M+H]^+$), which are within the theoretical masses of the tryptic peptide NSXSPILGYWK, where X represents BocLys and AlocLys, respectively (calcd: m/z 1,392.79 $[M+H]^+$, m/z 1,376.79 $[M+H]^+$). The molecular masses of these peptides are larger by 64.97 and 49.01 Da, respectively, than that of the wild-type tryptic peptide, NSYSPILGYWK (calcd: m/z 1,327.67 $[M+H]^+$, obsd: m/z 1,327.74 $[M+H]^+$) (Figure 3F). The sequence information from the spectrum demonstrated the site-specific incorporation of the non-natural amino acids into the GST protein.

The Y384F mutant PylRS was further tested for the incorporation of the natural substrate, pyrrolysine, into GST. The amount of the expressed GST protein produced by PylRS(Y384F) in the presence of 1 mM pyrrolysine in LB medium was increased, as compared with the amount produced by the wild-type PylRS (Figure 3G, upper panel). In these assays, the background of UAG expression in LB medium was high, and a substantial amount of GST was expressed in the absence of pyrrolysine. This high background appears to be due to some amino acid(s) in LB medium that is (are) incorporated into the amber-codon site in the absence of pyrrolysine. To test this possibility, we examined the amber suppression assay in M9 minimum medium. In fact, the amount of the expressed GST was much smaller in M9 medium than in LB medium, and was drastically increased with 1 mM pyrrolysine (Figure 3G, lower panel). It should be emphasized again that the screened Y384F mutation of PylRS resulted in much better amber suppression or expression of GST in the presence of pyrrolysine than in its absence. Therefore, the Y384F mutation increases the aminoacylation activity of PylRS, regardless of the amino acid substrate.

Effect of the Y384F Mutation on the PylRS Activity

The replacement of Tyr384 with Phe increased the in vivo suppression activity for non-natural lysine derivatives. Due to disorder, the Tyr384 residue at the tip of the β7–β8 hairpin (Figures S1, S2A, and S2B) was invisible in the present BocLys•AMPPNP-bound form as well as in the pyrrolysine•AMPPNP-bound form (Yanagisawa et al., 2008). Only the main chain of Tyr384 in the present AlocLys-AMP•PNP-bound form is visible, and it lies far from the active site, in a similar manner to that in the apo form (Yanagisawa et al., 2008). In the pyrrolysyl-AMP-bound form, Tyr384 is located much closer to the active site, as if it seals the amino acid binding tunnel (Kavran et al., 2007), because of a drastic conformational change of the β7–β8 hairpin (Yanagisawa et al., 2008). The conformational transition of the β7–β8 hairpin is not induced by pyrrolysine-binding, but instead occurs randomly (Yanagisawa et al., 2008).

In a prolonged reaction at a high enzyme concentration, the PylRS(Y384F) mutant protein produced the same levels of pyrrolysyl-$tRNA^{Pyl}$ and BocLys-$tRNA^{Pyl}$, as compared with those generated by the wild-type PylRS (Figure 4A). However, a time course analysis of the aminoacylation assay revealed that the PylRS Y384F mutant esterifies $tRNA^{Pyl}$ with BocLys and pyrrolysine significantly faster than the wild-type PylRS does (Figures 4B and 4C). The specific activities of aminoacylation (pmol aminoacyl-tRNA formed/pmol enzyme/min) for BocLys and pyrrolysine by the wild-type PylRS are low (0.09 and 0.04 min^{-1}, respectively), whereas those for BocLys and pyrrolysine by PylRS(Y384F) are 0.18 and 0.13 min^{-1}, respectively. According to the tRNA-docking model, it is difficult for the terminal adenosine of tRNA to access the pyrrolysyladenylate bound within the closed amino acid binding tunnel of the PylRS active site (Yanagisawa et al., 2008). We imagine that the Y384F mutation affects the opening/closing of the amino acid binding tunnel, so that $tRNA^{Pyl}$ could access the enzyme-bound pyrrolysyladenylate, and thus the mutation could increase the aminoacylation activity of PylRS. All of these results indicate that the hydrogen bonds between Tyr384 and the pyrrolysyl moiety of pyrrolysyl-AMP are not essential for the PylRS activity. In contrast to the PylRS(Y384F) mutant, the PylRS(Y384A) mutation destroyed the in vivo suppression activity for BocLys (Figure 4D). The hydroxyl group of Tyr384 is not necessary, but the aromatic ring at position 384 is required for the in vitro PylRS activity.

Structure-Based Mutation of PylRS, and Site-Specific Incorporation of ZLys into a Protein

Intriguingly, one of the PylRS(c270) active-site mutants, Y306A, can esterify $tRNA^{Pyl}$ with ZLys, which is the largest lysine derivative tested in this study, much more efficiently than the parent PylRS(c270) (Figures 5A, 5B, and 5C). The mutation of Tyr306 to Ala306 generated an appropriate cavity for accommodating the benzyloxycarbonyl group (Figures 5A and 5B). The PylRS(c270)Y306A mutant exhibits no aminoacylation activity for ZLys derivatives, such as N^{ε}-Z-D-lysine and N^{ε}-Z-L-ornithine (unpublished data). Therefore, the stereochemistry of the Cα carbon, the length of the alkyl chain, and the location of the N^{ε}-carbonyl group are important for the amino acid recognition by PylRS. In addition, the PylRS(c270)Y306A mutant showed lower aminoacylation activities for NmaLys and NicLys than those of the wild-type PylRS(c270) (data not shown). The flexibility of the linker between the N^{ε}-carbonyl group and the aromatic ring might be important for the recognition by PylRS.

Mutations were introduced into the amino acid binding tunnel of PylRS, to test a variety of lysine derivatives for their in vivo site-specific incorporation into proteins. Among these mutants, the efficient incorporation of the lysine derivative ZLys into the GST protein was achieved by using the double mutant PylRS(Y306A•Y384F). When the PylRS(Y306A•Y384F) double mutant was coexpressed with $tRNA^{Pyl}$ in the presence of 5 mM ZLys, the full-length GST protein was produced (Figure 5D, upper panel), as confirmed by a western blot analysis (Figure 5D, lower panel). The GST protein was purified with a glutathione-affinity column and yielded 1.97 mg per liter medium, a 6-fold higher amount than that (0.33 mg protein per liter medium) obtained

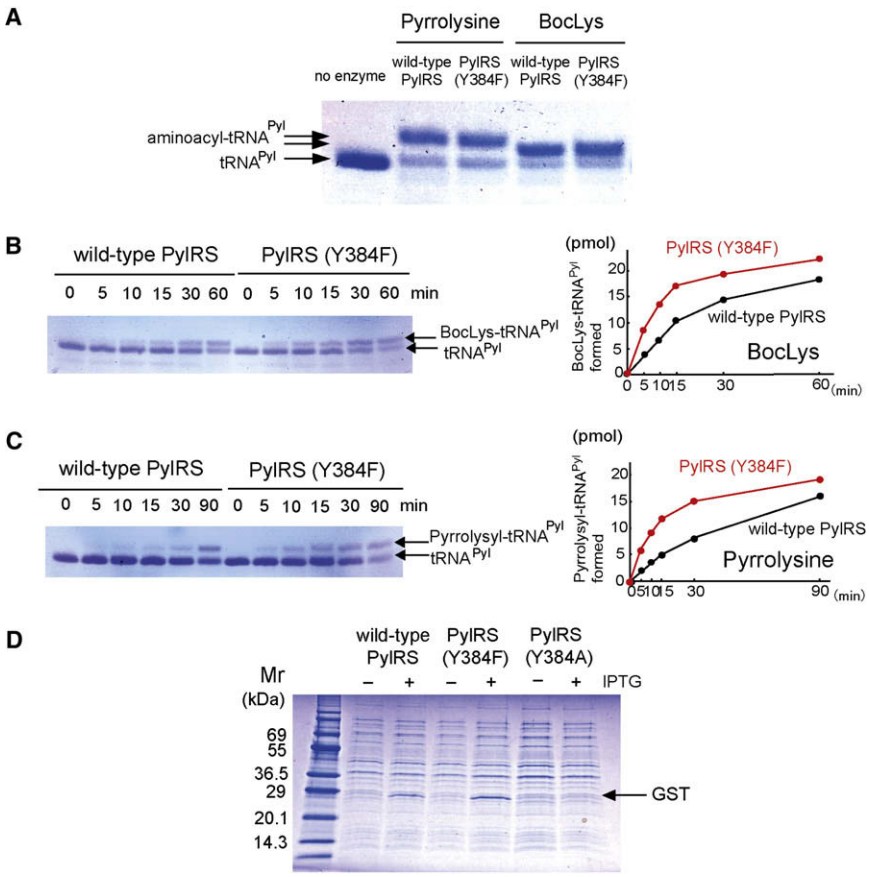

Figure 4. Effect of Tyr384 Mutations on In Vitro and In Vivo Activities

(A) The mutant PylRS(Y384F) yielded the same levels of pyrrolysyl-tRNA^{Pyl} and BocLys-tRNA^{Pyl} as those yielded by the wild-type PylRS, after a long duration (1 hr) under similar conditions to those for the results shown in Figure 1 (Experimental Procedures). Starting from the left, each lane shows a reaction with the following: no enzyme with pyrrolysine (1 mM); the wild-type PylRS with pyrrolysine (1 mM); PylRS(Y384F) with pyrrolysine (1 mM); the wild-type PylRS with BocLys (1 mM); PylRS(Y384F) with BocLys (1 mM).

(B and C) The mutant PylRS(Y384F) esterifies tRNA^{Pyl} with BocLys and pyrrolysine significantly faster than the wild-type PylRS does, at a lower enzyme concentration. The aminoacylation assay conditions are described in Experimental Procedures. (B, left panel) Starting from the left, each lane shows a reaction with the following: the wild-type PylRS with BocLys (0 to 60 min); PylRS(Y384F) with BocLys (0 to 60 min). (B, right panel) Time course analysis of the BocLys-tRNA^{Pyl} catalyzed by wild-type PylRS (black line) and PylRS(Y384F) (red line). The total amounts (pmol) of aminoacyl-tRNAs synthesized in the reaction mixture were calculated from the band intensities, as processed with the NIH imageJ software, and are shown in the graphs. (C, left panel) Starting from the left, each lane shows a reaction with the following: the wild-type PylRS with pyrrolysine (0–90 min); PylRS(Y384F) with pyrrolysine (0–90 min). (C, right panel) Time course analysis of the pyrrolysyl-tRNA^{Pyl} syntheses catalyzed by wild-type PylRS (black line) and PylRS(Y384F) (red line).

(D) BocLys-dependent amber suppression in *E. coli*. When the wild-type and Y384F PylRSs were coexpressed with tRNA^{Pyl} in the presence of BocLys, full-length GST proteins were produced. When the Y384A PylRS was coexpressed with tRNA^{Pyl} in the presence of BocLys, no readthrough products were detected.

with the single mutant PylRS(Y384F) in the presence of ZLys (Figure 5E). The amounts of GST proteins produced by the PylRS(Y384F) and PylRS(Y306A•Y384F) mutants in the absence of added amino acids were 0.2 and 0.3 mg per liter medium, respectively (Figure 5E). The identity of the amino acid inserted in response to the in-frame UAG codon was confirmed by mass spectrometry (obsd: m/z 1,426.70 $[M+H]^+$, calcd: m/z 1,426.80 $[M+H]^+$) (Figure 5F, Figure S3, and Table S1). The molecular masses of these peptides are 98.96 Da larger than that of the wild-type tryptic peptide (Figure 5F).

Incorporation of an Azido Group-Containing Non-natural Amino Acid and Site-Specific Fluorescent Labeling of a Protein

According to the docking model of the PylRS catalytic fragment with the ZLys derivatives, *ortho* substituents on the benzene ring can be accommodated in the PylRS active site tunnel with minor steric hindrance (Figures 6A and 6B). To provide an example of a useful ZLys derivative that can be productively incorporated into a protein by using the PylRS(Y306A•Y384F)/tRNA^{Pyl} system, we examined the in vivo amber suppression activity for N^{ε}-(*o*-azidobenzyloxycarbonyl)-L-lysine (AzZLys) (Figures 6A and 6B). When PylRS(Y306A•Y384F) was expressed in the presence of 1 mM AzZLys, a large amount of the full-length GST protein was produced. The GST protein was purified and yielded 10.46 mg per liter medium, a 5-fold higher amount than that obtained in the presence of ZLys (Figure 6C), and a 7.5-fold higher than the amount of the AzPhe-containing GST (1.4 mg per liter medium) produced by the mutant TyrRS/tRNA^{Tyr} system (Chin et al., 2002). The identity of the amino acid inserted in response to the amber codon was confirmed (Figure 6D). Azido group-containing proteins can be modified posttranslationally with useful functionalities (e.g., fluorescent groups, heavy atoms, reactive groups, and spin labels), and can also be used as photocrosslinkers to explore the interacting protein targets in cells. To demonstrate the utility of azido amino acids, we used selective modification with a fluorescein-triarylphosphine reagent by means of Staudinger ligation (Saxon and Bertozzi, 2000; Kiick et al., 2002), as described in Supplemental Experimental Procedures (Figure 6E, Figure S4). Only the band for the AzZLys-containing GST was fluorescent, whereas no fluorescence was detected from the ZLys-containing GST (Figure 6F). Furthermore, no fluorescence was detected from the unlabeled AzZLys-containing GST (data not shown), showing that the fluorescein-triarylphosphine derivative does not react with any existing side chains from the natural amino acids. As compared with the AzPhe-containing GST, the efficiency of the site-specific fluorescent labeling of the protein was higher at 37°C (Figure 6G).

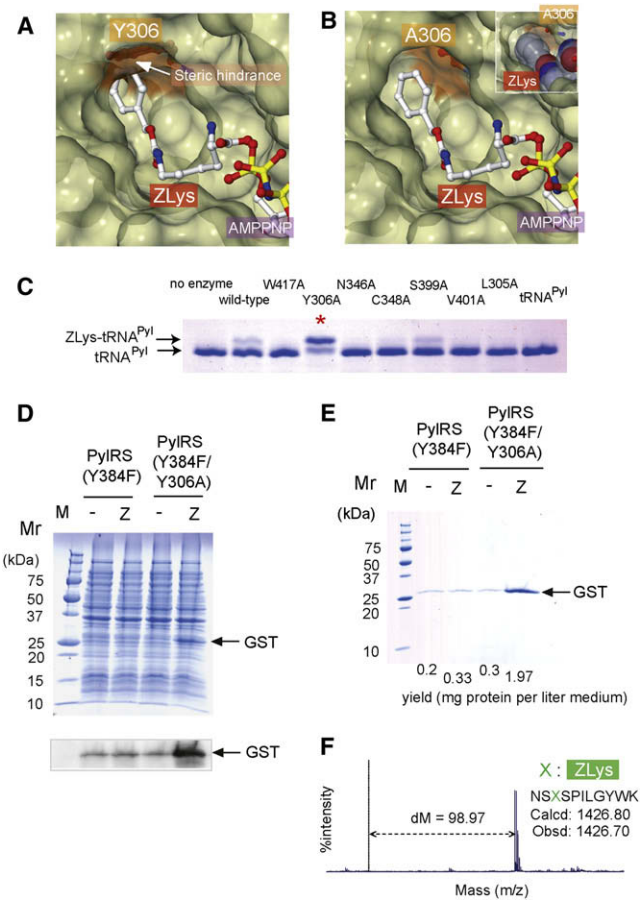

Figure 5. Structure-Based Mutations Affect In Vitro and In Vivo Activities for ZLys

(A and B) Docking models of the PylRS(c270) (A) and PylRS(c270)(Y306A) (B) active sites with ZLys. The substitution of Tyr 306 with Ala reduces the steric hindrance between Tyr306 and the benzyloxycarbonyl group of ZLys. Amino acid residues that would interact with ZLys are represented as ball-and-stick models. The PylRS structures are presented as surface models. In the inset, ZLys docked with PylRS(c270)(Y306A) is also shown as a CPK model. Tyr306 in PylRS(c270) and Ala306 in PylRS(c270)(Y306A) are shown in orange.

(C) The PylRS(c270) mutant Y306A efficiently esterifies $tRNA^{Pyl}$ with ZLys (3.5 mM). A 9 μM concentration of the PylRS(c270)(Y306A) protein was used. The lane with Y306A is highlighted by a red asterisk. Starting from the left, each lane shows a reaction with the following: no enzyme; wild-type PylRS(c270); W417A; Y306A; N346A; C348A; S399A; V401A; L305A; control $tRNA^{Pyl}$.

(D and E) The increased readthrough products were detected by using the PylRS(Y306A•Y384F)•$tRNA^{Pyl}$ system for ZLys (D, upper panel). These products are the GST protein, as confirmed by western blotting (D, lower panel). (E) The purified GST proteins, from in vivo amber suppression by using PylRS(Y384F) and PylRS(Y306A•Y384F), in the absence (–) or presence of ZLys (Z).

(F) In-gel tryptic digests of the purified GST proteins were analyzed by mass spectrometry.

DISCUSSION

In addition to pyrrolysine, a variety of lysine derivatives, from AcLys to ZLys, were esterified to some extent to $tRNA^{Pyl}$ by the wild-type PylRS. Among them, four non-natural lysine derivatives, BocLys, AlocLys, ZLys, and AzZLys, were efficiently incorporated at the targeted site of a protein with the developed mutants of PylRS, and the "alloproteins" were obtained in milligram quantities. Multistep engineering of PylRS enabled the genetic encoding of AzZLys *in E. coli*. The single mutation Y384F of PylRS, obtained from a random mutation library, enhanced the in vitro and in vivo activities for BocLys and pyrrolysine, as compared with the wild-type PylRS, and PylRS(Y384F) might be useful for the site-specific incorporation of a broad range of lysine derivatives into proteins. A single mutation of PylRS, Y306A, obtained by structure-based design, drastically enhanced the in vitro aminoacylation activity for ZLys. The combination of the Y306A and Y384F mutations of PylRS enabled the production of a protein containing ZLys. Furthermore, the amount of the AzZLys-containing protein produced by the PylRS(Y306A•Y384F)/$tRNA^{Pyl}$ system was increased 5-fold, as compared with that with ZLys. AzZLys is a much better substrate of PylRS(Y306A•Y384F) than ZLys because AzZLys snugly occupies the vacant space in the catalytic pocket of PylRS(Y306A•Y384F) (Figure 6B). The in vitro site-specific labeling of the AzZLys-containing protein with a fluorescent probe was readily accomplished, and the labeling efficiency was significantly higher than that of the AzPhe-containing protein produced by the mutant TyrRS/$tRNA^{Tyr}$ system (Chin et al., 2002). The present PylRS(Y306A•Y384F) mutant has a larger catalytic pocket than that of the PylRS(R61K•G131E•L309A•C348V•Y384F) mutant (Mukai et al., 2008), which is specialized for ZLys, rather than for AzZLys (the ZLys- and AzZLys-containing GSTs produced by the previous PylRS mutant are 2.2 and 0.5 mg per liter medium, respectively). Therefore, PylRS(Y306A•Y384F) is more useful than PylRS(R61K•G131E•L309A•C348V•Y384F), and could be expected to esterify $tRNA^{Pyl}$ with bulky ZLys derivatives bearing a wide variety of functionalities at the *ortho* position that are designed to fit within the amino acid binding tunnel.

Together with the present study, the molecular basis of the amino acid recognition by PylRS has been revealed by biochemical (Polycarpo et al., 2006) and crystallographic studies (Kavran et al., 2007; Yanagisawa et al., 2008). PylRS selects substrate amino acids by its large, hydrophobic pocket, to accommodate the bulky, hydrophobic moiety. A bulky, hydrophobic moiety with a suitable size and shape, and an adjacent N^{ε}-carbonyl group, are the most important determinants for amino acid selection. In addition, the N^{ε}-carbonyl-linked hydrophilic moiety is important, because the N^{ε}-carbonyl is rotated and the adjacent oxygen atom forms a hydrogen bond with the side chain of Asn346, as observed in the AlocLys-AMP-bound form. BocLys, ZLys, AzZLys, and pyrrolysine have a hydrophilic oxygen or nitrogen atom adjacent to the N^{ε}-carbonyl group, whereas NicLys and NmaLys do not. NicLys and NmaLys have bulkier aromatic (phenyl or nicotine) groups that are linked directly to the N^{ε}-carbonyl group. In fact, these two derivatives were only inefficiently esterified to $tRNA^{Pyl}$ by PylRS, and therefore failed to be incorporated into proteins (data not shown). *M. barkeri* PylRS activates and esterifies $tRNA^{Pyl}$ with (D-prolyl)-L-lysine, but not (L-prolyl)-L-lysine (Polycarpo et al., 2006). However, N^{ε}-(3-me-2,3-dihydro-pyrrol-2-yl)carbonyl-L-lysine, which is the stereo-isomer of pyrrolysine, can be recognized as well as pyrrolysine (Polycarpo et al., 2006). Our docking model provides the structural bases for these recognition properties of PylRS.

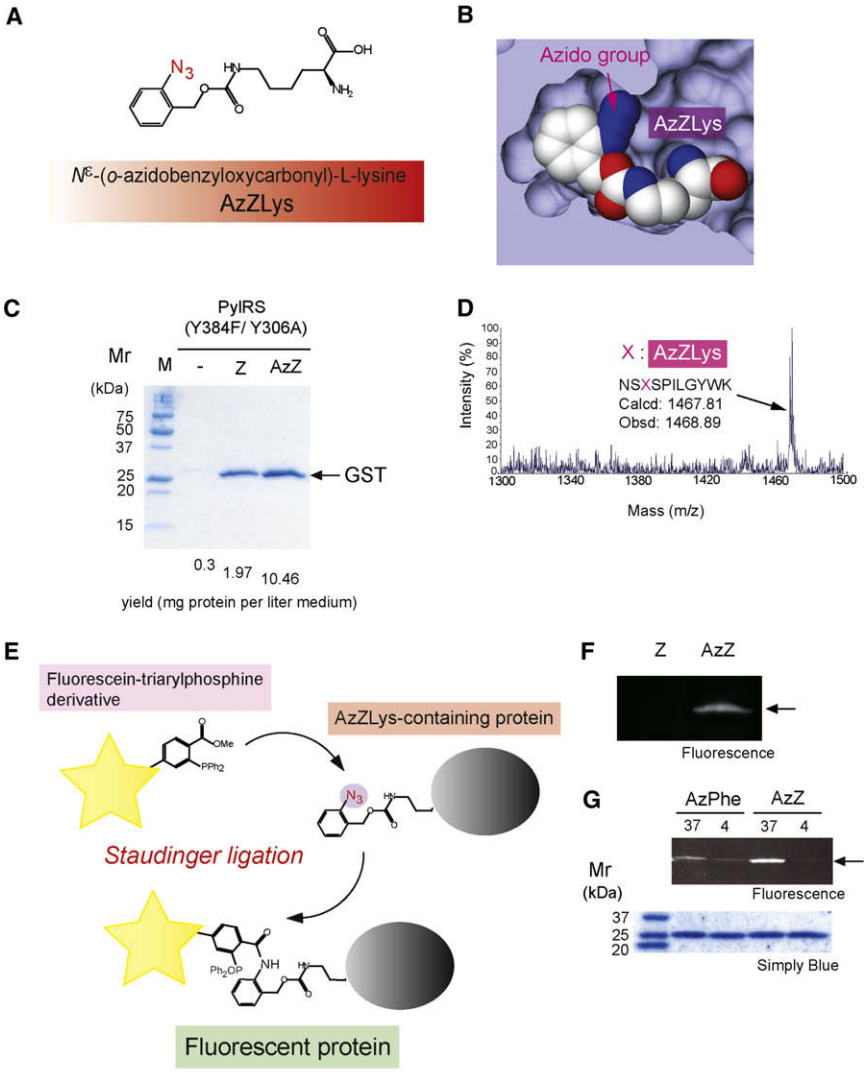

Figure 6. Productive Incorporation of a Useful ZLys Derivative into a Protein and In Vitro Labeling of a Protein with a Fluorescent Probe

(A) Chemical structure of AzZLys.

(B) Docking model of PylRS(c270)(Y306A) with AzZLys. AzZLys (shown as a CPK model) can be accommodated within the catalytic pocket without steric hindrance.

(C) The purified GST proteins, from in vivo amber suppression by using PylRS(Y306A•Y384F), in the absence (–) or presence of 1 mM ZLys (Z) and AzZLys (AzZ).

(D) In-gel tryptic digests of the GST protein were analyzed by mass spectrometry.

(E) Scheme of the fluorescent labeling of a protein containing AzZLys by Staudinger ligation.

(F) Fluorescence was detected with the AzZLys-containing GST (AzZ), whereas no fluorescence was detected with the ZLys-containing GST (Z).

(G) Fluorescent labeling of GSTs containing AzPhe (left two lanes) and AzZLys (right two lanes) at 37°C for 1 hr (37) and at 4°C for overnight (4). Fluorescent scanning (upper panel). Simply Blue staining (lower panel).

In many respects, this engineered PylRS•tRNA^{Pyl} system is suitable for the site-specific genetic encoding of non-natural amino acids. First, PylRS and tRNA^{Pyl} constitute an excellent orthogonal aaRS•tRNA pair (Neumann et al., 2008; Mukai et al., 2008). Second, the specificity of PylRS can be expanded for an unusually broad range of substrate amino acids, and PylRS shows almost no recognition of the tRNA anticodon (Ambrogelly et al., 2007; Yanagisawa et al., 2008). The weak recognition of the tRNA anticodon by PylRS could lead to a tRNA^{Pyl} complementary to a four-base codon (Hohsaka et al., 2001; Magliery et al., 2001). The site-specific incorporation of AcLys into proteins has already been achieved (Neumann et al., 2008; Mukai et al., 2008). The site-specific incorporation of BocLys and ZLys into proteins has been accomplished on a practical level in yeast and mammalian cells (Mukai et al., 2008).

In the present study, we carried out the multistep engineering of PylRS, and successfully achieved the productive incorporation of the azido-containing AzZLys site-specifically into a protein in *E. coli*. Furthermore, the azido group of the produced AzZLys-containing protein was selectively modified with a fluorescent phosphine derivative by Staudinger ligation. In this AzZLys system, the productivity and the fluorescent labeling efficiency were both several-fold higher than those in the AzPhe system, using a mutant TyrRS and an amber suppressor tRNA from *M. jannaschii* (Chin et al., 2002). The remarkably higher performance of AzZLys as a target site for the specific chemical modification is likely due to the longer and more flexible linker between the azido group and the protein backbone, as compared with AzPhe. In addition to Staudinger ligation, the azido group of AzZLys could be utilized for Huisgen [3+2]-cycloaddition. Moreover, the site-specific azido group could be used as a photocrosslinker to find interacting proteins (Krieg et al., 1986; Chen et al., 1994). The long, flexible linker moiety of AzZLys is expected to be advantageous for efficient crosslinking. In addition, in contrast to AzPhe, the products photocrosslinked via the azido group of AzZLys are cleavable in the downstream analysis, because the benzyloxycarbonyl group is readily removed by palladium-catalyzed hydrogenation. The engineered PylRS•tRNA^{Pyl} pair for the genetic encoding of AzZLys might also be suitable for other useful ZLys derivatives, as long as the substituent is at the *ortho* position of the benzene ring. Such ZLys derivatives should retain the advantages conferred by the flexible linker and the cleavability of the Z group. The ZLys-charging mutant of PylRS could be engineered further to fit larger substituents. Thus, we have prepared the technical platform for the genetic encoding of ZLys derivatives with useful functionalities.

SIGNIFICANCE

Pyrrolysine is a lysine derivative with a bulky pyrroline ring. Pyrrolysine is directly esterified to its specific tRNA ($tRNA^{Pyl}$), which has the anticodon complementary to the UAG codon, by pyrrolysyl-tRNA synthetase (PylRS). The wild-type PylRS can esterify $tRNA^{Pyl}$ with several lysine derivatives, and some of them have been translationally incorporated into proteins by the PylRS•$tRNA^{Pyl}$ system. Engineering of the PylRS•$tRNA^{Pyl}$ system has enabled the genetic encoding of acetyl-lysine (AcLys). In this study, we performed a multistep engineering of *Methanosarcina mazei* PylRS, and successfully achieved the genetic encoding of a useful, much larger lysine derivative, N^{ε}-(*o*-azidobenzyloxycarbonyl)-L-lysine (AzZLys), in *Escherichia coli*, as follows. First, with respect to the wild-type PylRS, the recognition mechanisms of N^{ε}-(*tert*-butyloxycarbonyl)-L-lysine (BocLys) and N^{ε}-allyloxycarbonyl-L-lysyladenylate (AlocLys-AMP) were crystallographically and biochemically elucidated, and provided the structural basis for the efficient esterification of these non-natural lysine derivatives to $tRNA^{Pyl}$. Second, the PylRS•$tRNA^{Pyl}$ system was evolved, by random screening, to increase the in vivo productivities of BocLys- and AlocLys-containing proteins. The obtained mutant, PylRS(Y384F), showed enhanced in vitro aminoacylation activities not only for BocLys and AlocLys but also for pyrrolysine. Third, the mutant PylRS(Y306A) was developed, based on the crystal structures, in order to drastically enhance the in vitro aminoacylation activity for N^{ε}-benzyloxycarbonyl-L-lysine (ZLys). Fourth, the double mutant PylRS(Y306A•Y384F), together with $tRNA^{Pyl}$, provided a good yield of the in vivo amber-suppression product containing ZLys. Fifth, the PylRS(Y306A•Y384F)/$tRNA^{Pyl}$ system turned out to be much more efficient for the useful, non-natural amino acid AzZLys. Site-specific labeling of an AzZLys-containing protein with a fluorescent probe was achieved. As compared with *p*-azido-L-phenylalanine (AzPhe), the efficiency of the site-specific modification of AzZLys was significantly higher. The present study has opened the door to the possibility of the efficient and productive incorporation of a wide range of non-natural amino acids, with bulky, useful functionalities, into proteins for structural and functional analyses.

EXPERIMENTAL PROCEDURES

Sample preparation and crystallization, data collection, and structure determination and refinement are described in Supplemental Experimental Procedures.

Aminoacylation Assays

Mutations were generated with a QuikChange mutagenesis kit (Stratagene). The mutant PylRS(c270) proteins were overexpressed and purified with a HisTrap column (Amersham Biosciences). The aminoacylation reactions were performed for 1 hr at 37°C. The standard aminoacylation assay solution (20 μl) contained 2.8–5.2 μM purified *M. mazei* PylRS (or 9 μM PylRS[c270]), 10 mM $MgCl_2$, 2 mM ATP, 4 mM DTT, 2.1 μM *M mazei* $tRNA^{Pyl}$ transcript, and the appropriate amino acid concentrations, in 100 mM Na-HEPES buffer (pH 7.2). The time course of the aminoacylation reaction (20 μl total volume) was measured in the presence of 0.47 μM PylRS, 2.1 μM *M. mazei* $tRNA^{Pyl}$, and 1 mM BocLys or pyrrolysine at 37°C. Unaminoacylated and aminoacylated $tRNA^{Pyl}$s were subjected to electrophoresis on a 10% denaturing polyacrylamide gel under acidic conditions (pH 5.0) for 18 hr at 4°C (Varshney et al., 1991), and were stained with either methylene blue or toluidine blue.

ACCESSION NUMBERS

The atomic coordinates and the structure factors for PylRS(c270)•BocLys•AMPPNP and PylRS(c270)•AlocLys-AMP•PNP have been deposited in the Protein Data Bank under the accession codes 2ZIN and 2ZIO, respectively.

SUPPLEMENTAL DATA

Supplemental Data include Supplemental Discussion, Supplemental Experimental Procedures, four figures, one table, and Supplemental References and can be found with this article online at http://www.chembiol.org/cgi/content/full/15/11/1187/DC1/.

ACKNOWLEDGMENTS

We would like to thank the staff of the beamline BL41XU at SPring-8 (Harima, Japan), as well as the staff of the BL5A and AR-NW12 beamlines at the Photon Factory (Tsukuba, Japan). We also thank Shun-ichi Sekine, Takuhiro Ito (The University of Tokyo), and Toru Sengoku (RIKEN) for assisting with the data collection, as well as for helpful discussions. We are grateful to Ko-ichiro Kodama (The University of Tokyo) for helpful discussions about the chemistry of non-natural amino acids. We would like to thank Yutaka Muto (RIKEN) for help with the NMR measurement. We thank Miwako Asanuma, Ryogo Akasaka, Machiko Yamaguchi-Hirafuji, Nobuko Maoka, Takahito Mukai, Takaho Terada, Mikako Shirouzu, and Hiroshi Hirota (RIKEN) for mass spectrometry analyses. We would also like to thank Tomomi Sumida for technical assistance, and Azusa Ishii, Kiyomi Yajima, and Tomoko Nakayama for clerical assistance. This work was supported in part by Grants-in-Aid for Scientific Research in Priority Areas from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan, the Project for the Development of Fundamental Technology for Protein Analyses, the Targeted Proteins Research Program (TPRP), and the RIKEN Structural Genomics/Proteomics Initiative (RSGI), in the National Project on Protein Structural and Functional Analyses, MEXT. T.K. was supported by the special postdoctoral researcher program of RIKEN. T.Y. expresses his deepest gratitude to the late Emiko Fusatomi for her continuous encouragement and the peace of mind she gave him.

Received: March 7, 2008
Revised: September 4, 2008
Accepted: October 2, 2008
Published: November 21, 2008

REFERENCES

Ambrogelly, A., Gundllapalli, S., Herring, S., Polycarpo, C., Frauer, C., and Söll, D. (2007). Pyrrolysine is not hardwired for cotranslational insertion at UAG codons. Proc. Natl. Acad. Sci. USA *104*, 3141–3146.

Atkins, J.F., and Gesteland, R. (2002). Biochemistry, The 22nd amino acid. Science *296*, 1409–1410.

Blight, S.K., Larue, R.C., Mahapatra, A., Longstaff, D.G., Chang, E., Zhao, G., Kang, P.T., Green-Church, K.B., Chan, M.K., and Krzycki, J.A. (2004). Direct charging of tRNA(CUA) with pyrrolysine *in vitro* and *in vivo*. Nature *431*, 333–335.

Carter, C.W., Jr. (1993). Cognition, mechanism, and evolutionary relationships in aminoacyl-tRNA synthetases. Annu. Rev. Biochem. *62*, 715–748.

Chen, Y., Ebright, Y.W., and Ebright, R.H. (1994). Identification of the target of a transcription activator protein by protein-protein photocrosslinking. Science *265*, 90–92.

Chin, J.W., Santoro, S.W., Martin, A.B., King, D.S., Wang, L., and Schultz, P.G. (2002). Addition of *p*-azido-L-phenylalanine to the genetic code of *Escherichia coli*. J. Am. Chem. Soc. *124*, 9026–9027.

Commans, S., and Böck, A. (1999). Selenocysteine inserting tRNAs: an overview. FEMS Microbiol. Rev. *23*, 335–351.

Cusack, S. (1995). Eleven down and nine to go. Nat. Struct. Biol. *2*, 824–831.

Cusack, S., Berthet-Colominas, C., Hartlein, M., Nassar, N., and Leberman, R. (1990). A second class of synthetase structure revealed by X-ray analysis of *Escherichia coli* seryl-tRNA synthetase at 2.5 Å. Nature *347*, 249–255.

Deiters, A., Cropp, T.A., Mukherji, M., Chin, J.W., Anderson, J.C., and Schultz, P.G. (2003). Adding amino acids with novel reactivity to the genetic code of *Saccharomyces cerevisiae*. J. Am. Chem. Soc. *125*, 11782–11783.

Desogus, G., Todone, F., Brick, P., and Onesti, S. (2000). Active site of lysyl-tRNA synthetase: structural studies of the adenylation reaction. Biochemistry *39*, 8418–8425.

Eriani, G., Delarue, M., Poch, O., Gangloff, J., and Moras, D. (1990). Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. Nature *347*, 203–206.

Gladyshev, V.N., Jeang, K.T., and Stadtman, T.C. (1996). Selenocysteine, identified as the penultimate C-terminal residue in human T-cell thioredoxin reductase, corresponds to TGA in the human placental gene. Proc. Natl. Acad. Sci. USA *93*, 6146–6151.

Hao, B., Gong, W., Ferguson, T.K., James, C.M., Krzycki, J.A., and Chan, M.K. (2002). A new UAG-encoded residue in the structure of a methanogen methyltransferase. Science *296*, 1462–1466.

Hendrickson, T.L., de Crecy-Lagard, V., and Schimmel, P. (2004). Incorporation of nonnatural amino acids into proteins. Annu. Rev. Biochem. *73*, 147–176.

Hohsaka, T., Ashizuka, Y., Taira, H., Murakami, H., and Sisido, M. (2001). Incorporation of nonnatural amino acids into proteins by using various four-base codons in an *Escherichia coli in vitro* translation system. Biochemistry *40*, 11060–11064.

Ibba, M., and Söll, D. (2000). Aminoacyl-tRNA synthesis. Annu. Rev. Biochem. *69*, 617–650.

Ibba, M., and Söll, D. (2002). Genetic code: introducing pyrrolysine. Curr. Biol. *12*, R464–R466.

Kavran, J.M., Gundllapalli, S., O'Donoghue, P., Englert, M., Söll, D., and Steitz, T.A. (2007). Structure of pyrrolysyl-tRNA synthetase, an archaeal enzyme for genetic code innovation. Proc. Natl. Acad. Sci. USA *104*, 11268–11273.

Kiick, K.L., Saxon, E., Tirrell, D.A., and Bertozzi, C.R. (2002). Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. Proc. Natl. Acad. Sci. USA *99*, 19–24.

Krieg, U.C., Walter, P., and Johnson, A.E. (1986). Photocrosslinking of the signal sequence of nascent preprolactin to the 54-kilodalton polypeptide of the signal recognition particle. Proc. Natl. Acad. Sci. USA *83*, 8604–8608.

Link, A.J., and Tirrell, D.A. (2003). Cell surface labeling of *Escherichia coli* via copper(I)-catalyzed [3+2] cycloaddition. J. Am. Chem. Soc. *125*, 11164–11165.

Magliery, T.J., Anderson, J.C., and Schultz, P.G. (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.

Mukai, T., Kobayashi, T., Hino, N., Yanagisawa, T., Sakamoto, K., and Yokoyama, S. (2008). Adding L-lysine derivatives to the genetic code of mammalian cells with engineered pyrrolysyl-tRNA synthetase. Biochem. Biophys. Res. Commun. *371*, 818–822.

Neumann, H., Peak-Chew, S.Y., and Chin, J.W. (2008). Genetically encoding N^{ε}-acetyllysine in recombinant proteins. Nat. Chem. Biol. *4*, 232–234.

Polycarpo, C., Ambrogelly, A., Berube, A., Winbush, S.M., McCloskey, J.A., Crain, P.F., Wood, J.L., and Söll, D. (2004). An aminoacyl-tRNA synthetase that specifically activates pyrrolysine. Proc. Natl. Acad. Sci. USA *101*, 12450–12454.

Polycarpo, C.R., Herring, S., Berube, A., Wood, J.L., Söll, D., and Ambrogelly, A. (2006). Pyrrolysine analogues as substrates for pyrrolysyl-tRNA synthetase. FEBS Lett. *580*, 6695–6700.

Ruff, M., Krishnaswamy, S., Boeglin, M., Poterszman, A., Mitschler, A., Podjarny, A., Rees, B., Thierry, J.C., and Moras, D. (1991). Class II aminoacyl transfer RNA synthetases: crystal structure of yeast aspartyl-tRNA synthetase complexed with tRNA(Asp). Science *252*, 1682–1689.

Saxon, E., and Bertozzi, C.R. (2000). Cell surface engineering by a modified Staudinger reaction. Science *287*, 2007–2010.

Schimmel, P. (1987). Aminoacyl tRNA synthetases: general scheme of structure-function relationships in the polypeptides and recognition of transfer RNAs. Annu. Rev. Biochem. *56*, 125–158.

Speers, A.E., Adam, G.C., and Cravatt, B.F. (2003). Activity-based protein profiling *in vivo* using a copper(i)-catalyzed azide-alkyne [3 + 2] cycloaddition. J. Am. Chem. Soc. *125*, 4686–4687.

Srinivasan, G., James, C.M., and Krzycki, J.A. (2002). Pyrrolysine encoded by UAG in Archaea: charging of a UAG-decoding specialized tRNA. Science *296*, 1459–1462.

Theobald-Dietrich, A., Frugier, M., Giege, R., and Rudinger-Thirion, J. (2004). Atypical archaeal tRNA pyrrolysine transcript behaves towards EF-Tu as a typical elongator tRNA. Nucleic Acids Res. *32*, 1091–1096.

Tsao, M.L., Tian, F., and Schultz, P.G. (2005). Selective Staudinger modification of proteins containing *p*-azidophenylalanine. ChemBioChem *6*, 2147–2149.

Varshney, U., Lee, C.P., and RajBhandary, U.L. (1991). Direct analysis of aminoacylation levels of tRNAs *in vivo*. Application to studying recognition of *Escherichia coli* initiator tRNA mutants by glutaminyl-tRNA synthetase. J. Biol. Chem. *266*, 24712–24718.

Wang, Q., Chan, T.R., Hilgraf, R., Fokin, V.V., Sharpless, K.B., and Finn, M.G. (2003). Bioconjugation by copper(I)-catalyzed azide-alkyne [3 + 2] cycloaddition. J. Am. Chem. Soc. *125*, 3192–3193.

Wang, L., Xie, J., and Schultz, P.G. (2006). Expanding the genetic code. Annu. Rev. Biophys. Biomol. Struct. *35*, 225–249.

Yanagisawa, T., Ishii, R., Fukunaga, R., Nureki, O., and Yokoyama, S. (2006). Crystallization and preliminary X-ray crystallographic analysis of the catalytic domain of pyrrolysyl-tRNA synthetase from the methanogenic archaeon *Methanosarcina mazei*. Acta Crystallogr. F Struct. Biol. Cryst. Commun. *62*, 1031–1033.

Yanagisawa, T., Ishii, R., Fukunaga, R., Kobayashi, T., Sakamoto, K., and Yokoyama, S. (2008). Crystallographic studies on multiple conformational states of active-site loops in pyrrolysyl-tRNA synthetase. J. Mol. Biol. *378*, 634–652.