Korean Journal of Microbiology (2024) Vol. 60, No. 1, pp. 1-11
DOI https://doi.org/10.7845/kjm.2024.4009

eISSN 2383-9902
pISSN 0440-2413

A genetic incorporation of *p*-azido-*L*-phenylalanine by variants derived from the *Saccharomyces cerevisiae* tyrosyl-tRNA synthetase in *E. coli*§

Eyob Tekalign# and Jungchan Park*

Department of Bioscience and Biotechnology, Hankuk University of Foreign Studies, Yongin 17035, Republic of Korea
#Present address: National Institute of Health, Bethesda, Maryland, USA

대장균에서 효모균 타이로신-tRNA 합성효소 돌연변이 단백질을 이용한 비천연 아미노산 *p*-azido-*L*-phenylalanine의 위치 특이적 삽입§

이욥테칼린# · 박중찬*

한국외국어대학교 생명공학과, #현재 소속: 미국 국립보건원

(Received February 8, 2024; Revised February 21, 2024; Accepted February 23, 2024)

The incorporation of unnatural amino acids (UAAs) into proteins at specific sites presents a promising avenue for the generation of modified proteins with diverse applications. This study focuses on genetically incorporating *p*-azido-*L*-phenylalanine (AzPhe) into proteins using an orthogonal pair of *Saccharomyces cerevisiae* tyrosyl-tRNA synthetase (*Sc* TyrRS) and a variant of *Escherichia coli* initiator $tRNA_2^{fMet}$ (*fMam*$tRNA_{CUA}$). Employing a rational design approach, three different mutants of *Sc* TyrRS were constructed through site-directed mutagenesis at putative active site residues, guided by analogy with AzPhe-incorporating synthetases derived from *Methanocaldococcus jannaschii* (*Mj*) TyrRS. The three mutants, designated as AzPheRS-1, -2, and -3, exhibited 2–3-fold higher values of chloramphenicol-resistant growth and β-galactosidase activities under AzPheRS-mediated amber suppression in the presence of AzPhe. *In vitro* aminoacylation assay with purified recombinant aminoacyl-tRNA synthetase (aaRS) proteins revealed that the three AzPheRS mutants displayed enhanced aminoacylation activities toward AzPhe compared to tyrosine, with catalytic efficiencies (k_{cat}/K_m) of tyrosylation approximately 10-fold lower than that of the wild-type *Sc* TyrRS. All three mutants successfully incorporated AzPhe into a reporter protein $GFP_{(TAG)}$ in *E. coli*, with AzPheRS-2 and AzPheRS-3 demonstrating higher efficiency than AzPheRS-1. These findings highlight the utility of the orthogonal pair of *Sc* TyrRS and *fMam*$tRNA_{CUA}$ for expanding the genetic code and incorporating UAAs, providing a platform for creating proteins with tailored properties and novel functionalities.

Keywords: *Saccharomyces cerevisiae*, aminoacylation, *p*-azido-phenylalanine, tyrosyl-tRNA synthetase, unnatural amino acids

The expansion of an organism's genetic code to incorporate unnatural amino acids into protein is typically achieved within cells using a pair of orthogonal aminoacyl-tRNA synthetase (aaRS) and its cognate tRNA. The orthogonal synthetase only recognizes its cognate orthogonal tRNA, excluding any other endogenous tRNAs present in the cell. It then selectively transfers the unnatural amino acid to its cognate orthogonal tRNA, which is designed to recognize a nonsense codon, commonly the amber stop codon (Liu and Schultz, 1999). The orthogonal synthetase/tRNA pairs must exhibit minimal cross-reactivity with the host organism's endogenous tRNAs and aaRSs to ensure selectivity.

While various orthogonal aaRS/tRNA pairs derived from yeast, archea, and hybrid forms have been introduced for this

***For correspondence.** E-mail: jpark@hufs.ac.kr;
Tel.: +82-31-330-4355; Fax: +82-31-330-4566
§Supplemental material for this article may be found at http://www.kjom.org/main.html

purpose in prokaryotic cells, the TyrRS/tRNA^{Tyr} pair from *Methanocaldococcus jannaschii* (*Mj*) has been predominantly utilized to enable a system for incorporating genetically unnatural amino acids (Liu and Schultz, 2010; Chin, 2017). Remarkably, numerous variants derived from *Mj* TyrRS have demonstrated successful incorporation of a wide range of unnatural amino acids, including *o*-methyl-*L*-tyrosine, *p*-azido-*L*-phenylalanine, *p*-benzoyl-*L*-phenylalanine, *p*-acetyl-*L*-phenylalanine, and *L*-3-(2-naphthyl)alanine (Wang *et al*., 2001; Chin *et al*., 2002a, 2002b; Wang *et al*., 2002, 2003).

We have used a hybrid pair of *Saccharomyces cerevisiae* (*Sc*) TyrRS and *fMam*tRNA$_{CUA}$, a variant of *E. coli* initiator tRNA$_2^{fMet}$ (Kowal *et al*., 2001; Tekalign *et al*., 2018) to broaden genetic codes for site-specific incorporation of unnatural amino acids in *E. coli*. In our attempt to introduce *p*-azidophenylalanine (AzPhe) using this orthogonal pair, a rational design strategy was employed based on the crystal structures of the TyrRS-tRNA^{Tyr}-tyrosine complexes and the sequence similarity of TyrRS proteins from both *Mj* and *Sc* (Kobayashi *et al*., 2003; Tsunoda *et al.,* 2007). Five residues-Y43, Y127, D177, E178, and L181- constituting the amino acid binding pocket of *Sc* TyrRS, were partially or completely mutated to the corresponding amino acid residues in *p*-azidophenylalanyl-tRNA synthetase (AzPheRS) derived from *Mj* (Chin *et al*., 2002b). Through site-directed mutagenesis, we generated three distinct mutants of *Sc* TyrRS, designated as AzPheRS-1, AzPheRS-2, and AzPheRS-3. Their *in vivo* and *in vitro* enzymatic activities for incorporating AzPhe into proteins were then evaluated. The comprehensive analysis presented herein contributes valuable insights into the rational design of *Sc* TyrRS mutants and expansion of usable orthogonal aaRS/tRNA pairs for site-specific incorporation of unnatural amino acids.

Materials and Methods

Escherichia coli strains, culture media, and plasmids

Escherichia coli strain DH10B (Invitrogen) and a genetically modified strain DH10B (Tn:*lacZ*am) which contains an amber-mutated *lacZ* gene in the DH10B chromosome were used in this

Table 1. Characteristics of bacterial strains and plasmids utilized in this study

Strains/Plasmids	Characteristics	Sources
***Escherichia coli* strains**		
DH10B	F-, *mcr*A, Δ (*mrr-hsd*RMS-*mcr*BC), Φ80d *lacZ*ΔM15, Δ*lac*X74, *deo*R, *rec*A1, *end*A1, *ara*D139, Δ (*ara*, *leu*)7697, *gal*U, *gal*K, λ-*rps*L, *nup*G	Invitrogen
DH10B (Tn:*lacZ*am)	DH10B containing chromosomal integration of *lacZ*am-Km^R cassette	Kim *et al*. (2009)
M15 (pREP4)	Nal^s, Str^s, Rif^s, Thi$^-$, Lac$^-$, Mtl$^-$, F$^-$, [pREP4 Km^R], used for production of aaRS recombinant proteins	Qiagen
Plasmids		
pACamG	CAT gene with an amber mutation at the 27th residue and *E. coli* mutant initiator tRNA$_2^{fMet}$ (*trnfMT2:A71/T35A36/G72*) called *fMam*tRNA$_{CUA}$, pACYC184 frame, Tet^R	Chow and RajBhabdary (1993)
pQE-30	Commercial bacterial expression vector, Amp^R	Qiagen
pQE-YRS	Sc YRS expression vector. Cm^R gene (CAT) was removed from pQE-30 and Sc YRS gene was cloned with N-terminal 6xHis tag	Tekalign *et al*. (2018)
pQE-AzPheRS-1, -2, and -3	Vector expressing AzPheRS proteins that were generated by site-directed mutagenesis with pQE-YRS as template	This study
pACamG-YRS	Vector expressing both *Sc* YRS and *fMam*tRNA$_{CUA}$. *Sc* YRS gene was subclone into pACamG	This study
pACamG-AzPheRS-1, -2, and, -3	Vector expressing both *Sc* AzPheRS mutants and *fMam*tRNA$_{CUA}$. *Sc* AzPheRS genes were subcloned into pACamG	This study
pCNFRS	Vector expressing both AzPhe-incorporating aaRS derived from *Mj* YRS and suppressor tRNA, araBAD promoter, Cm^R.	Miyake-Stoner *et al*. (2010)
pBad-sfGFP-150TAG	Reporter plasmid, superfold GFP gene with an amber mutation at the 150th residue, araBAD promoter, 6xHis tag at the C-terminus, Amp^R	Miyake-Stoner *et al*. (2010)

Abbreviations: CAT, chloramphenicol acetyl transferase; aaRS, aminoacyl-tRNA synthetase; *Sc*, *Saccharomyces cerevisiae*; *Mj*, *Methanocaldococcus jannaschii*; Cm^R, chloramphenicol resistance; Amp^R, ampicillin resistance; Tet^R, tetracycline resistance.

Table 2. Summary of oligonucleotide primers used in this study

Primers	Nucleotide sequence (5' →3')	Usage
SDM-1	CATTTGAAATTAACATGGGGTACCGCG	Site-directed mutagenesis, Y43T
SDM-2	GTTGGCTCTTCTTTCCAGCTAACTCC	Site-directed mutagenesis, Y127F
SDM-3	GCAAGCGTTGCCACTTCAATTCGGTGATGTTGATTGCCAATTTG GTGGTGTC GACCA	Site-directed mutagenesis, D177P, I178L, and L181G
SDM-4	GCAAGCGTTGCCACTTCAATTCCAGGATGTTGATTGCCAATTTG GTGGTGTC GACCA	Site-directed mutagenesis, D177P, I178L, and L181Q
SDM-5	GATAGCGTCAGCAACAATTTTTAGCGCCGCTGGGGACAATTTTTC	Site-directed mutagenesis, P320A and D321A
E-3	CAATCGAGCTCATGTCCTCTGCTGCCACGGTTGAC	Forward primer for generation of pACamG-AzPheRS-1, -2, or -3
E-5	CAAGCCTGCAGTTACAATTTGGTTTCCTCTAGTTTCG	Reverse primer for generation of pACamG-AzPheRS-1, -2, or -3

study (Kim *et al*., 2009). LB broth and GMML media were used for bacterial culture and an assay for amber suppression-mediated chloramphenicol-resistant growth (Edan *et al*., 2014). Auto-induction media previously described (Hammill *et al.* 2007) was used for AzPhe incorporation into the reporter protein $GFP_{(TAG)}$.

The *Sc* TyrRS expression vector, pQE-YRS (Tekalign *et al.,* 2018), and pACamG, encoding a chloramphenicol acetyl transferase (CAT) gene containing an amber mutation at the 27th residue and a suppressor tRNA derived from mutant *E. coli* initiator tRNA gene *trnfMT2:A71/T35A36/G72* (*fMam*$tRNA_{CUA}$), were previously documented (Chow and RajBhandary, 1993). pCNFRS, an expression plasmid of AzPhe-incorporating aaRS derived from *Mj* TyrRS, and pBad-sfGFP-150TAG, a reporter plasmid producing green fluorescence protein (GFP) by amber suppression, were generously provided by Dr. Mehl (Miyake-Stoner *et al.,* 2010). Detailed information on all bacterial strains and plasmids utilized in this study is summarized in Table 1.

Site-directed mutagenesis

Site-directed mutagenesis was conducted using QuickChange Multisite-directed mutagenesis kit (Agilent Technologies), following the manufacturer's protocol. Initially, pQE-YRS served as the template for generating AzPheRS-1, which carried Y43T and Y127F mutations. This was achieved by employing SDM-1 and SDM-2 primers. Subsequently, pQE-AzPheRS-1 was utilized as a template for introducing additional mutations at D177, I178, and L181 residues. Two different primers, SDM-3 and SDM-4, were employed in this step, resulting in the creation of pQE-AzPheRS-2 and pQE-AzPheRS-3, respectively. Detailed information on the primers used in site-directed mutagenesis is summarized in Table 2, while Table 3 provides an overview of the amino acid changes in AzPheRS mutants.

Table 3. Summary of amino acid changes in AzPheRS mutants

aaRS	Number of amino acid residue				
	43	127	177	178	181
Sc TyrRS	Tyr	Tyr	Asp	Glu	Leu
AzPheRS-1	Thr	Phe	Asp	Glu	Leu
AzPheRS-2	Thr	Phe	Pro	Leu	Gly
AzPheRS-3	Thr	Phe	Pro	Leu	Gln

Introduction of P320A and D321A mutations into AzPheRS-3 was achieved using the same method. Using pACamG-AzPheRS-3 as template, Quick-change multisite-directed mutagenesis kit was employed with SDM-5 primer.

Assays for amber suppression activity; Chloramphenicol-resistant bacteria growth and β-galactosidase activity

To test the activity of AzPhe incorporation by AzPheRS mutants, pACamG and an aaRS expression vector (pQE-YRS, or pQE-AzPheRS-1,-2, -3) were cotransformed into DH10B. After overnight seed culture, each transformant was inoculate at 1:100 ratio into fresh GMML media containing 120 μg/ml chloramphenicol with or without 1 mM AzPhe. Optical densities of bacterial growth at 600 nm were measured after 48 h culture at 37°C.

The *in vivo* AzPhe incorporation was also tested by measuring the enzymatic activity of β-galactosidase that is produced by amber suppression. pAcamG and an aaRS expression vector (pQE-YRS, or pQE-AzPheRS-1, -2, -3) were cotransformed into DH10B (Tn:*LacZ*am). After overnight seed culture, each

transformant was sub-inoculated (1:100) into 3 ml of LB broth containing 100 μg/ml ampicillin and 25 μg/ml tetracycline in the presence or absence of 1 mM AzPhe. Bacterial culture was harvested after 12 h culture at 37°C and its β-galactosidase activity was measured as described previously (Tekalign *et al.*, 2018). Briefly, bacterial culture (1.5 ml) was harvested and washed once with Z buffer (60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$, 50 mM 2-mercaptoethanol, pH 7.0). Cell pellet was resuspended in 50 μl of lysis buffer (1 mg/ml lysozyme, 30 mM Tris-Cl pH 8.0, 1 mM EDTA, 20% sucrose) and incubated for 4 min at room temperature. After adding 100 μl of Z buffer, 15 μl of 0.1% SDS, and 30 μl of chloroform to the cell resuspension, cells were broken by vortexing for 10 sec. After centrifugation for 10 min, the aqueous layer was taken for β-galactosidase assay. Enzyme reaction mixture consisting of 100 μl of cell extract, 200 μl of Z buffer, and 30 μl of 5 mM 4-methylumbelliferyl-β-D-galactopyranoside (MUG) was incubated for 30 min at 37°C. Reaction was stopped by adding 500 μl of 1 M Na_2CO_3 and fluorescence was measured at 360 nm excitation and 440 nm emission. In addition, protein quantitation was carried out with the same cell extracts by the Bradford method (Bio-Rad protein assay) and the protein concentration of each cell extract was used for normalization of β-galactosidase activity.

Purification of recombinant AzPheRS proteins

Four aaRS expression vectors (pQE-YRS, pQE-AzPheRS-1, -2, and -3) were separately transformed into the *E. coli* M15 strain. Each transformed clone was cultured in 500 ml of LB broth supplemented with 100 μg/ml of ampicillin and 20 μg/ml of kanamycin at 37°C. Protein expression was induced with 1 mM IPTG for 4 h. Cells were harvested by centrifugation. Purification of recombinant aaRS proteins was carried out by the manufacturer's protocol of HisPur cobalt resin (Thermo Scientific). Briefly, cell pellets were re-suspended in 15 ml of lysis buffer (30 mM HEPES, 300 mM NaCl, 10 mM 2-mercaptoethnol, pH 7.4) and 1 mg/ml of lysozyme was added. After incubation on ice for 30 min, cells were lysed by sonication on ice. After centrifugation at 12,000 rpm for 20 min at 4°C, 1.5 ml of cobalt resin was added to the cell extract and incubated at 4°C for 30 min with gentle rocking. The cobalt resin was packed in a column, and column was washed with 10 volumes of washing buffer (30 mM HEPES, 300 mM NaCl and 10 mM imidazole, pH 7.4). TyrRS and AzPheRS proteins were eluted with elution buffer (30 mM HEPES, 300 mM NaCl and 150 mM imidazole, pH 7.4). Protein concentration were measured by Bio-Rad assay kit using bovine serum albumin (BSA) as standard. Purification quality of the recombinant proteins was tested by SDS-PAGE analysis.

In vitro aminoacylation assay

To test *in vitro* aminoacylation activity of AzPheRS proteins, we have used a pyrophosphate-exchange assay that quantitatively measures pyrophosphates produced in aminoacylation reaction (Cestari and Stuart, 2012). Each aaRS (0.22 μM) was added into the reaction buffer containing 30 mM HEPES, 140 mM NaCl, 30 mM KCl, 40 mM $MgCl_2$, 1 mM DTT, 200 μM ATP, 200 U/ml inorganic pyrophosphatase (PPiase), 1.6 μM *fMam*$tRNA_{CUA}$ and 1 mM of either tyrosine or AzPhe. The aminoacylation reactions (30 μl total volume) were performed in flat-bottom 96-well microplate and incubated for 30 min at 37°C. Reactions were stopped by addition of 0.3 M EDTA, and 100 μl malachite green (Echelon Biosciences) was added to detect the inorganic phosphate released during aminoacylation reaction. After incubation for 30 min at RT, absorbance was measured at 620 nm using Infinite M200 Pro Spectrophotometer (Nano Quent). Inorganic phosphate production was calculated from the standard curve produced according to manufacturer's protocol.

In addition, to compare the aminoacylation kinetics of AzPheRS proteins against Tyr and AzPhe, *in vitro* aminoacylation reactions were performed in increasing concentrations (0–2 mM) of either Tyr or AzPhe. Lineweaver-Burk plot was drawn base on the assay results and K_m, K_{cat}, and K_{cat}/K_m values were calculated.

Site-specific *in vivo* incorporation of AzPhe into $GFP_{(TAG)}$ by AzPheRS mutants

To examine the *in vivo* site-specific integration of AzPhe into a reporter protein, we conducted the cloning of *Sc* TyrRS and AzPheRS genes into the plasmid pACamG, which resulting in the incorporation of both an aaRS gene and the *fMam*$tRNA_{CUA}$ gene within a single plasmid. The TyrRS and AzPheRS genes from pQE expression vectors, specifically the pQE-YRS and pQE-AzPheRS plasmids, were amplified through PCR using

E-3 and E-5 primers detailed in Table 2. Subsequently, the PCR products were subcloned into pACamG utilizing *SacI* and *PstI* restriction sites, and the successful cloning of TyrRS and AzPheRS genes was verified through sequencing analysis. Each of the aaRS/suptRNA expression vectors (pACamG-YRS, pACamG-AzPheRS-1, -2, & -3, or pCNFRS) was cotransformed into DH10B along with a reporter plasmid, pBad-sfGFP-150TAG. Transformants were cultured in auto-induction media supplemented with 100 μg/ml ampicillin, 25 μg/ml tetracycline, and either 1 mM AzPhe or none, at 37°C for 36 h. Following this, 1 ml of cell culture was harvested, resuspended in phosphate-buffered saline, and the fluorescence intensity of $GFP_{(TAG)}$ protein was measured at 485 nm excitation and 528 nm emission using Infinite Pro-200 Spectrophotometry (Nano-Quent).

To purify $GFP_{(TAG)}$ proteins produced by amber suppression, a consistent amount of cell culture (500 ml) was harvested. The protein purification process followed the same method employed for purifying AzPheRS proteins. The elution of $GFP_{(TAG)}$ proteins was carried out using 1.5 ml of elution buffer (30 mM HEPES, 300 mM NaCl, and 150 mM imidazole). Subsequently, an equivalent volume (15 μl) of the purified $GFP_{(TAG)}$ proteins was subjected to analysis on SDS-PAGE to compare the production quantity of $GFP_{(TAG)}$ in each bacterial culture. The concentrations of the purified proteins were determined using a Bio-Rad assay kit with BSA as a standard.

MALDI mass spectrometry

For tryptic digestion, $GFP_{(TAG)}$ proteins purified from bacteria cultured in the presence of 1 mM AzPhe were incubated with 25 ng/ml trypsin gold (Promega) in 50 mM Tris-HCl (pH 8) at 37°C overnight. Reaction was terminated with 1% trifluoroacetic acid and dried in a vacuum centrifuge. Peptides were resuspended in trifluoroacetic acid and analyzed in the MS/MS mode using an ABI 4800 Plus MALDI-TOF/TOF analyzer (Applied Biosystems). CHCA (α-cyano-4-hydroxycinnamic acid) matrix were used.

Statistical analysis

The experiments were repeated at least more than 3 times to obtain reproducibility, the mean, and standard deviation. For statistical analysis, Student's unpaired *t*-test was performed.

Results

Rational Design of *Sc* TyrRS mutants for AzPhe incorporation

Mj TyrRS was the first successful aaRS that was used for genetically incorporating diverse unnatural amino acids in *E. coli* by modifying its amino acid binding pocket (Liu and Schultz, 2010). Crystal structures of the TyrRS-$tRNA^{Tyr}$-*L*-tyrosine complexes from both *Mj* and *Sc* are informative for understanding specific amino acid residues that interact with tyrosine-AMP at the active site (Kobayashi *et al*., 2003; Tsunoda *et al.,* 2007). Protein sequence alignment between wild-type *Mj* TyrRS and *Sc* TyrRS reveals a high level of similarity in the residues involved in tyrosine recognition (Supplementary data Fig. S1). Thus, based on amino acid changes in *Mj* TyrRS mutants for AzPhe incorporation, we have chosen 5 amino acid residues (Y43, Y127, D177, E178, and L181) of *Sc* TyrRS composing binding pocket of amino acid substrates and mutated them to the corresponding amino acids present in *Mj* AzPheRS.

We have generated 3 different mutants to test the effects of amino acid changes on AzPhe incorporation. The first one, AzPheRS-1, contains mutations at only 2 residues, Y43T and Y127F, while the other 2 mutants have all five mutations but with one difference at L181 residue. The amino acid changes are shown in Table 3.

AzPhe-dependent *in vivo* amber suppression activity of AzPheRS mutants

To test whether the rationally designed mutants had activity in incorporating AzPhe, we performed a chloramphenicol resistance assay of bacterial growth mediated by amber suppression of the CAT gene. A plasmid encoding a suppressor tRNA (*fMam*$tRNA_{CUA}$) and CAT amber mutant, pACamG, was cotransformed into DH10B cells with an aaRS expression vector, such as pQE-YRS, pQE-AzPheRS-1, -2, or -3. Chloramphenicol resistance of the transformants was measured from the bacterial culture growing in GMML media containing 120 μg/ml chloramphenicol with or without 1 mM AzPhe. Optical densities of each bacterial culture were measured after 48 h, and their growth rates were compared (Fig. 1A). The growth rate of bacteria expressing *fMam*$tRNA_{CUA}$ and wild-type *Sc* TyrRS (pQE-YRS) was not affected by the presence of AzPhe. On the other hand, the growth rates of bacteria expressing

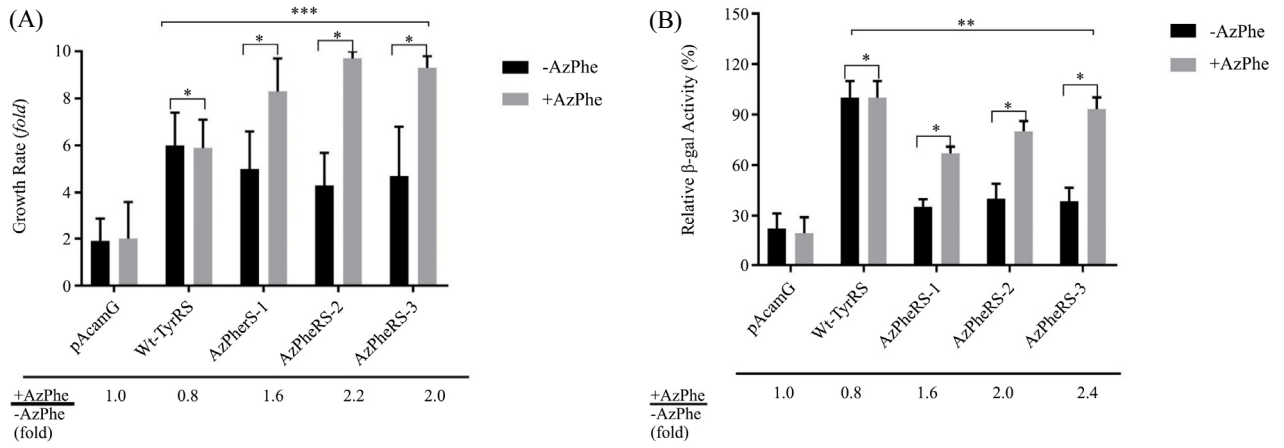

Fig. 1. *In vivo* amber suppression activity of AzPheRS mutants is dependent on AzPhe. (A) Chloramphenicol-resistant growth depends on AzPhe. Bacteria transformed with pACamG plus an aaRS expression plasmid as indicated were cultured in GMML media. After 48 h culture in the presence or absence of 1 mM AzPhe, the chloramphenicol-resistant growth of bacteria was measured. (B) The activities of β-galactosidase produced by amber suppression in DH10B (Tn:*lac*Zam) were compared among bacteria containing either *Sc* TyrRS or AzPheRS mutants in the presence or absence of AzPhe. *$P < 0.01$; ** $P < 0.05$.

AzPheRS mutants (pQE-AzPheRSs) were approximately 1.6–2.2 fold higher when cultured in the presence of AzPhe. Among the three AzPheRS mutants, AzPheRS-1, which contains mutations at two residues, showed the smallest increase in chloramphenicol resistance.

To validate AzPhe dependence on amber suppression mediated by AzPheRS mutants, we used a reporter strain, *E. coli* DH10B (Tn*:lacZ*am), which contains a single copy of an amber-mutated *lacZ* gene in the genomic DNA (Kim *et al*., 2009). *E. coli* DH10B(Tn*:lacZ*am) was co-transformed with pACamG and one of the aaRS expression vectors, and the transformants were cultured in the presence or absence of AzPhe for 24 h and harvested for a β-galactosidase assay (Fig. 1B). Transformants expressing *Sc* TyrRS did not show any significant difference in β-galactosidase activity upon the presence of AzPhe. On the other hand, all three AzPheRS mutants showed increased β -galactosidase activity of 1.6–2.4 times in the presence of AzPhe.

In vitro aminoacylation activity of AzPheRS mutants

To assess the direct aminoacylation activities of AzPheRS mutants for both AzPhe and tyrosine, we purified recombinant *Sc* TyrRS and AzPheRS mutant proteins from *E. coli.* SDS-PAGE analysis showed the proper purification of all recombinant aaRSs with no significant contamination (Fig. 2A). Using these purified recombinant aaRSs, an *in vitro* aminoacylation assay was conducted with *fMam*tRNA$_{CUA}$ in the presence of either AzPhe or tyrosine (Fig. 2B). The wild-type *Sc* TyrRS exhibited a 5-fold higher aminoacylation activity for tyrosine than for AzPhe. In contrast, compared to the wild-type TyrRS, the three AzPheRS mutants displayed 70% lower activities of aminoacylation for tyrosine and enhanced aminoacylation activities for AzPhe. AzPheRS-2 and -3 mutants exhibited more than a 2.5-fold higher aminoacylation activity for AzPhe than for tyrosine, while the AzPheRS-1 mutant showed a minor increase of 1.6-fold.

Kinetic analysis of aminoacylation for AzPhe and tyrosine catalyzed by the recombinant aaRSs revealed similar results (Table 4). The K_m values of AzPheRS mutants for tyrosine were approximately 6.7-fold higher than that of *Sc* TyrRS. The k_{cat}/K_m values of AzPheRS mutants for tyrosine were approximately 11.5-fold lower than that of *Sc* TyrRS. The decreased efficiency of aminoacylation for tyrosine catalyzed by AzPheRS mutants was primarily attributed to the high K_m values of AzPheRS for tyrosine, as the decrease in the k_{cat} values of AzPheRS mutants for tyrosine was less than 2-fold compared to that of *Sc* TyrRS. Concerning the kinetics of AzPhe aminoacylation, all three AzPheRS mutants showed similar K_m values for AzPhe, but their k_{cat}/K_m values were slightly different. The k_{cat}/K_m values of AzPheRS-2 and -3 mutants were approximately 40% higher than that of AzPheRS-1. *Sc* TyrRS could not be analyzed for the kinetics of AzPhe aminoacylation because there was no

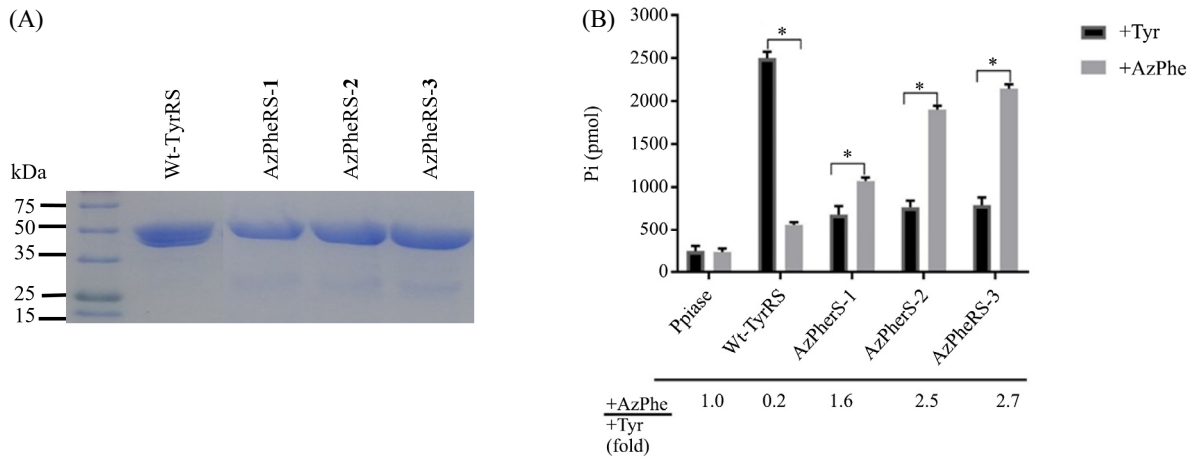

Fig. 2. *In vitro* aminoacylation activity of AzPheRS mutants. (A) SDS-PAGE analysis of purified recombinant aaRS proteins. Recombinant proteins of *Sc* TyrRS and AzPheRS mutants were produced in *E. coli* and purified as described in the section 'Materials and Methods.' (B) *in vitro* aminoacylation activities of the purified recombinant aaRSs against both tyrosine and AzPhe. **P* < 0.01.

Table 4. Kinetic parameters for *in vitro* aminoacylation of *Sc* TyrRS and AzPheRS mutants

	TyrRS		AzPheRS-1		AzPheRS-2		AzPheRS-3	
	+ Tyr	+ AzPhe	+ Tyr	+ AzPhe	+ Tyr	+ AzPhe	+ Tyr	+ AzPhe
K_m (μM)	18	n.d.	119	112	122	108	124	116
k_{cat} (min^{-1})	114	n.d.	57	103	71	126	76	142
k_{cat}/K_m	6.31	n.d.	0.48	0.92	0.58	1.17	0.61	1.23

change in the reaction product upon an increase in substrate concentrations (Supplementary data Fig. S2).

In vivo incorporation of AzPhe into a reporter protein GFP$_{(TAG)}$

To assess AzPhe-dependent production of a reporter protein GFP$_{(TAG)}$ mediated by amber suppression of AzPheRS mutants, we co-transformed DH10B cells with a vector expressing both *fMam*tRNA$_{CUA}$ and aaRS (pACamG-TyrRS or pACamG-AzPheRS-1, -2, and -3) and an expression vector of the reporter protein GFP$_{(TAG)}$ (pBad-sfGFP-150TAG). Each transformant was cultured in auto-induction media in the presence or absence of AzPhe, and GFP$_{(TAG)}$ production was evaluated by measuring the fluorescence intensity of the whole cell harvest (Fig. 3A). Transformants producing only *fMam*tRNA$_{CUA}$ (pACamG) were unable to produce GFP$_{(TAG)}$ proteins. In contrast, bacterial cells producing a previously known orthogonal pair of tRNA$_{CUA}$ and AzPheRS derived form *Mj* TyrRS (pCNFRS) produced a 3.5-fold higher amount of GFP$_{(TAG)}$ in the presence of AzPhe. Our AzPheRS mutants derived from *Sc* TyrRS also produced increased amounts of GFP$_{(TAG)}$ in the presence of AzPhe compared to the absence of AzPhe. However, among the three AzPheRS mutants, AzPheRS-1 showed only a minor increase of GFP$_{(TAG)}$ production compared to AzPheRS-2 and -3, which showed more than a 2-fold increase in the presence of AzPhe. These data were consistent with our previous experimental results of AzPhe-dependent β-galactosidase activity as well as *in vitro* aminoacylation assays.

We also conducted a standard assessment of protein production in the presence and absence of AzPhe. Bacterial cells transformed with both pBad-sfGFP-150TAG and one of the plasmids expressing an othogonal pair of tRNA$_{CUA}$/aaRS (pACamG-TyrRS, pACamG-AzPheRS-1, -2, and -3, or pCNFRS), were cultured in the same volume of auto-induction media in the presence and absence of 1mM AzPhe for 36 h. Cells were harvested and purified with a standard protocol described in the section 'Materials and Methods.' To compare the quantity of GFP$_{(TAG)}$ production of each bacterial culture, the same volume of GFP purification was tested in SDS-PAGE analysis (Fig. 3B). GFP$_{(TAG)}$ production in bacteria containing an expression vector

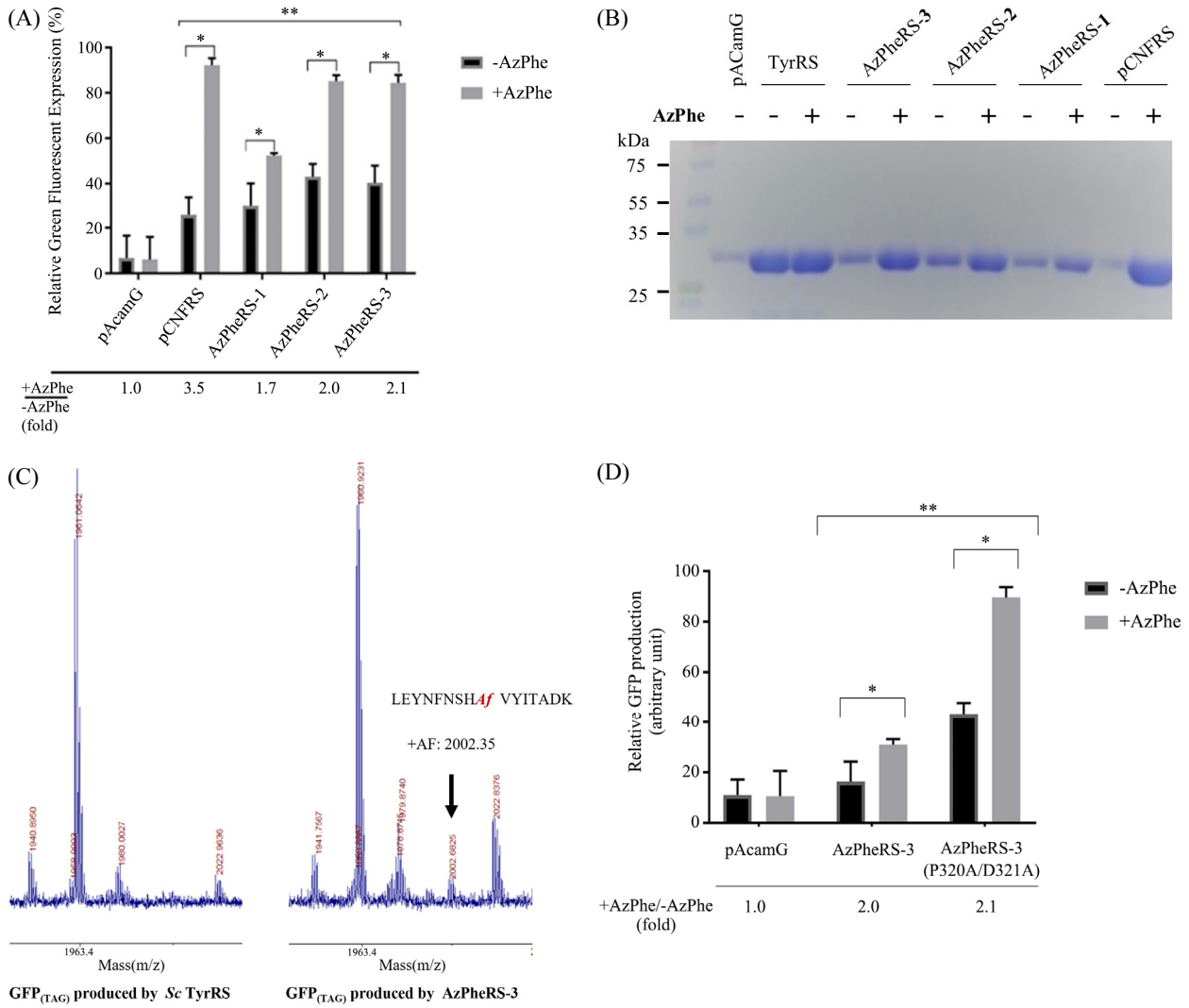

Fig. 3. *In vivo* incorporation of AzPhe into a reporter protein, $GFP_{(TAG)}$ by AzPheRS-mediated amber suppression (A) GFP fluorescence of bacterial cells cultured in the presence or absence of 1 mM AzPhe. Each label indicates the following: pAcamG, bacteria expressing only *fMam*$tRNA_{CUA}$; pCNFRS, bacteria expressing AzPheRS derived from *Mj* and its cognate suppressor tRNA; AzPheRS-1, -2, and -3, bacteria expressing AzPheRSs derived from *Sc* and *fMam*$tRNA_{CUA}$. (B) SDS-PAGE analysis of purified recombinant $GFP_{(TAG)}$ proteins produced by amber suppression in the presence or absence of 1 mM AzPhe. (C) MALDI-TOF mass spectra of tryptic peptides digested from $GFP_{(TAG)}$ proteins produced by either *Sc* TyrRS or AzPheRS-3. The spectrum of AzPheRS-3 sample shows a peak at m/z 2002.68, which matches the expected mass 2002.35 for the AzPhe-containing peptide. This peak was not observed in the spectrum of the *Sc* TyrRS sample. The amino acid sequence of the AzPhe-containing tryptic peptide is indicated. (D) Enhanced amber suppression activity of AzPheRS-3 with P320A/D321A mutations. P320A and D321A mutations were introduced into the AzPheRS-3 gene, and its amber suppression activity was compared to its parental type AzPheRS-3 by measuring the fluorescence produced through amber suppression of $GFP_{(TAG)}$. $*P < 0.01$; $**P < 0.005$.

of *Sc* TyrRS did not show any difference between the presence and the absence of AzPhe. On the other hand, AzPheRS mutants showed a larger amount of $GFP_{(TAG)}$ production in the presence of AzPhe, and the production yield was higher in bacteria expressing AzPheRS-2 or -3 than in the one expressing AzPheRS-1. The protein concentrations of purified $GFP_{(TAG)}$ produced in the presence of AzPhe were measured as 80 mg/L, 120 mg/L, 150 mg/L for AzPheRS-1, -2, and -3, respectively, and 200 mg/L for *Mj* AzPheRS (pCNFRS).

However, all three AzPheRS mutants still produced small amounts of $GFP_{(TAG)}$ proteins in the absence of AzPhe, and the production quantities among the mutants were very similar. On the other hand, *Mj* AzPheRS showed a very tiny amount of $GFP_{(TAG)}$ production in the absence of AzPhe, and produced a

larger amount of $GFP_{(TAG)}$ in the presence of AzPhe than our AzPheRS mutants. These results indicate that our AzPheRS mutants still have a substantial amount of aminoacylation activity for tyrosine, and that the previously known *Mj* AzPheRS has better specificity and aminoacylation activity for AzPhe than our AzPheRS mutants.

To verify AzPhe incorporation into the reporter protein $GFP_{(TAG)}$, we performed MALDI-TOF mass spectrometry analysis with tryptic peptides of purified $GFP_{(TAG)}$ proteins (Fig. 3C). The $GFP_{(TAG)}$ coding sequence contains an amber codon at the 150th residue, and the molecular mass of the tryptic peptide containing AzPhe at the amber codon is calculated as 2002.35 Da. Mass spectrum of tryptic peptides from $GFP_{(TAG)}$ produced by AzPheRS-3 showed a peak at 2002.68 m/z, which matched with the theoretical mass of the AzPhe-containing peptide, whereas the same peak was not observed at the mass spectrum done with tryptic peptides of $GFP_{(TAG)}$ produced by *Sc* TyrRS.

In a previous study, we characterized that P320A and D321A mutations in *Sc* TyrRS resulted in a 6.5-fold increase in amber suppression activity compared to the wild-type *Sc* TyrRS, owing to an augmented binding affinity to *fMam*$tRNA_{CUA}$ (Tekalign *et al.,* 2018). To explore whether these mutations could similarly enhance AzPhe incorporation in response to an amber codon, we introduced the P320A and D321A mutations into AzPheRS-3 using site-directed mutagenesis. Plasmids expressing AzPheRS-3 or AzPheRS-3(P320A/D321A) were co-transformed into *E. coli* DH10B along with pBad-sfGFP-150TAG. Transformants were cultured in the presence or absence of 1 mM AzPhe for 36 h, and the resulting GFP fluorescence in bacteria was quantified (Fig. 3D). AzPheRS-3(P320A/D321A) exhibited an approximately 3.2-fold increase in $GFP_{(TAG)}$ production compared to its parental synthetase, AzPheRS-3. This improvement primarily stemmed from the heightened binding affinity of the AzPheRS-3(P320A/D321A) mutant to *fMam*$tRNA_{CUA}$, as indicated by the proportional increase in $GFP_{(TAG)}$ production from the same bacteria cultured in the absence of AzPhe. Consequently, these findings not only reaffirm our earlier report that P320A/D321A mutations enhance the amber suppression activity of *Sc* TyrRS by increasing binding affinity to *fMam*$tRNA_{CUA}$ but also suggest the potential utility of these mutations for efficiently incorporating unnatural amino acids, such as AzPhe.

Discussion

Based on the structural similarity between *Sc* TyrRS and *Mj* TyrRS, as well as information from AzPheRS variants derived from *Mj* TyrRS, we strategically designed three AzPheRS mutants originating from *Sc* TyrRS. Both i*n vitro* aminoacylation assays and *in vivo* incorporation of AzPhe revealed that AzPheRS-1 exhibited significantly weaker activity compared to the other two mutants, AzPheRS-2 and -3. However, all three mutants displayed a similar loss of tyrosylation activity and had comparable K_m values for tyrosine. These findings suggest that mutations at two residues, Y43T and Y127F (common to all three AzPheRS mutants), notably compromised tyrosine recognition, albeit not completely abolishing it.

Crystal structure of a *Sc* TyrRS-tRNA-tyrosine-AMP analog complex revealed the critical roles of Y43 and D177 in tyrosine recognition, forming hydrogen bonds with the hydroxyl group of the tyrosine moiety (Tsunoda *et al*., 2007). In addition, Y43G mutation in *Sc* TyrRS resulted in a dramatic 85-fold increase in the K_m value for tyrosine, enabling the aminoacylation of 3-ioido-*L*-tyrosine (Ohno *et al*., 2001). Although our experimental data align with these reports, the K_m values of our AzPheRS mutants increased only about 7-fold compared to that of *Sc* TyrRS. Both *in vivo* and *in vitro* assays demonstrated residual tyrosylation activities in all three AzPheRS mutants. Given that the removal of tyrosylation activity is crucial for the effective application of this technique, additional development, such as screening a random mutation library, particularly at the Y43 residue, might enhance the selectivity of AzPheRS.

While Y43T and Y127F mutations compromised tyrosine recognition, they were insufficient for the effective acylation of AzPhe, as AzPheRS-1 showed only a minor increase in AzPhe acylation activity. In contrast, AzPheRS-2 and -3 mutants, containing additional mutations at three residues (D177, E178, and L181) of Sc TyrRS, exhibited a significant increase in AzPhe acylation. Kinetic analysis of the AzPheRS mutants revealed that the additional three mutations contributed to an increase in K_{cat} values without affecting binding affinities of AzPheRS mutants for AzPhe. This suggests that mutations at these three residues are involved in the efficient acylation reaction of AzPhe. The only distinction between AzPheRS-2 and -3 is an amino acid substitution at L181, specifically

L181G in AzPheRS-2 and L181Q in AzPheRS-3. Despite the structural difference between glycine and glutamine, these mutants exhibited similar AzPhe acylation activities, implying that L181 may not be crucial for the specific recognition and acylation of AzPhe by our AzPheRS mutants. Further studies will be necessary to clarify this interpretation.

In conclusion, we have established a system that genetically incorporates AzPhe at the amber codon of a target protein in *E. coli*, utilizing a new orthogonal pair of *Sc* TyrRS/*fMam*tRNA$_{CUA}$, distinct from the conventionally used orthogonal pair of *Mj* TyrRS/tRNA^{Tyr}. Although this system requires further enhancement of aminoacylation activity and specificity to AzPhe for diverse applications, our results demonstrate the potential utility of the orthogonal pair of *Sc* TyrRS/*fMam*tRNA$_{CUA}$ in expanding the genetic code for the incorporation of unnatural amino acids in *E. coli.*

적 요

단백질의 특정 위치에 비천연아미노산을 삽입함으로써 변형된 단백질의 기능을 다양한 분야에서 활용할 수 있다. 이 연구는 대장균에서 효모균 타이로신-tRNA 합성효소(*Sc* TyrRS)와 변형된 대장균 시작 tRNA (*fMam*tRNA$_{CUA}$)의 이종 쌍을 사용하여 *p*-아지도페닐알라닌(AzPhe)을 단백질에 유전적으로 삽입하는 방법을 개발하는데 있다. 우리는 *Methanocaldococcus jannaschii* (*Mj*) TyrRS와 *Sc* TyrRS의 구조적 유사성과 *Mj* TyrRS를 돌연변이 시켜 유래된 AzPhe-tRNA 합성효소와의 유사성을 가정하여 세 가지 다른 *Sc* TyrRS 변이체를 만들었다. 세 가지 변이체는 AzPheRS-1, -2 및 -3으로 명명되었으며, 이들의 AzPhe 삽입 능력을 엠버서프레션에 의해 나타나는 클로람페니콜 항생제 저항성과 β-갈락토시다제 활성을 측정하여 평가하였다. 이 변이체를 가진 대장균은 AzPhe 존재시 2~3배 높은 항생제 저항성과 β-갈락토시다제 활성을 보였다. 또한 정제된 재조합 아미노아실-tRNA 합성효소(aaRS) 단백질을 사용한 *in vitro* 아미노아실화 실험에서 세 가지 AzPheRS 변이체는 *Sc* TyrRS와 비교하여 AzPhe에 대한 아미노아실화 활성이 향상되었으며, 티로신의 촉매 효율성(k_{cat}/K_m)은 대략 10배 낮아졌다. 대장균에서 세 가지 변이체 모두 AzPhe를 리포터 단백질 GFP$_{(TAG)}$에 성공적으로 삽입하였으며, AzPheRS-2 및 AzPheRS-3이 AzPheRS-1보다 높은 효율을 나타냈다. 이러한 결과는 *Sc* TyrRS 및 *fMam*tRNA$_{CUA}$의 이종 쌍이 대장균에서 유전적 코드를 확장하여 비천연 아미노산을 단백질의 특정 위치에 삽입할 수 있는 추가적인 플랫폼이 될 수 있음을 보여주며, 새로운 기능을 가진 단백질을 생성하는 데 도움이 될 수 있다.

Acknowledgments

The authors thank Dr. RajBhandary for the gifts of pACamG and Dr. Mehl for pCNFRS and pBad-sfGFP-150TAG. This research was supported by the grant of GRRC and internal research grant from HUFS.

Conflict of Interest

The authors have no conflict of interest to report.

References

Cestari I and Stuart K. 2012. A spectrophotometric assay for quantitative measurement of aminoacyl-tRNA synthetase activity. *J. Biomol. Screen*. **18**, 490–497.

Chin JW. 2017. Expanding and reprogramming the genetic code. *Nature* **550**, 53–60.

Chin JW, Martin AB, King DS, Wang L, and Schultz PG. 2002a. Addition of a photocrosslinking amino acid to the genetic code of *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **99**, 11020–11024.

Chin JW, Santoro SW, Martin AB, King DS, Wang L, and Schultz PG. 2002b. Addition of *p*-azido-L-phenylalanine to the genetic code of *Escherichia coli*. *J. Am. Chem. Soc.* **124**, 9026–9027.

Chow CM and RajBhandary UL. 1993. *Saccharomyces cerevisiae* cytoplasmic tyrosyl-tRNA synthetase gene. *J. Biol. Chem*. **268**, 12855–12863.

Edan DS, Choi I, and Park J. 2014. Establishment of a selection system for the site-specific incorporation of unnatural amino acids into protein. *Korean J. Microbiol*. **50**, 1–7.

Hammill JT, Miyake-Stoner S, Hazen JL, Jackson JC, and Mehl RA. 2007. Preparation of site-specifically labeled fluorinated proteins for ^{19}F-NMR structural characterization. *Nat. Protoc.* **2**, 2601–2607.

Kim K, Park MY, and Park J. 2009. Establishment of an *in vivo* report system for the evaluation of amber suppression activity in *Escherichia coli*. *Korean J. Microbiol.* **45**, 215–221.

Kobayashi T, Nureki O, Ishitani R, Yaremchuk A, Tukalo M, Cusack S, Sakamoto K, and Yokoyama S. 2003. Structural basis for

orthogonal tRNA specificities of tyrosyl-tRNA synthetases for genetic code expansion. *Nature Struct. Mol. Biol.* **10**, 425–432.

Kowal AK, Kohrer C, and RajBhandary UL. 2001. Twenty-first aminoacyl-tRNA synthetase-suppressor tRNA pairs for possible use in site-specific incorporation of amino acid analogues into proteins in eukaryotes and in eubacteria. *Proc. Natl. Acad. Sci. USA* **98**, 2268–2273.

Liu DR and Schultz PG. 1999. Progress toward the evolution of an organism with an expanded genetic code. *Proc. Natl. Acad. Sci. USA* **96**, 4780–4785.

Liu CC and Schultz PG. 2010. Adding new chemistries to the genetic code. *Annu. Rev. Biochem*. **79**, 413–444.

Miyake-Stoner SJ, Refakis CA, Hammill JT, Lusic H, Hazen JL, Deiters A, and Mehl RA. 2010. Generating permissive site-specific unnatural aminoacyl-tRNA synthetases. *Biochemistry* **49**, 1667–1677.

Ohno S, Yokogawa T, and Nishikawa K. 2001. Changing the amino acid specificity of yeast tyrosyl-tRNA synthetase by genetic engineering. *J. Biochem*. **130**, 417–423.

Tekalign E, Oh JE, and Park J. 2018. Improving amber suppression activity of an orthogonal pair of *Saccharomyces cerevisiae* tyrosyl-tRNA synthetase and a variant of *E. coli* initiator tRNA, *fMam* $tRNA_{CUA}$, for the efficient incorporation of unnatural amino acids. *Korean J. Microbiol.* **54**, 420–427.

Tsunoda M, Kusakabe Y, Tanaka N, Ohno S, Nakamura M, Senda T, Moriguchi T, Asai N, Sekine M, Yokogawa T, *et al*. 2007. Structural basis for recognition of cognate tRNA by tyrosyl-tRNA synthetase from three kingdoms *Nucleic Acids Res.* **35**, 4289–4300.

Wang L, Brock A, Herberich B, and Schultz PG. 2001. Expanding the genetic code of *Escherichia coli*. *Science* **292**, 498–500.

Wang L, Brock A, and Schultz PG. 2002. Adding L-3-(2-naphthyl)-alanine to the genetic code of *E. coli. J. Am. Chem. Soc*. **124**, 1836–1837.

Wang L, Zhang Z, Brock A, and Schultz PG. 2003. Addition of the keto functional group to the genetic code of *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **100**, 56–61.