J|A|C|S
ARTICLES

Design of a Bacterial Host for Site-Specific Incorporation of *p*-Bromophenylalanine into Recombinant Proteins

Inchan Kwon, Pin Wang, and David A. Tirrell*

Contribution from the Division of Chemistry and Chemical Engineering, California Institute of Technology, 1200 East California Boulevard, Pasadena, California 91125

Received April 14, 2006; E-mail: tirrell@caltech.edu

Abstract: Introduction of a yeast suppressor tRNA (ytRNA$^{\text{Phe}}_{\text{CUA}}$) and a mutant yeast phenylalanyl-tRNA synthetase (yPheRS (T415G)) into an *Escherichia coli* expression host allows in vivo incorporation of phenylalanine analogues into recombinant proteins in response to amber stop codons. However, high-fidelity incorporation of non-natural amino acids is precluded in this system by mischarging of ytRNA$^{\text{Phe}}_{\text{CUA}}$ with tryptophan (Trp) and lysine (Lys). Here we show that ytRNA$^{\text{Phe}}_{\text{CUA}}$ and yPheRS can be redesigned to achieve high-fidelity amber codon suppression through delivery of *p*-bromophenylalanine (pBrF). Two strategies were used to reduce misincorporation of Trp and Lys. First, Lys misincorporation was eliminated by disruption of a Watson−Crick base pair between nucleotides 30 and 40 in ytRNA$^{\text{Phe}}_{\text{CUA}}$. Loss of this base pair reduces mischarging by the *E. coli* lysyl-tRNA synthetase. Second, the binding site of yPheRS was redesigned to enhance specificity for pBrF. Specifically, we used the T415A variant, which exhibits 5-fold higher activity toward pBrF as compared to Trp in ATP−PP$_i$ exchange assays. Combining mutant ytRNA$^{\text{Phe}}_{\text{CUA}}$ and yPheRS (T415A) allowed incorporation of pBrF into murine dihydrofolate reductase in response to an amber codon with at least 98% fidelity.

Introduction

Non-natural amino acids carrying a wide variety of novel functional groups have been incorporated into recombinant proteins in both prokaryotic and eukaryotic cells.[1−24] Although global replacement of natural amino acids with non-natural analogues is useful for many purposes,[5,23,24] there are situations in which single-site substitution by a non-natural amino acid is required. In such circumstances, a codon must be assigned uniquely to the non-natural amino acid. Amber (UAG),[7,8,11,12,20,25−27] ochre (UAA),[28,29] and opal (UGA) stop codons[29,30] and four-base codons[30,31] have been explored for this purpose. The amber codon has been used most widely,[7,8,20,25−27] because it is the least common stop codon in *Escherichia coli* and because several naturally occurring suppressor tRNAs recognize it efficiently.[32,33] Use of the amber codon to encode non-natural amino acids requires outfitting the cell with an “orthogonal pair” comprising a suppressor tRNA and a cognate aminoacyl-tRNA synthetase (aaRS) that operate

---

(1) Link, A. J.; Mock, M. L.; Tirrell, D. A. *Curr. Opin. Biotechnol.* **2003**, *14*, 603−609.
(2) Xie, J.; Schultz, P. G. *Curr. Opin. Chem. Biol.* **2005**, *9*, 548−554.
(3) Hendrickson, T. L.; de Crecy-Lagard, V.; Schimmel, P. *Annu. Rev. Biochem.* **2004**, *73*, 147−176.
(4) Budisa, N.; Pal, P. P. *Biol. Chem.* **2004**, *385*, 893−904.
(5) Beatty, K. E.; Xie, F.; Wang, Q.; Tirrell, D. A. *J. Am. Chem. Soc.* **2005**, *127*, 14150−14151.
(6) Link, A. J.; Vink, M. K. S.; Tirrell, D. A. *J. Am. Chem. Soc.* **2004**, *126*, 10598−10602.
(7) Bose, M.; Groff, D.; Xie. J.; Brustad, E.; Schultz, P. G. *J. Am. Chem. Soc.* **2006**, *128*, 388−389.
(8) Deiters, A.; Geierstanger, B. H.; Schultz, P. G. *ChemBioChem* **2005**, *6*, 55−58.
(9) Kiick, K. L.; Weberskirch, R.; Tirrell, D. A. *FEBS Lett.* **2001**, *502*, 25−30.
(10) Dedkova, L. M.; Fahmi, N. E.; Golovine, S. Y.; Hecht, S. M. *J. Am. Chem. Soc.* **2003**, *125*, 6616−6617.
(11) Chin, J. W.; Cropp, T. A.; Anderson, J. C.; Mukherji, M.; Zhang, Z. W.; Schultz, P. G. *Science* **2003**, *301*, 964−967.
(12) Deiters, A.; Cropp, T. A.; Mukherji, M.; Chin, J. W.; Anderson, J. C.; Schultz, P. G. *J. Am. Chem. Soc.* **2003**, *125*, 11782−11783.
(13) Kwon, I.; Kirshenbaum, K.; Tirrell, D. A. *J. Am. Chem. Soc.* **2003**, *125*, 7512−7513.
(14) Zhang, Z. W.; Gildersleeve, J.; Yang, Y. Y.; Xu, R.; Loo, J. A.; Uryu, S.; Wong, C. H.; Schultz, P. G. *Science* **2004**, *303*, 371−373.
(15) Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G. *Science* **2001**, *292*, 498−500.
(16) Wang, L.; Xie, J.; Deniz, A. A.; Schultz, P. G. *J. Org. Chem.* **2003**, *68*, 174−176.
(17) Sharma, N.; Furter, R.; Kast, P.; Tirrell, D. A. *FEBS Lett.* **2000**, *467*, 37−40.
(18) Ibba, M.; Stathopoulos, C.; Soll, D. *Curr. Biol.* **2001**, *11*, R563−R565.
(19) Kirshenbaum, K.; Carrico, I. S.; Tirrell, D. A. *ChemBioChem* **2002**, *3*, 235−237.
(20) Furter, R. *Protein Sci.* **1998**, *7*, 419−426.
(21) Sakamoto, K.; Hayashi, A.; Sakamoto, A.; Kiga, D.; Nakayama, H.; Soma, A.; Kobayashi, T.; Kitabatake, M.; Takio, K.; Saito, K.; Shirouzu, M.; Hirao, I.; Yokoyama, S. *Nucleic Acids Res.* **2002**, *30*, 4692−4699.
(22) Datta, D.; Wang, P.; Carrico, I. S.; Mayo, S. L.; Tirrell, D. A. *J. Am. Chem. Soc.* **2002**, *124*, 5652−5653.
(23) Tang, Y.; Ghirlanda, G.; Petka, W. A.; Nakajima, T.; DeGrado, W. F.; Tirrell, D. A. *Angew. Chem., Int. Ed.* **2001**, *40*, 1494−1496.
(24) Link, A. J.; Tirrell, D. A. *J. Am. Chem. Soc.* **2003**, *125*, 11164−11165.
(25) Cload, S. T.; Liu, D. R.; Froland, W. A.; Schultz, P. G. *Chem. Biol.* **1996**, *3*, 1033−1038.
(26) Cornish, V. W.; Benson, D. R.; Altenbach, C. A.; Hideg, K.; Hubbell, W. L.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 2910−2914.
(27) Mendel, D.; Cornish, V. W.; Schultz, P. G. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 435−462.
(28) Kohrer, C.; Yoo, J.; Bennett, M.; Schaack, J.; RajBhandary, U. L. *Chem. Biol.* **2003**, *10*, 1095−1102.
(29) Kohrer, C.; Sullivan, E. L.; RajBhandary, U. L. *Nucleic Acids Res.* **2004**, *32*, 6200−6211.
(30) Anderson, J. C.; Schultz, P. G. *Biochemistry* **2003**, *42*, 9598−9608.
(31) Hohsaka, T.; Muranaka, N.; Komiyama, C.; Matsui, K.; Takaura, S.; Abe, R.; Murakami, H.; Sisido, M. *FEBS Lett.* **2004**, *560*, 173−177.
(32) Yoshimura, M.; Inokuchi, H.; Ozeki, H. *J. Mol. Biol.* **1984**, *177*, 627−644.
(33) Thorbjarnardottir, S.; Uemura, H.; Dingermann, T.; Rafnar, T.; Thorsteinsdottir, S.; Soll, D.; Eggertsson, G. *J. Bacteriol.* **1985**, *161*, 207−211.

10.1021/ja0626281 CCC: $33.50

independently of the endogenous synthetase-tRNA pairs in *E. coli*. The site-specific incorporation of a non-natural amino acid into recombinant proteins via this strategy was reported by Furter in 1998,[20] and the Schultz laboratory has developed powerful selection methods to identify heterologous synthetases and tRNAs.[7,16,30,34,35] The tyrosyl pair derived from the archaebacterium *Methanococcus jannaschii* has been especially useful in this regard.[36]

In a complementary approach, rational modification and virtual screening methods have been used to design the amino acid binding pocket of the *E. coli* PheRS (ePheRS).[22,37] On the basis of the crystal structure of the PheRS (tPheRS) from *Thermus thermophilus*, Safro and colleagues proposed that Val 261 and Ala 314 in the amino acid binding pocket are critical in the discrimination of Phe from its amino acid competitors.[38] Sequence alignment indicates that Ala 314 in tPheRS corresponds to Ala 294 in ePheRS, and the Hennecke group showed that the substrate specificity of ePheRS can be relaxed by a point mutation at Ala 294. The A294G mutant was shown to enable incorporation of *p*-chlorophenylalanine into recombinant proteins.[39] A subsequent computational simulation, consistent with the Safro prediction, identified two cavity-forming mutations (T251G and A294G) in the ePheRS binding pocket. These two mutations led to relaxed substrate specificity and efficient in vivo replacement of Phe by *p*-acetylphenylalanine (pAcF).[22] Sequence alignment shows that Thr 415 in yPheRS is equivalent to Thr 251 in ePheRS. We therefore anticipated that the yPheRS (T415G) variant would activate a variety of Phe analogues.[40] However, when the T415G variant was co-transformed with ytRNA$^{Phe}_{CUA}$ into an *E. coli* host, substantial misincorporation of Trp and Lys was observed. Use of a triple auxotrophic *E. coli* host (Phe, Trp, and Lys) led to good incorporation (85−95%) of a variety of novel aromatic amino acids in response to amber codons,[40] but the limited specificity of this expression system prompted us to consider its refinement.

In this report, we describe the design of a bacterial host capable of high-fidelity site-specific incorporation of *p*-bromophenylalanine (pBrF) in response to an amber codon. Introduction of aryl halides, such as pBrF[16,17] or *p*-iodophenylalanine (pIF),[16,19] into recombinant proteins allows site-specific modification via versatile palladium-catalyzed cross-coupling reactions with terminal alkene or alkyne reaction partners.[41−46] In order to achieve high-fidelity incorporation of pBrF, two different strategies were applied to reduce misincorporation of Trp and Lys, while providing good yields of recombinant protein. First, Lys misincorporation was eliminated by modification of the sequence of ytRNA$^{Phe}_{CUA}$ to reduce mischarging of ytRNA$^{Phe}_{CUA}$ by *E. coli* lysyl-tRNA synthetase (eLysRS). Second, the binding site of yPheRS was redesigned to enhance specificity for pBrF. Finally, the combined use of these two strategies provided high-fidelity amber codon suppression through delivery of pBrF.

Materials and Methods

Materials. Unless otherwise noted, reagents were obtained from commercial suppliers and were used without further purification. Natural amino acids were from Sigma (St. Louis, MO), and [^{3}H]-amino acids from Amersham Pharmacia Biotech (Piscataway, NJ). pBrF was obtained from Chem-Impex (Wood Dale, IL). Purified total yeast tRNA was purchased from Roche Biochemical (Indianapolis, IN). The nickel−nitrilotriacetic acid (Ni−NTA) affinity column and repressor plasmid pREP4 were from Qiagen (Valencia, CA). *E. coli* strain BLR (*RecA*$^-$ derivative of BL21) was purchased from Novagen (Madison, WI).

Plasmid Construction for Synthetase Expression. Genes encoding the α- and β-subunits of yPheRS were amplified from template plasmid pUC-ASab2[20,47] and inserted between the *Bam*HI and *Kpn*I sites of pQE32 to give pQE32-yPheRS. pQE32-T415G and pQE32-T415A were constructed from pQE32-yPheRS by polymerase chain reaction (PCR) mutagenesis at position 415 of the α-subunit of yPheRS by use of a QuikChange mutagenesis kit (Stratagene). Two complementary oligonucleotides, 5′-CT ACC TAC AAT CCT TAC GCC GAG CCA TCA ATG GAA ATC-3′ for the forward primer and 5′-GAT TTC CAT TGA TGG CTC GGC GTA AGG ATT GTA GGT AG-3′ for the reverse primer, were used to introduce the T-to-A mutation at position 415 of the α-subunit of yPheRS. The entire yPheRS gene was verified by DNA sequencing for each of these constructs. Proofreading polymerase *Pfx* was used in all PCR described in this work. The *E. coli lysS* gene was amplified by PCR from template plasmid pXLysKS1[48] by using two primers, 5′-GCA CTG ACC ATG GCT GAA CAA CAC GCA CAG-3 (which includes an *Nco*I restriction site) and 5′-GGA CTT CGG ATC CTT TCT GTG GGC GCA TCG C-3′ (which carries a *Bam*HI restriction site). The resulting DNA was inserted between the *Nco*I and *Bam*HI sites of pQE60 to yield pQE60-eLysS. The cloned enzymes contained N-terminal or C-terminal hexa-histidine tags to facilitate protein purification.

Synthetase Expression and Purification. The plasmids pQE32-yPheRS, pQE32-T415G, pQE32-T415A, and pQE60-eLysS were individually co-transformed with repressor plasmid pREP4 into *E. coli* strain BLR to form expression strains BLR(pQE32-yFRS), BLR-(pQE32-T415G), BLR(pQE32-T415A), and BLR(pQE60-eLysS), respectively. Synthetase expression was conducted in 2xYT media with 200 μg/mL of ampicillin and 35 μg/mL of kanamycin. At an OD of 0.6, expression of each of the synthetase variants was induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG). After 4 h, cells were harvested and proteins were purified over a nickel−nitrilotriacetic acid affinity column under native conditions according to the manufacturer's protocol (Qiagen). The imidazole in the elution buffer was removed by a desalting column (Amersham Lifescience), and proteins were eluted into a buffer containing 50 mM Tris-HCl (pH = 7.5) and 1 mM dithiothreitol (DTT). Aliquots of proteins were stored in −80 °C in 50% glycerol. Concentrations of the yPheRS variants and eLysRS, the *lysS* gene product, were determined by UV absorbance at 280 nm using calculated extinction coefficients[49] of 99 060 and 29 420 cm^{-1} M^{-1} for yPheRS and eLysRS, respectively.

---

(34) Santoro, S. W.; Wang, L.; Herberich, B.; King, D. S.; Schultz, P. G. *Nat. Biotechnol.* **2002**, *20*, 1044−1048.
(35) Liu, D. R.; Magliery, T. J.; Pasternak, M.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 10092−10097.
(36) Cropp, T. A.; Schultz, P. G. *Trends Genet.* **2004**, *20*, 625−630.
(37) Wang, P.; Vaidehi, N.; Tirrell, D. A.; Goddard, W. A. *J. Am. Chem. Soc.* **2002**, *124*, 14442−14449.
(38) Fishman, R.; Ankilova, V.; Moor, N.; Safro, M. *Acta Crystallogr.* **2001**, *D57*, 1534−1544.
(39) Ibba, M.; Kast, P.; Hennecke, H. *Biochemistry* **1994**, *33*, 7107−7112.
(40) Wang, P. Ph.D. thesis, California Institute of Technology, Pasadena, CA, 2003.
(41) Nguefack, J. F.; Bolitt, V.; Sinou, D. *Tetrahedron Lett.* **1996**, *37*, 5527−5530.
(42) Kodama, K.; Fukuzawa, S.; Nakayama, H.; Kigawa, T.; Sakamoto, K.; Yabuki, T.; Matsuda, N.; Shirouzu, M.; Takio, K.; Tachibana, K.; Yokoyama, S. *ChemBioChem* **2006**, *7*, 134−139.
(43) Schoenbe.A.; Heck, R. F. *J. Org. Chem.* **1974**, *39*, 3327−3331.
(44) Sakakibara, Y.; Ido, Y.; Sasaki, K.; Sakai, M.; Uchino, N. *Bull. Chem. Soc. Jpn.* **1993**, *66*, 2776−2778.
(45) Yoneyama, M.; Kakimoto, M.; Imai, Y. *Macromolecules* **1988**, *21*, 1908−1911.
(46) Yoneyama, M.; Konishi, T.; Kakimoto, M. A.; Imai, Y. *Makromol. Chem., Rapid Commun.* **1990**, *11*, 381−386.
(47) Sanni, A.; Walter, P.; Ebel, J. P.; Fasiolo, F. *Nucleic Acids Res.* **1990**, *18*, 2087−2092.
(48) Leveque, F.; Plateau, P.; Dessen, P.; Blanquet, S. *Nucleic Acids Res.* **1990**, *18*, 305−312.
(49) Gill, S. C.; von Hippel, P. H. *Anal. Biochem.* **1989**, *182*, 319−326.

Amino Acid Activation Assay. The amino acid-dependent adenosine triphosphate (ATP)−[^{32}P]-pyrophosphate (PP$_i$) exchange reaction was used to determine the kinetics of activation of amino acid analogues by yPheRS. The reaction buffer consisted of 50 mM *N*-(2-hydroxethyl)-piperazine-*N*′-(2-ethanesulfonic acid) (potassium-HEPES) (pH = 6.2), 20 mM MgCl$_2$, 1 mM DTT, 2 mM ATP, and 2 mM PP$_i$ (NEN Life Science) with specific activity of 10−50 Ci/mol. The amino acid concentration varied from 10 μM to 2.5 mM, and the enzyme concentration varied from 10 nM to 100 nM. Aliquots (20 μL) were removed from the reaction solution at various time points and quenched into 500 μL of buffer solution containing 200 mM NaPP$_i$, 7% w/v HClO$_4$, and 3% w/v activated charcoal. The charcoal was sedimented by centrifugation and washed twice with 500 μL of 10 mM NaPP$_i$ and 0.5% HClO$_4$ solution. The [^{32}P]-labeled ATP absorbed on the charcoal was quantified via liquid scintillation methods. The specificity constants were calculated by a nonlinear regression fit of the data to a Michaelis−Menten model.

Plasmid Construction for Yeast tRNA$^{\text{Phe}}$ Expression. The mutant yeast amber suppressor tRNA (ytRNA$^{\text{Phe}}_{\text{CUA}}$) was constitutively expressed under control of an *lpp* promoter. The expression cassette for ytRNA$^{\text{Phe}}_{\text{CUA}}$ was inserted into repressor plasmid pREP4 to form pREP4-ytRNA as described by Furter.[20] A mutant yeast suppressor ytRNA$^{\text{Phe}}_{\text{CUA}}$_30U40G (ytRNA$^{\text{Phe}}_{\text{CUA\_UG}}$) was constructed from ytRNA$^{\text{Phe}}_{\text{CUA}}$ by use of a QuikChange mutagenesis kit. Two complementary oligonucleotides, designated primer UG_f (5′-GAA CAC AGG ACC TCC ACA TTT AGA GTA TGG CGC TCT CCC-3′) for the forward primer and primer UG_r (5′-GGG AGA GCG CCA TAC TCT AAA TGT GGA GGT CCT GTG TTC-3′) for the reverse primer, were used to introduce the desired mutations at positions 30 and 40 of the yeast suppressor tRNA. The resulting plasmid, carrying the gene encoding ytRNA$^{\text{Phe}}_{\text{CUA\_UG}}$, is designated pREP4-ytRNA_UG. In order to construct plasmids for in vitro transcription of ytRNA$^{\text{Phe}}$, the ytRNA$^{\text{Phe}}_{\text{CUA}}$ and ytRNA$^{\text{Phe}}_{\text{CUA\_UG}}$ genes were amplified from pREP4-ytRNA and pREP4-ytRNA_UG, respectively. At the end of the tRNA sequence, a *Bst*NI site was inserted to allow linearization prior to transcription. A T7 promoter sequence was added for in vitro transcription of ytRNA$^{\text{Phe}}$ by T7 RNA polymerase. The following primers were used for PCR: 5′-CTG GGT AAG CTT CGC TAA GGA TCT GCC CTG GTG CGA ACT CTG-3′ (which includes restriction sites for *Hin*dIII and *Bst*NI) and 5′-GAT TAC GGA TTC CTA ATA CGA CTC ACT ATA GCG GAC TTA GCT C-3′ (which carries an *Eco*RI restriction site and a T7 promoter sequence). The resulting DNA fragments were introduced between the *Hin*dIII and *Eco*RI sites of pUC18 to yield pUC18-ytRNA$^{\text{Phe}}_{\text{CUA}}$ and pUC18-ytRNA$^{\text{Phe}}_{\text{CUA\_UG}}$, respectively. In order to facilitate DNA manipulation, a *Bst*NI site close to the T7 promoter sequence of pUC18-ytRNA$^{\text{Phe}}_{\text{CUA}}$ was removed to increase the size of the DNA fragment containing the ytRNA$^{\text{Phe}}_{\text{CUA}}$ gene from 180 bp to 500 bp after *Bst*NI digestion. Two complementary oligonucleotides, 5′-CGG AAG CAG AAA GTG TAA AGA GCG GGG TGC CTA ATG AGT G-3′ for the forward primer and 5′-CAC TCA TTA GGC ACC CCG CTC TTT ACA CTT TAT GCT TCC G-3′ for the reverse primer, were used to introduce this mutation.

In Vitro Transcription. Linearized DNAs were prepared by *Bst*NI digestion of pUC18-ytRNA$^{\text{Phe}}_{\text{CUA}}$ and pUC18-ytRNA$^{\text{Phe}}_{\text{CUA\_UG}}$ as described previously.[50] In vitro transcription of linearized DNA templates and purification of transcripts were performed as described previously with minor alterations.[51] The in vitro transcription of linearized DNA to produce 76-mer tRNA transcripts was performed with the Ambion T7-MEGAshortscript kit. Transcripts were isolated by extraction with 25:24:1 phenol/CHCl$_3$/isoamyl alcohol. The organic layer was re-extracted with water, and a 24:1 CHCl$_3$/isoamyl alcohol extraction was performed on the aqueous layers. The water layer was then mixed with an equal volume of isopropanol, precipitated overnight at −20 °C, pelleted, dried, and redissolved in water. Unreacted nucleotides in the tRNA solution were eliminated using CHROMA SPIN-30 DEPC-H$_2$O (BD Bioscience) spin columns. Concentrations of the transcripts were determined by UV absorbance at 260 nm.

Aminoacylation Assay. Aminoacylation of wild-type ytRNA$^{\text{Phe}}_{\text{GAA}}$ with Phe and Trp by yPheRS variants was performed as described earlier.[50] Aminoacylation reactions were carried out in buffer (100 μL) containing 30 mM HEPES (pH = 7.45), 15 mM MgCl$_2$, 4 mM DTT, 25 mM KCl, and 2 mM ATP at 30 °C. Purified yeast total tRNA was used in the assay at a final concentration of 4 mg/mL (ytRNA$^{\text{Phe}}_{\text{GAA}}$ concentration approximately 2.24 M). For aminoacylation with Phe, 13.3 μM [^{3}H]-Phe (5.3 Ci/mmol) and 80 nM yPheRS variants were used; for aminoacylation with Trp, 3.3 μM [^{3}H]-Trp (30.0 Ci/mmol) and 160 nM yPheRS variants were used. Aminoacylation of ytRNA$^{\text{Phe}}$ transcripts with Lys by eLysRS was carried out in buffer (100 μL) containing 100 mM potassium-HEPES (pH = 7.4), 10 mM MgCl$_2$, 1 mM DTT, 0.2 mM ethylenediaminetetraacetic acid, 2 mM ATP, and 4 units/mL yeast inorganic pyrophosphatase (Sigma) at 37 °C. Concentrations of 4 μM of ytRNA$^{\text{Phe}}$ transcript, 1.1 μM [^{3}H]-Lys (91 Ci/mmol), and 80 nM eLysRS were used. The tRNAs were annealed before use by heating to 85 °C for 4 min in annealing buffer (60 mM Tris, pH = 7.8, 2 mM MgCl$_2$), followed by slow cooling to room temperature. Reactions were initiated by adding the enzyme, and 10 μL aliquots were quenched by spotting on Whatman filter disks soaked with 5% trichloroacetic acid (TCA). The filters were washed for three 10-min periods in ice-cold 5% TCA, washed in ice-cold 95% ethanol, and counted via liquid scintillation methods.

Strain and Plasmid Construction for in Vivo Incorporation of pBrF. A Phe/Trp double auxotrophic strain (AFW) and a Phe/Trp/Lys triple auxotrophic strain (AFWK) were constructed from the Phe auxotrophic strain AF[20] (K10, Hfr(Cavalli) *pheS13rel*-1 *tonA*22 thi *T2*$^{\text{R}}$ *pheA*18, *trpB*114) by P1 phage-mediated transposon transduction as described previously.[40] pQE16 (Qiagen) was chosen as the expression plasmid; pQE16 bears a gene encoding the marker protein murine dihydrofolate reducase (mDHFR) outfitted with a C-terminal hexahistidine tag under control of a bacteriophage T5 promoter and t$_0$ terminator. A QuikChange mutagenesis kit with two complementary oligonucleotides (5′-CCG CTC AGG AAC GAG TAG AAGTAC TTC CAA AGA ATG-3′; 5′-CAT TCT TTG GAA GTA CTT CTA CTC GTT CCT GAG CGG-3′) was used to place an amber codon (TAG) at the 38th position of mDHFR; the resulting plasmid was designated pQE16am. The mutant yPheRS genes T415G and T415A were amplified from pQE32-T415G and pQE32-T415A, respectively, and a constitutive *tac* promoter with an abolished *lac* repressor binding site was added upstream of the start codon of this gene.[20] The resulting expression cassettes were inserted into the *Pvu*II site of pQE16 to yield pQE16am-T415G and pQE16am-T415A.

In Vivo Incorporation of pBrF. The auxotrophic strains AF, AFW, and AFWK were transformed with pQE16am and pREP4-ytRNA vectors to enable pBrF incorporation. mDHFR expression was investigated in 20 mL cultures. The *E. coli* expression strains were grown in M9 minimal medium supplemented with glucose, thiamin, MgSO$_4$, CaCl$_2$, 20 amino acids (at 25 μg/mL), and antibiotics (35 μg/mL of kanamycin and 200 μg/mL of ampicillin). When the cultures reached an OD$_{600}$ of 0.8−1.0, cells were sedimented by centrifugation, washed twice with cold 0.9% NaCl, and shifted to expression media supplemented with 17 amino acids (at 25 μg/mL), 6 mM pBrF, and the indicated concentrations of Phe, Trp, and Lys. Protein expression was induced by addition of 1 mM IPTG. After 4 h, cells were pelleted by centrifugation, and the protein was purified on a Ni-NTA spin column under denaturing conditions according to the supplier's instructions (Qiagen). After purification, expression levels of mDHFR were determined by UV absorbance at 280 nm using a calculated extinction coefficient[49] of 24 750 cm^{-1} M^{-1}.

---

(50) Sampson, J. R.; Uhlenbeck, O. C. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 1033−1037.

(51) Nowak, M. W.; Gallivan, J. P.; Silverman, S. K.; Labarca, C. G.; Dougherty, D. A.; Lester, H. A. *Method Enzymol.* **1998**, *293*, 504−529.

Quantitative Analysis of pBrF Incorporation by Liquid Chromatography Mass Spectrometry (LC-MS). LC-MS and LC-MS/MS analyses of tryptic digests of mDHFR were conducted on a Finnigan LCQ ion trap mass spectrometer equipped with an HPLC pump and ESI probe. Mutant mDHFR was prepared in elution buffer (8 M urea, 100 mM NaH_2PO_4, 10 mM Tris, pH = 4.5). After concentration of the protein by ultrafiltration (Millipore), 10 μL of the concentrate was diluted into 90 μL of 75 mM NH_4HCO_3 for trypsin digestion. Modified trypsin (1 μL, Promega, 0.2 μg/ μL) was added. The sample was incubated at 37 °C for 2−6 h, and the reaction was stopped by addition of 12 μL of 5% trifluoroacetic acid (TFA) solution. Digested peptide solution was subjected to desalting on a C_{18} Vydac Microspin column (The Nest group) and eluted with 50 μL of 80% acetonitrile/20% formic acid (0.1% w/v). The peptide solution eluted from the Microspin column was dried, redissolved in 10% acetonitrile/90% TFA (0.1% w/v) solution, and injected into the HPLC. Peptides were separated on a Magic C_{18} column (Michrom, 300 Å, 0.3 × 150 mm) and eluted at a flow rate of 30 μL/min using a gradient of 10−95% of solvent A (90% acetonitrile/10% 0.1 M aqueous acetic acid solution) and solvent B (2% acetonitrile/98% 0.1 M aqueous acetic acid solution) for 30 min. The column eluent was introduced to the electrospray source, and sequencing was carried out by fragmentation of the precursor ion corresponding to the fragment bearing the residue at position 38 of mutant mDHFR.

Results and Discussion

Aminoacylation of ytRNA$^{\text{Phe}}_{\text{CUA}}$ and ytRNA$^{\text{Phe}}_{\text{CUA\_UG}}$ with Lys. When the yeast phenylalanine amber suppressor (ytRNA$^{\text{Phe}}_{\text{CUA\_WT}}$) was co-expressed with wild-type yPheRS in Furter's *E. coli* expression strain, 60% of the amber codon sites were decoded as Lys.[20] Use of the G37A mutant of ytRNA$^{\text{Phe}}_{\text{CUA\_WT}}$ (ytRNA$^{\text{Phe}}_{\text{CUA}}$) reduced the extent of Lys misincorporation to 5%.[20] Because insertion of Lys in response to the amber codon is likely a consequence of charging of ytRNA$^{\text{Phe}}_{\text{CUA}}$ by the *E. coli* lysyl-tRNA synthetase (eLysRS), we modified ytRNA$^{\text{Phe}}_{\text{CUA}}$ to eliminate cross-reactivity with eLysRS. We focused on the base pair between nucleotides 30 and 40, which is thought to enhance recognition of yeast isoleucine amber suppressor (ytRNA$^{\text{Ile}}_{\text{CUA}}$) by eLysRS.[52] Buttcher and co-workers showed that introduction of a wobble base pair between nucleotides 30 and 40 reduced charging of ytRNA$^{\text{Ile}}_{\text{CUA}}$ by eLysRS.[52] Similarly, we generated a mutant yeast phenylalanine amber suppressor (ytRNA$^{\text{Phe}}_{\text{CUA\_UG}}$) containing the wobble base pair 30U-40G and compared the rates of charging of ytRNA$^{\text{Phe}}_{\text{CUA}}$ and ytRNA$^{\text{Phe}}_{\text{CUA\_UG}}$ with Lys by eLysRS (Figure 1).[53] As expected, the rate of aminoacylation of ytRNA$^{\text{Phe}}_{\text{CUA\_UG}}$ was about 3-fold lower than that of ytRNA$^{\text{Phe}}_{\text{CUA}}$ (Figure 1); however, there was only about a 40% reduction in the rate of aminoacylation of ytRNAPhe$_{\text{CUA\_UG}}$ with Phe as compared to that of ytRNA$^{\text{Phe}}_{\text{CUA}}$. Since ytRNA$^{\text{Phe}}_{\text{CUA}}$ variants are competitively aminoacylated by either eLysRS or yPheRS, the 3-fold reduction in the rate of aminoacylation by eLysRS was expected to reduce the level of Lys-charged ytRNA$^{\text{Phe}}$.

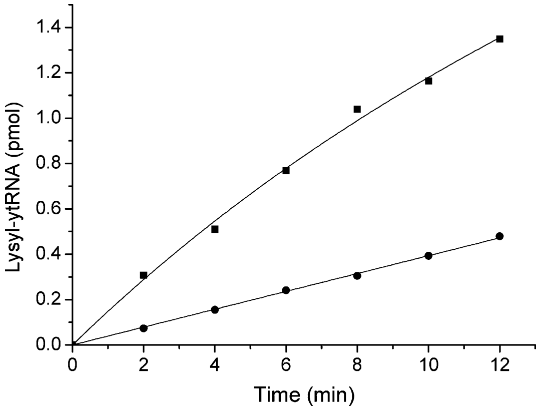

Figure 1. Aminoacylation of ytRNA$^{\text{Phe}}_{\text{CUA}}$ (■) and ytRNA$^{\text{Phe}}_{\text{CUA\_UG}}$ (●) with Lys by eLysRS. The reaction was carried out at 37 °C in mixtures containing 2 mM ATP, 1.1 μM [^{3}H]-Lys, 80 nM eLysRS, and 4 μM ytRNA$^{\text{Phe}}_{\text{CUA}}$ or ytRNA$^{\text{Phe}}_{\text{CUA\_UG}}$.

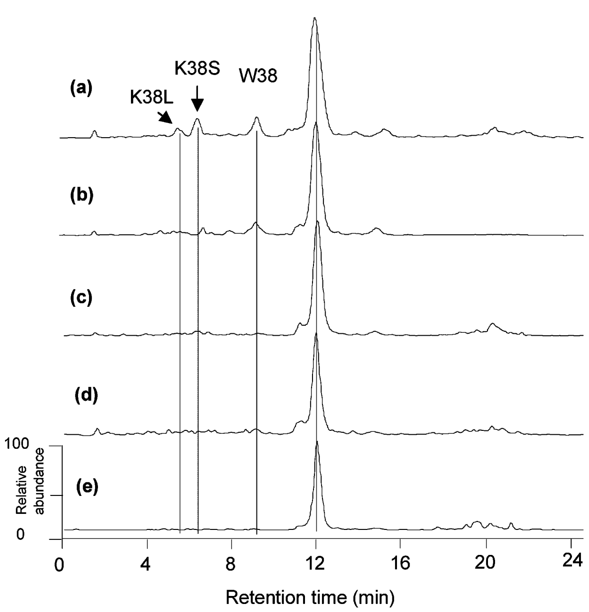

Figure 2. LC-MS chromatograms of tryptic digests of mDHFR. Peptide 38 (residues 26−39; NGDLPWPPLRNEAmK; Am indicates an amber codon) contains an amber codon at the 38th position. Peptide 38 variants containing Lys (peptides K38S and K38L), Trp (peptide W38), and pBrF (peptide Z38) were separated and detected by MS. mDHFRs were synthesized in a triple auxotrophic expression host outfitted with (a) ytRNA$^{\text{Phe}}_{\text{CUA}}$ and yPheRS (T415G); (b) ytRNA$^{\text{Phe}}_{\text{CUA\_UG}}$ and yPheRS (T415G); (c) ytRNA$^{\text{Phe}}_{\text{CUA}}$ and yPheRS (T415A); (d) ytRNA$^{\text{Phe}}_{\text{CUA\_UG}}$ and yPheRS (T415A); or in a single auxotrophic strain with ytRNA$^{\text{Phe}}_{\text{CUA\_UG}}$ and yPheRS (T415A) (e). The expression media were supplemented with 6.0 mM pBrF, 0.01 mM Trp, 1.0 mM Lys, 0.03 mM Phe (a−c) or 0.01 mM Phe (d−f), and 25 μg/mL of 17 amino acids.

Occupancy of the Amber Site. Occupancy of the amber site in mDHFR-38Am was analyzed by liquid chromatography tandem mass spectrometry (LC-MS/MS) (Figure 2). The results are summarized in Table 1. LC-MS/MS has been used successfully for quantitative analysis of protein modification.[42] Mutant mDHFR containing the amber codon at the 38th position (mDHFR−38Am) was expressed in *E. coli* strains AF, AFW, and AFWK co-transformed with pQE16_yPheRS (T415G) and either pREP4_ytRNA$^{\text{Phe}}_{\text{CUA}}$ or pREP4_ytRNA$^{\text{Phe}}_{\text{CUA\_UG}}$. We focused on peptide 38 (residues 26−39; NGDLPWPPLRNEAmK), which contains the amber site. Peptide 38 variants containing Trp and pBrF are designated peptides W38 and Z38, respectively. When Lys is incorporated in response to the amber codon, there are two consecutive lysines at the C-terminus of peptide 38, and the C-terminal Lys can be further cleaved by trypsin. The fully cleaved (NGDLPWPPLRNEK) and partially

---

(52) Buttcher, V.; Senger, B.; Schumacher, S.; Reinbolt, J.; Fasiolo, F. *Biochem. Biophys. Res. Commun.* **1994**, *200*, 370−377.

(53) Nishikawa and co-workers used an alternative strategy to reduce misacylation of yeast amber suppressor tRNA$^{\text{Tyr}}$ by eLysRS. Fukunaga, J.; Yokogawa, T.; Ohno; Nishikawa, K. *J. Biochem.* **2006**, *139*, 689−696.

Table 1. Occupancy of Amber Sites and Expression Yields

			occupancy of amber sites (%)			
host strain	yPheRS	ytRNA^{Phe}	Lys	Trp	pBrF	yield (mg/L)^a
AFWK^b	T415G	CUA^c	7.6	9.5	83	1.4 ± 0.4
AFWK	T415G	CUA	2.7^d	10.2	87	1.5 ± 0.2
AFWK	T415G	CUA_UG^e	ND^f	10.4	90	1.4 ± 0.1
AFWK	T415A	CUA	2.7	ND	97	3.1 ± 0.3
AFWK	T415A	CUA_UG	ND	ND	>98	2.0 ± 0.2
AF^g	T415A	CUA_UG	ND	ND	>98	4.3 ± 0.4

[a] Volumetric yields are given as mg of purified mDHFR−38Am per liter of culture volume. [b] Triple (Phe, Lys, and Trp) auxotrophic *E. coli* strain. [c] The ytRNA$^{Phe}{}_{CUA}$ is designated as CUA. [d] 0.01 mM of Lys was supplemented into the media rather than 1.0 mM. [e] The ytRNA$^{Phe}{}_{CUA\_UG}$ is designated as CUA_UG. [f] Not detected in either LC/MS or MALDI analyses. [g] Single (Phe) auxotrophic *E. coli* strain.

cleaved (NGDLPWPPLRNEKK) variants of peptide K38 were designated K38S (short) and K38L (long), respectively. Peptides K38S, K38L, W38, and Z38 were readily separated by liquid chromatography (Figure 2a), and the relative yields of the four variants were determined by comparing the integrated areas of the corresponding chromatographic signals. Furthermore, the fragment ion masses could be unambiguously assigned (data not shown), confirming the identity of the amino acid inserted in response to the amber codon. The amplitude of the most intense signal in the chromatograms in Figure 2 is ca. 1×10^8 (instrument-dependent arbitrary units), compared to a noise level of ca. 5×10^5, suggesting that misincorporation of natural amino acids can be analyzed at the 0.5% level.

Reduced Lys Misincorporation by ytRNA$^{Phe}{}_{CUA\_UG}$. When mDHFR−38Am was expressed in a triple auxotrophic *E. coli* host (AFWK) outfitted with ytRNA$^{Phe}{}_{CUA}$ and yPheRS (T415G) in medium supplemented with 0.03 mM Phe, 0.01 mM Trp, 1 mM Lys, and 6 mM pBrF, Lys misincorporation was observed at a level of 7.6% (Figure 2a). The level of Lys misincorporation could be reduced to 2.7% by restricting the Lys concentration to 0.01 mM (Table 1). However, when ytRNA$^{Phe}{}_{CUA\_UG}$ was co-transformed into *E. coli* strain AFWK with yPheRS (T415G), Lys misincorporation was not detected, even in media supplemented with 1.0 mM Lys (Figure 2b). Considering the detection limit of the ESI MS detector, this establishes an upper limit of 0.5% Lys misincorporation. Similar results were obtained with Phe auxotrophic strain AF outfitted with ytRNA$^{Phe}{}_{CUA\_UG}$ (Figure 2e); the Lys auxotrophic strain is not required to eliminate Lys misincorporation.

Redesign of Phenylalanyl-tRNA Synthetase. Although use of ytRNA$^{Phe}{}_{CUA\_UG}$ eliminated Lys at the amber site, ca. 10% Trp misincorporation was still detected (Figure 2b). Trp misincorporation could result from mischarging of ytRNA$^{Phe}{}_{CUA}$ with Trp either by yPheRS or by *E. coli* tryptophanyl-tRNA synthetase (eTrpRS). Because misincorporation of Trp was not detected in the absence of yPheRS (T415G), mischarging of ytRNA$^{Phe}{}_{CUA}$ by eTrpRS is probably negligible. However, the Trp-dependent ATP−PP$_i$ exchange rate of yPheRS (T415G) was comparable to that observed for pBrF (Table 2). Therefore, mischarging of ytRNA$^{Phe}{}_{CUA}$ with Trp by yPheRS (T415G) is the more likely explanation for Trp misincorporation at the amber codon. We speculated that yPheRS (T415A) might exclude Trp.

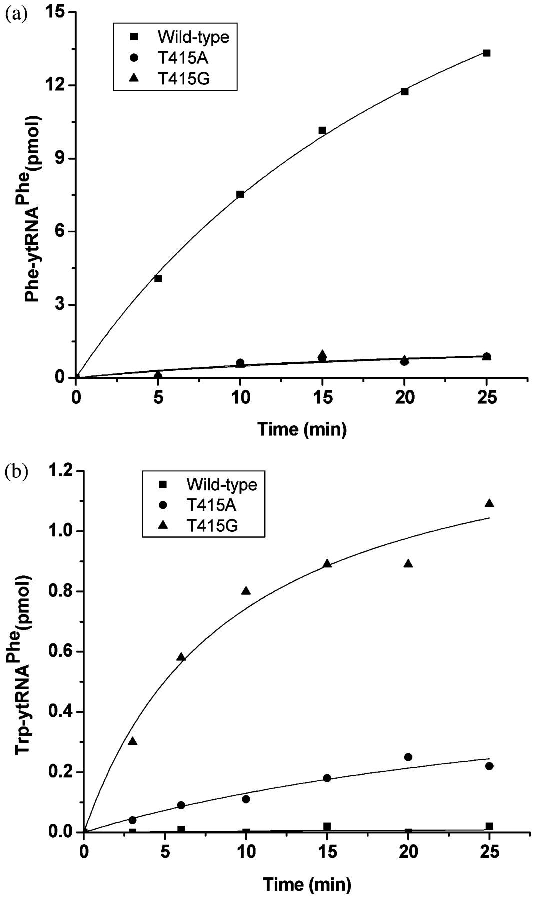

Figure 3. Charging of Phe (a) and Trp (b) by wild-type yPheRS (■), yPheRS (T415G) (▲), and yPheRS (T415A) (●). Reactions were carried out at 30 °C in mixtures containing 2.24 μM ytRNA$^{Phe}{}_{GAA}$, 2 mM ATP, and either (a) 13.3 μm [^{3}H]-Phe and 80 nM yPheRS variants or (b) 3.3 μm [^{3}H]-Trp and 160 nM yPheRS variants in aminoacylation buffer.

ATP−PP$_i$ Exchange and Aminoacylation of ytRNA^{Phe} Catalyzed by yPheRS Variants. The kinetics of amino acid-dependent ATP−PP$_i$ exchange catalyzed by three variants of yPheRS (wild-type, T415G, and T415A) were measured; the

Table 2. Kinetic Parameters for the ATP−PP$_i$ Exchange of Amino Acids by the Wild-Type yPheRS, yPheRS (T415G), and yPheRS (T415A) Variants

amino acid	enzyme	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m (M^{-1} s^{-1})	k_{cat}/K_m (rel)^a
Phe	T415A	185 ± 56	0.207 ± 0.02	1163 ± 240	2.1
Trp	T415A	913 ± 310	0.129 ± 0.01	152.4 ± 49	0.28
pBrF	T415A	255 ± 73	0.217 ± 0.03	892 ± 220	1.6
Phe	T415G	499 ± 51	0.278 ± 0.06	553.0 ± 79	1.0
Trp	T415G	55.0 ± 24	0.261 ± 0.08	5001 ± 1000	9.0
pBrF	T415G	36.3 ± 5.0	0.211 ± 0.09	6116 ± 3400	11
Phe	wild-type	29.2 ± 9.8	0.302 ± 0.13	10669 ± 3700	19
pBrF	wild-type	ND^b	ND	ND	ND

[a] Relative to k_{cat}/K_m for Phe by T415G. [b] Not detected.

kinetic parameters are reported in Table 2. Wild-type yPheRS activates neither Trp nor pBrF, whereas the T415G variant shows 10-fold higher activity toward pBrF and Trp as compared to Phe. As expected, the Trp-dependent ATP−PP_i exchange rate of yPheRS (T415A) is ca. 30-fold lower than that of yPheRS (T415G). In contrast, the effects of the G415A mutation on the rates of activation of Phe and pBrF are much smaller (Table 2). Aminoacylation of $ytRNA^{Phe}_{CUA}$ with Phe was comparable for yPheRS (T415G) and yPheRS (T415A) (Figure 3a); however, aminoacylation of $ytRNA^{Phe}_{CUA}$ with Trp by yPheRS (T415A) is much slower than that by yPheRS (T415G) (Figure 3b).

The improved selectivity of yPheRS (T415A) with respect to exclusion of Trp may be understood, at least in part, on the basis of the side-chain volumes of the amino acid substrates and the active-site residues of the synthetase. Replacement of Gly415 by Ala reduces the volume of the substrate-binding site by ca. 16 Å^3.[54] Occlusion of the binding site reduces the binding affinity of both Trp and pBrF, but because Trp is the larger of the two amino acids (by ca. 5 Å^3), it suffers the larger reduction in affinity. The G415A mutation increases K_m by a factor of ca. 17 for Trp, as compared to ca. 7 for pBrF (Table 2).

High-Fidelity Site-Specific Incorporation of pBrF. Use of expression hosts outfitted with yPheRS (T415A) reduced misincorporation of Trp at the amber site to levels below the limit of detection (Figure 2c−e). Our initial studies used the triple auxotrophic expression host AFWK. However, construction of the triple auxotroph was time-consuming, and the limited cellular pools of several of the natural amino acids reduced protein yield. In order to overcome these problems, a single (Phe) auxotrophic expression host (AF) outfitted with the yPheRS (T415A) and $ytRNA^{Phe}_{CUA\_UG}$ was tested for pBrF incorporation. High-fidelity incorporation of pBrF in response to the amber codon was observed (Figure 2e and Table 1), indicating that the triple auxotrophic expression host is not necessary. Furthermore, the AF expression host outfitted with yPheRS (T415A) and $ytRNA^{Phe}_{CUA\_UG}$ yielded twice as much protein as the triple auxotroph (4.3 mg/L vs 2.0 mg/L) (Table 1).

Conclusion

We describe the engineering of a bacterial expression host that allows high-fidelity incorporation of pBrF into recombinant proteins in response to amber stop codons. Rational modification of the yeast amber suppressor tRNA eliminated cross-reactivity with the *E. coli* LysRS, while the substrate specificity of yPheRS was enhanced by a point mutation in the amino acid binding pocket. When the modified yeast amber suppressor tRNA and the mutant yPheRS were co-expressed in a Phe auxotrophic *E. coli* expression host, at least 98% of the amber sites in full-length recombinant mDHFR were occupied by pBrF. Aryl bromides at programmed positions should enable chemoselective ligation of proteins with terminal alkene or alkyne reaction partners.[42]

Acknowledgment. We thank Dr. P. Plateau for providing the plasmid pXLysKS1 and Dr. M. Shahgholi for assistance with mass spectrometry. We are grateful to Dr. K. Shin, Dr. J. Petersson, and Dr. D. Beene for help with in vitro transcription of tRNA. We thank I. C. Tanrikulu for help with analysis of the PheRS crystal structure. We also thank Dr. A. J. Link and K. E. Beatty for discussion and valuable comments. This work was supported by National Institutes of Health grant GM62523.

JA0626281

---

(54) Amino acid volumes were estimated as solvent-excluded volumes (1.4 Å probe) by using Chem3D Pro 5.0 (CambridgeSoft, Cambridge, MA).