ACS SyntheticBiology

Research Article

pubs.acs.org/synthbio

Tailoring the Substrate Specificity of Yeast Phenylalanyl-tRNA Synthetase toward a Phenylalanine Analog Using Multiple-Site-Specific Incorporation

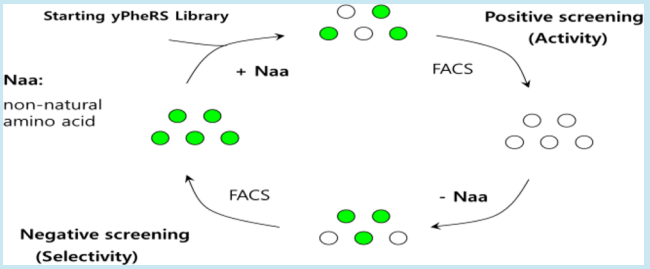
Inchan Kwon*[,†,‡] and Sung In Lim[‡]

[†]School of Materials Science and Engineering, Gwangju Institute of Science and Technology (GIST), Gwangju 500-712, Republic of Korea

[‡]Department of Chemical Engineering, University of Virginia, Charlottesville, Virginia 22904-4741, United States

Supporting Information

ABSTRACT: A yeast phenylalanyl-tRNA synthetase variant with T415G mutation (yPheRS (T415G)) was rationally designed to recognize various phenylalanine (Phe) analogs allowing site-specific incorporation into an amber site of a protein in *E. coli*. However, the relaxed substrate specificity of yPheRS (T415G) led to a significant tryptophan (Trp) misincorporation restricting the utility of yPheRS for biosynthesis of proteins containing a Phe analog. In order to obtain yPheRS variants with high substrate-specificity toward a Phe analog, we developed a general high-throughput screening method. This method uses fluorescence reduction of green fluorescence protein (GFP) upon efficient introduction of a Phe analog into multiple sites of GFP by breaking the degeneracy of the Phe codons. Combined use of positive and negative screenings of a yPheRS saturation library led to a yPheRS variant (yPheRS_naph) very selective toward 2-L-naphthylalanine (2Nal), a model Phe analog. The yPheRS_naph exhibited 6-fold higher relative activity toward 2Nal (vs Trp) in ATP-PPi exchange assays and led to high-fidelity incorporation of 2Nal into an amber site of murine dihydrofolate reductase in both minimal and rich media. These results successfully demonstrate that the high-throughput screening method developed can be used to evolve yPheRS to be very selective toward a Phe analog.



KEYWORDS: *yeast phenylalanyl-tRNA synthetase, substrate specificity, saturation mutagenesis, 2-L-naphthylalanine, high-throughput screening, multiple-site-specific incorporation,*

Non-natural amino acids have unique physical, chemical, or biological properties that are distinct from natural amino acids. Incorporating non-natural amino acids has resulted in recombinant proteins with unusual properties for use in diverse applications such as X-ray crystallography, proteomics, therapeutic protein engineering, and *in vivo* imaging.[1−9] To incorporate a non-natural amino acid into a specific site of a target protein *in vivo*, the protein biosynthesis machinery of a host organism expressing a target protein containing a non-natural amino acid should be engineered. Briefly, introduction of a heterologous orthogonal pair of aminoacyl-tRNA and aminoacyl-tRNA synthetase (aaRS) into expression hosts, a synthetic biological procedure, is required.[10−15] For *E. coli* expression hosts, an orthogonal pair of tyrosyl-tRNA and tyrosyl-tRNA synthetase (*Mj*TyrRS) derived from *Methanococcus jannaschii* were successfully used. So far, diverse phenylalanine (Phe) and tyrosine (Tyr) analogs have been incorporated into a protein in *E. coli* in a site-specific manner. This is due, in part, to the active site of *Mj*TyrRS being very flexible in accommodating Phe and Tyr analogs and the development of a powerful screening method to change *Mj*TyrRS substrate specificity.[11,16−19] Phenylalanyl-tRNA synthetase derived from yeast (yPheRS), as well as its cognate tRNA (ytRNA^{Phe}), was also used to achieve site-specific incorporation of several Phe analogs into a protein in *E. coli*.[20,21] However, compared to the numerous applications of *Mj*TyrRS, the utility of yPheRS variants has been limited, at least in part, because of their moderate selectivity toward non-natural amino acids over natural ones. The rationally designed yPheRS (T415G) efficiently activates a natural amino acid, Trp, as well as various Phe analogs (*p*-bromo-L-phenylalanine, 2-L-naphthylalanine, etc.) due to the enlarged binding pocket generated by the T415G mutation (see the chemical structures of Phe, Trp, 2-L-naphthylalanine in Supporting Information Figure S1).[22] yPheRS (T415G) variant showed 10-fold higher activation activity toward Trp than 2-L-naphthylalanine.[22] The relaxed substrate specificity of yPheRS (T415G) presents several drawbacks. Cell growth rate decreased, likely due to misincorporation of Trp at Phe sites in essential proteins. Furthermore, misincorporation of Trp is observed at an amber codon, preventing high-fidelity site-specific incorporation of

Received: August 27, 2014
Published: September 30, 2014

pBrF. These drawbacks prompted the pursuit of more selective yPheRS variants. It has been reported that another rationally designed yPheRS (T415A) showed 10-fold higher activation activity toward pBrF than Trp. However, yPheRS (T415A) did not exhibit enhanced selectivity toward pBrF over Phe.[20] Therefore, in this study, we explored a high-throughput screening of yPheRS variants to obtain highly selective yPheRS variants toward Phe analogs.

Such a high-throughput screening method of yPheRS variants will greatly facilitate identification of novel yPheRS variants very specific to many Phe analogs. This will make an orthogonal pair of yPheRS and ytRNA^{Phe} a practical alternative for Phe analog incorporation. The availability of an alternate orthogonal pair has several benefits. First, it will increase the number of living organisms producing proteins containing non-natural amino acids. Currently, only a limited number of living organisms are being used for site-specific incorporation of non-natural amino acids, including *E. coli*, *Saccharomyces cerevisiae*, CHO cells, *Drosophila melanogaster*, and *C. elegans*.[2,14,15,23−25] As the number of applications of site-specific incorporation of non-natural amino acids continues to grow, it is expected that researchers will explore other living organisms for non-natural amino acid incorporation. Then, there is a possibility that the orthogonal pair of yPheRS and ytRNA^{Phe} is suitable for site-specific non-natural amino acid incorporation for a certain living organism. Second, by cointroducing multiple mutually orthogonal pairs into a single living organism, site-specific incorporation of multiple non-natural amino acids into a protein can be achieved. Neumann et al. demonstrated that the introduction of two mutually orthogonal pairs, *Mj*TyrRS/*Mj*tRNA and the *Methanosarcina barkeri* pyrrolysyl-tRNA synthetase/tRNA allowed introduction of two non-natural amino acids (*p*-azidophenylalanine and N$^{\epsilon}$-*tert*-butyl-oxy-carbonyl-L-lysine) into a single protein expressed in *E. coli*.[26] Since the incorporation of multiple non-natural amino acids into a protein requires multiple mutually orthogonal pairs, the availability of an alternative aaRS could expand the number of feasible combinations of non-natural amino acids to incorporate.

In this study, we describe the high-throughput screening of a yPheRS library to obtain yPheRS variants of which substrate specificity is changed to a Phe analog, 2-L-naphthylalanine (2Nal). We developed a screening strategy based on fluorescence loss of GFP upon multiple-site-specific incorporation of 2Nal. Since GFP fluorescence requires a correctly folded structure, the perturbation of the folded structure, such as circular permutation, fusion to an aggregation-prone protein, or residue-specific incorporation of 5,5,5-trifluoroleucin, reduces its fluorescence.[27−30] Although the incorporation of a non-natural amino acid into a single "nonpermissive site" of GFP may result in a substantial loss of its fluorescence, we chose multiple-site-specific incorporation of a non-natural amino acid into GFP to ensure a substantial structural perturbation leading to a significant fluorescence loss. Multiple-site-specific incorporation *in vivo* was previously realized by breaking the degeneracy of the Phe codons.[21] The combined use of yPheRS (T415G) variant and ytRNA variant containing modified AAA anticodon (ytRNA$^{Phe}_{AAA}$) favoring UUU Phe codons over UUC Phe codons, achieved efficient incorporation of 2Nal into mDHFR at five Phe sites encoded as UUU.[21] Therefore, in this study, we chose 2Nal as a model Phe analog to validate a high-throughput screening of a yPheRS library using multiple-site-specific incorporation of 2Nal.

In order to construct a yPheRS variant library, saturation mutagenesis was performed at four residues in the active site of yPheRS in close proximity of a substrate Phe. *E. coli* expression hosts were used to prepare an expression library of yPheRS variants. The screening of yPheRS libraries consists of multiple rounds of positive and negative screenings. As a positive screening to enrich yPheRS variants capable of 2Nal incorporation into GFP, *E. coli* cells exhibiting low cellular fluorescence were enriched using fluorescence activated cell sorting (FACS). As a negative screening to exclude *E. coli* cells that cannot express GFP, only *E. coli* cells with high cellular fluorescence were enriched using FACS.

RESULTS AND DISCUSSION

Relaxed Substrate Specificity of yPheRS (T415G) Variant. In order to evaluate the substrate specificity of yPheRS (T415G) on 2Nal and other natural amino acids, an amber codon was introduced to position 38 of mDHFR and the occupancy of the amber codon by 2Nal and other natural amino acids were analyzed. The modified mDHFR (mDHFR (38Am)) was expressed in AFWK [pQE16_mDHFR (38Am)_yPheRS (T415G)/pREP4_ytRNA^{Phe}_UG] expression hosts in minimal medium supplemented with 17 amino acids (MM17_FW), 2.5 μM Phe, 3.0 μM Trp, 50 μM Lys, and 3 mM 2Nal. The tryptic digests of the modified mDHFR were subjected to MALDI-TOF analysis. The MALDI-TOF spectrum of peptide **1Am** (residues 26−39; NGDLPWPPLRNEAmK) with an amber codon (Am) at position 38 is shown in Figure 1a. One major peak indicates peptide **1Am** containing 2Nal at position 38 (**1Z**) (actual m/z = 1732.4; expected m/z = 1732.1). The minor peak indicates **1Am** containing Trp at position 38 (**1W**) (actual m/z = 1721.4; expected m/z = 1721.1) implying that yPheRS (T415G) activates Trp as well as 2Nal. In order to confirm this, mDHFR (38Am) was expressed in the same minimal medium except 3 mM Trp instead of 3 μM Trp. The MALDI-TOF spectrum shows an enhanced intensity of **1W** signal (Figure 1b) clearly demonstrating the moderate substrate specificity of yPheRS (T415G) toward 2Nal compared to Trp.

Similar Trp misincorporation was observed in the case of multiple-site-specific incorporation of 2Nal at UUU Phe codons of mDHFR. The mDHFR was expressed in AFW [pQE16_mDHFR_yPheRS (T415G)/pREP4_ytRNA^{Phe}_AAA] expression hosts in minimal medium supplemented with 18 amino acids (MM18_FW), 2.5 μM Phe, 3 mM Trp. Occupancy of the UUU codon sites was determined by LC-MS analysis of tryptic digests of mDHFR expressed with or without 3 mM 2Nal. We focused on peptide **2** (residues 140−144; LF$_{UUU}$VTR), which contains a Phe residue encoded as UUU codon at position 141. LC-MS analysis of peptide **2** variants indicated that 10% of position 141 was occupied by Trp (Figure 1d), while the original peptide **2** has only Phe at position 141 (Figure 1c) supporting the relaxed substrate specificity of yPheRS (T415G).

Construction of yPheRS Library. In order to enhance the specificity of yPheRS with respect to 2Nal, we explored high-throughput screening of a yPheRS library. Based on the crystal structure of the homologous PheRS (tPheRS) from *Thermus thermophiles*[31] and sequence alignment between tPheRS and yPheRS, a homology model of yPheRS was constructed (Supporting Information Figure S2). Four residues (N412, S418, T415, and S437) within 7 Å of the *para* position of the phenyl-ring of Phe bound to yPheRS were subjected to

dx.doi.org/10.1021/sb500309r | ACS Synth. Biol. 2015, 4, 634−643

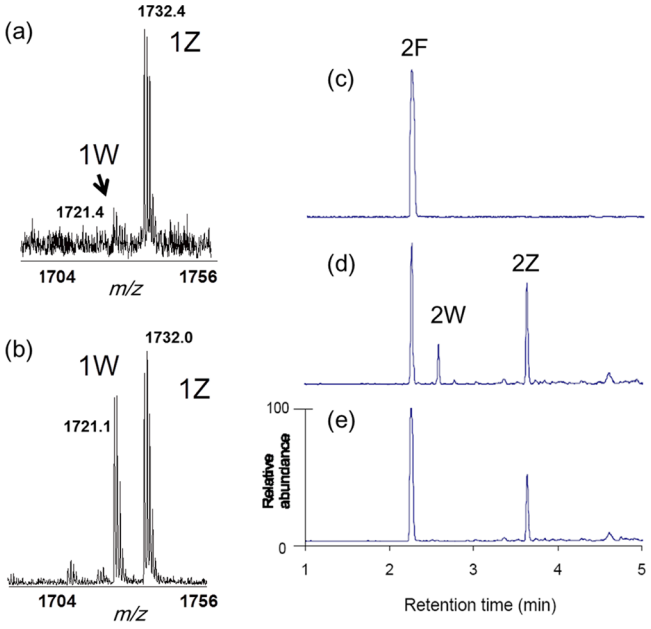


Figure 1. MALDI-TOF spectra (a and b) and LC-MS chromatogram (c–e) of tryptic digests of mDHFR. Peptide **1Am** (residues 26–39; NGDLPWPPLRNEAmK) variants of which Am (amber codon) at position 38 is occupied by Trp or 2Nal are designated **1W** or **1Z**, respectively. The modified mDHFR was synthesized in a Phe/Trp/Lys autoxtrophic expression host (AFWK) outfitted with ytRNA$^{Phe}_{CUA_UG}$ and yPheRS (T415G). The minimal expression medium was supplemented with 17 amino acids (25 μg/mL), 2.5 μM Phe, 50 μM Lys, 3 mM 2Nal, and 3.0 μM Trp (a) or 3 mM Trp (b). Peptide **2** (residues 140–144; LF$_{UUU}$VTR) contains a Phe residue encoded as UUU. Peptide **2** variants containing Phe, Trp, and 2Nal were designated **2F**, **2W**, and **2Z**, respectively. These peptides were separated by LC and detected by MS. Unmodified mDHFR was synthesized in a Phe/Trp auxotrophic expression host in 2× YT medium (c). Modified mDHFRs were synthesized in a Phe/Trp auxotrophic expression host outfitted with ytRNA$^{Phe}_{AAA}$ and yPheRS (T415G) (d) or yPheRS_naph (e). The minimal expression medium was supplemented with 18 amino acids (25 μg/mL), 2.5 μM Phe, 3 mM Trp, and 3 mM 2Nal (d and e).

saturation mutagenesis. Another homology model of yPheRS was constructed by the SWISS-MODEL server using the human cytosol phenylalanyl-tRNA synthetase with the sequence identify of 48.98 as a template. Even in the new homology model of yPheRS, the four residues (N412, S418, T415, and S437) are located within 7 Å of the *para* position of the phenyl-ring of Phe (Figure 2). PCR mutagenesis generated 6×10^6 yPheRS transformants (see Materials and Methods), a population larger than the theoretical library size ($10^6 \sim 32$ (NNK = $4 \times 4 \times 2$) $\times 32 \times 32 \times 32$). By transforming yPheRS library plasmids into DHF expression hosts, 1×10^7 transformants were obtained and pooled to construct the yPheRS expression library (L_01).

High-Throughput Screening of yPheRS Library. Screening of the yPheRS library entailed both positive and negative screenings to obtain active and selective yPheRS variants, respectively (Figure 3). Active synthetases allow incorporation of either 2Nal or natural amino acids at multiple nonpermissive sites of GFP6, where amino acids other than Phe lead to reduction in the fluorescence of cells. Our preliminary study showed that the incorporation of 2Nal at 12 Phe sites of GFP6 resulted in more than 20-fold reduction in the fluorescence of cells, likely due to the perturbed folded

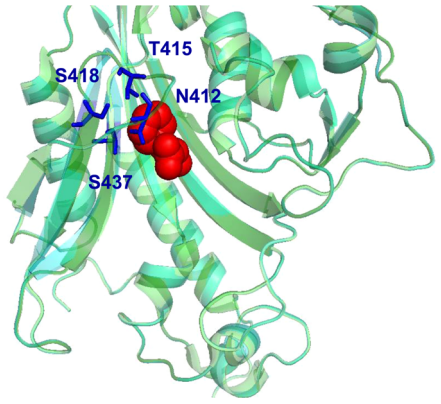


Figure 2. Phe substrate (red) and four residues (blue) within 7 Å of the *para*-position of the phenyl ring of the substrate inside the binding pocket of a homology model of yPheRS (cyan). The crystal structure of human cytosol phenylalanyl-tRNA synthetase (PDB ID: 3L4G) (green), the template of homology modeling, is overlapped with the homology model of yPheRS. These four residues were subjected to mutagenesis to generate a yPheRS library.

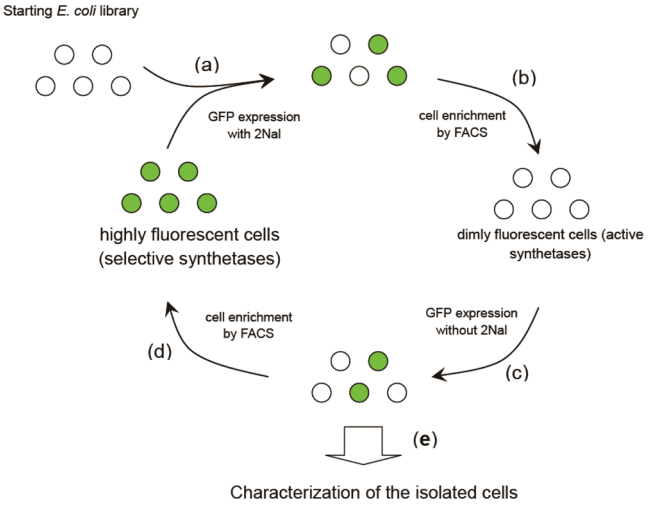


Figure 3. Screening scheme for yPheRS library. GFP6 was expressed in yPheRS expression library *E. coli* cells outfitted with ytRNA$^{Phe}_{AAA}$ and yPheRS library in the presence of 2Nal (a). Dimly fluorescent cells expressing GFP6 expected to be active yPheRS variants toward 2Nal were enriched by FACS (b). GFP6 was expressed in the enriched cells in the absence of 2Nal (c). Highly fluorescent cells expressing GFP6 expected to be selective yPheRS variants were enriched by FACS (d). After two rounds of screening, ten colonies were isolated from the enriched cells and characterized (e).

structure of GFP6 (data not shown). For a positive screening, the yPheRS library was induced with 1 mM IPTG to express GFP6 in minimal medium supplemented with 3 mM 2Nal. The fluorescence histograms of yPheRS library cells expressing GFP6 before and after the induction were obtained by flow cytometry (Figure 4a and b). The gate in the fluorescence channel was set to collect bottom 1% of cells. In the positive screening, any yPheRS variant that efficiently activates Phe is excluded, because those variants can generate an intact GFP6 with full fluorescence intensity. The 5×10^4 cells collected from the first positive screening (L_01_P01) were regrown in 2× YT medium and subjected to a negative screening to enrich cells containing yPheRS variants selective toward 2Nal (Figure 4c). The negative screening was based on the fact that

dx.doi.org/10.1021/sb500309r | ACS Synth. Biol. 2015, 4, 634–643

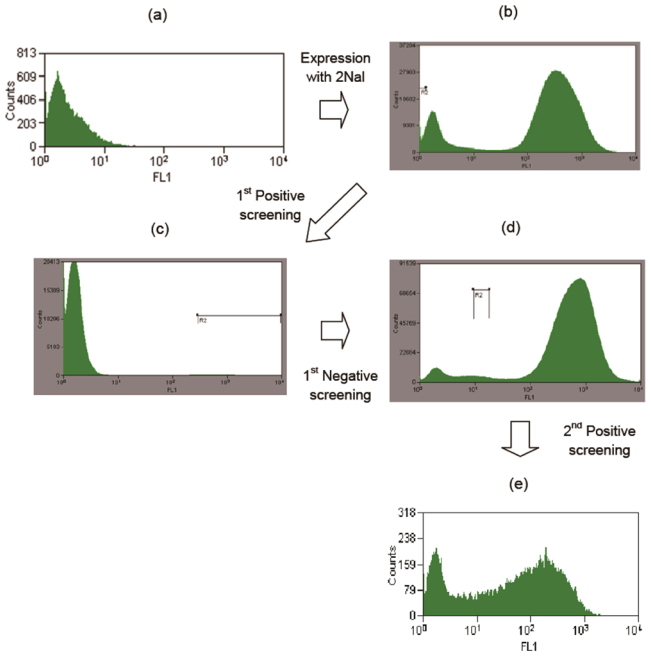


Figure 4. Actual yPheRS library screening results. yPheRS expression library (L_01) prior to expression of GFP6 (a). yPheRS library (L_01) expressing GFP6 in the presence of 2Nal (b). The enriched cells after first positive screening (L_01_P01) (c). The enriched cells after first negative screening (L_01_P01_N01) (d). The enriched cells after second positive screening (L_01_P02_N01) (e).

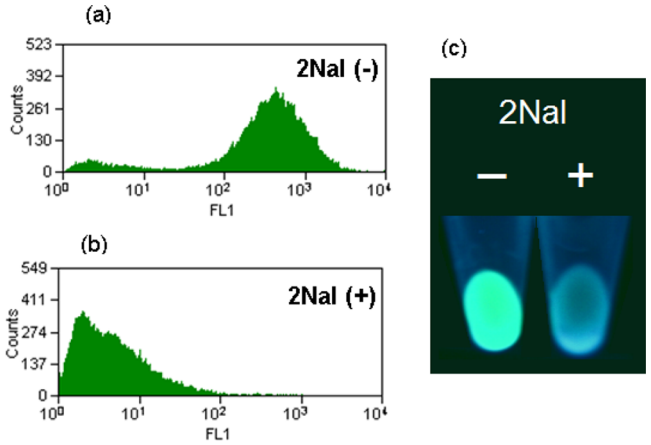


Figure 5. Fluorescence intensities of the clone (L_01_P02_N01_c2) expressing GFP analyzed by flow cytometry (a and b). Visual comparison of fluorescence of the clone (L_01_P02_N01_c2) expressing GFP6 with or without 2Nal (c). The expression medium was supplemented with 18 amino acids (25 μg/mL), 5 μM Phe, 50 μM Trp, and 3 mM 2Nal (b and c right) or no 2Nal (a and c left).

synthetases highly selective toward 2Nal over natural amino acids will not misincorporate any natural amino acid except Phe at multiple Phe sites in GFP6. Therefore, by combining both the positive and negative screening methods, it is expected that yPheRS variants selective toward 2Nal will be enriched. For the negative screening, GFP6 was expressed in minimal medium in the absence of 2Nal. The top 1% of cells in the fluorescence channel was collected to obtain 5×10^4 cells. The collected cells (L_01_P01_N01) were regrown in 2× YT medium and entered the next round of screening (Figure 4d). One more positive screening was carried out to enrich active yPheRS variants toward 2Nal (Figure 4e). The gate in the fluorescence channel was set to collect 1% of cells for which fluorescence was slightly above background. A portion of the collected cells (L_01_P02_N01) were spread on agar plates containing suitable antibiotics and ten colonies (L_01_P2_N01_c1 ~ c10) were subjected to further characterization.

Characterization of the Selective yPheRS Variant. One out of ten clones (L_01_P02_N01_c2) showed the fluorescence characteristics expected for a selective yPheRS variant. The other clones were false positives having two peaks that coincidently overlapped with the gates for sorting in the histogram of fluorescence. We speculate that the multiple rounds of cultures led to modifications in the DNA sequence of the plasmid inside *E. coli*s or *E. coli* genetic information resulting in two populations with different intensities in cellular fluorescence. The choice of *E. coli* expression hosts with an enhanced genetic stability may avoid the generation of such false positives. In the absence of 2Nal, cells expressing GFP6 retained full intensity of fluorescence (Figure 5a). However, in the presence of 2Nal, cells expressing GFP6 showed substantially reduced intensity of fluorescence (Figure 5b). Similar trends were observed in the visual inspection of fluorescence of cell pellets (Figure 5c). The plasmid DNA coding the yPheRS variant selective toward 2Nal was isolated from the culture of the clone (L_01_P02_N01_c2) and subjected to further analysis. DNA sequencing analysis of the plasmid showed that there were mutations of all four positions (N412G, T415G, S418C, and S437F) in the binding pocket of yPheRS (Supporting Information Figure S2). The isolated yPheRS variant was designated yPheRS_naph.

Amino Acid Activation Analysis by yPheRS_naph. Activation of Phe, Trp, and 2Nal by both yPheRS (T415G) and yPheRS_naph were examined *in vitro*. The kinetic parameters are listed in Table 1. yPheRS (T415G) showed 10-fold higher activity toward Trp than 2Nal. However, the yPheRS_naph showed 6-fold higher activity toward 2Nal than Trp, which translates into 60-fold enhanced selectivity toward 2Nal (vs Trp) compared to yPheRS (T415G). The yPheRS_naph also showed 18-fold higher activity toward 2Nal than Phe, while yPheRS (T415G) showed 8-fold higher activity. Previously, we reported that rationally designed yPheRS (T415A) showed 10-fold higher activity toward pBrF than Trp. However, the activation of Phe by yPheRS (T415A) was comparable to that of pBrF. Therefore, yPheRS_naph is the first yPheRS variant that efficiently discriminates a Phe analog from all 20 natural amino acids. It is noteworthy that the kinetic efficiencies of yPheRS_naph on 2Nal is substantially lower than that of yPheRS (T415G), indicating the enhanced selectivity toward 2Nal was accompanied by a significant loss in the activity. Similar to our findings, it was reported that the F130S mutation in *E. coli* tyrosyl-tRNA synthetase (eTyrRS) resulted in the enhanced specificity toward azatyrosine but the significant loss in its activity.[32] According to the amino acid activation assay *in vitro*, wild-type eTyrRS exhibits the 36-fold higher kinetic efficiency toward Tyr over azatyrosine.[32] The eTyrRS variant with F130S mutation (eTyrRS (F130S)) exhibits the 19-fold higher kinetic efficiency toward Tyr over azatyrosine indicating 2-fold enhanced substrate specificity toward azatyrosine. Despite the moderate substrate specificity enhancement, the kinetic efficiencies of eTyrRS (F130S) toward Tyr and azatyrosine decreased 11-fold and 6-fold, respectively, compared to those of wild-type eTyrRS.[32] Therefore, we speculate

dx.doi.org/10.1021/sb500309r | *ACS Synth. Biol.* 2015, 4, 634–643

Table 1. Kinetic Parameters for ATP-PPi Exchange by yPheRS (T415G) and yPheRS_naph

amino acid	enzyme	K_m (μM)[a]	k_{cat} (s^{-1})[a]	k_{cat}/K_m (M^{-1} s^{-1})	k_{cat}/K_m (rel)	ref
Phe	T415G	264 ± 42	0.05 ± 0.002	189 ± 30	**1**[b]	22
Trp	T415G	20 ± 5.0	0.30 ± 006	15 000 ± 4000	**79**[b]	22
2Nal	T415G	27 ± 2.0	0.04 ± 0.001	1481 ± 100	**7.8**[b]	22
Phe	Naph	11 000 ± 2700	0.0095 ± 0.0021	0.863 ± 0.007	**1**[c]	this study
Trp	Naph	1400 ± 600	0.0035 ± 0.0009	2.5 ± 0.4	**2.9**[c]	this study
2Nal	Naph	2000 ± 700	0.030 ± 0.018	15.5 ± 4.2	**18**[c]	this study

[a]Values ± regression standard error. [b]Relative to k_{cat}/K_m for Phe by yPheRS (T415G). [c]Relative to k_{cat}/K_m for Phe by yPheRS_naph.

that the engineering of synthetases to enhance both substrate specificity and kinetic efficiency is not trivial.

The poor binding of Phe and Trp by yPheRS_naph may be understood by loss of favorable interactions between the aromatic rings of the substrate and active-site residues of the synthetase. In a crystal structure of tPheRS,[31] the side chain of Phe258 (which corresponds to Asn412 in yPheRS) makes direct contact with the phenyl ring of the substrate. Since the interaction makes the substrate recognition highly specific and very favorable energetically, we can readily assume that N412G mutation in yPheRS would lead to poor binding of Phe and Trp by the PheRS. Analysis of fluorescence changes revealed that the omission of the N412G mutation in yPheRS_naph leads to misincorporation of Trp but retained high level incorporation of 2Nal (Figure 6). Therefore, N412G mutation

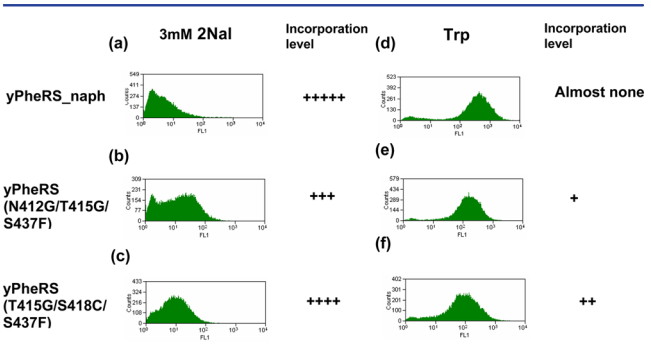


Figure 6. Fluorescence of cells expressing GFP6 in the presence of 2Nal with various yPheRS variants; yPheRS_naph (a); yPheRS (N412G/T415G/S437F) (b); yPheRS (T415G/S418C/S437F) (c). Fluorescence of cells expressing GFP6 in the absence of 2Nal with various yPheRS variants; yPheRS_naph (d); yPheRS (N412G/T415G/S437F) (e); yPheRS (T415G/S418C/S437F) (f). Reduction of cellular fluorescence can be explained by 2Nal incorporation (a, b, and c) or Trp misincorporation (d, e, and f). Incorporation level of 2Nal or Trp is represented by the number of '+'. The genes of the yPheRS variant containing the three mutations were prepared by PCR mutagenesis using the plasmid encoding the gene of yPheRS_naph. DHF expression hosts expressing the yPheRS variant and $ytRNA^{Phe}_{AAA}$ were induced to express GFP6 in MM18_FW medium supplemented with 2.5 μM Phe, 50 μM Trp, and 3 mM 2Nal (a, b, and c) or no 2Nal (d, e, and f).

was thought to play a key role in discriminating 2Nal from both Phe and Trp. However, the omission of S418C mutation significantly reduced the fluorescence loss upon incorporation of 2Nal indicating a weaker activity toward 2Nal than yPheRS_naph (Figure 6). Therefore, the S418C mutation was thought to be critical for the activation of 2Nal.

Elimination of Trp Misincorporation at UUU Codons by yPheRS_naph *In Vivo*. AFW [pQE16_mDHFR_yPheRS_naph/pREP4_$ytRNA^{Phe}$_AAA] cells were induced to express mDHFR in MM18_FW medium supplemented with 2.5 μM Phe, 50 μM Trp and 3 mM 2Nal. Peptide **2** (residues 140−144; $LF_{UUU}VTR$), one of the tryptic fragments of mDHFR, contains a Phe residue encoded as UUU. Occupancy of the UUU site in peptide **2** was investigated by LC-MS analysis. The results revealed that use of yPheRS_naph completely eliminated the misincorporation of Trp at the UUU site (Figure 1e), while yPheRS (T415G) allowed misincorporation of Trp at the UUU site (Figure 1d). Similar results were obtained by N-terminal sequencing of purified GFP6 (2UUU) variant containing a Phe site encoded by UUU codon at position 2. DHF [pQE9_GFP6_lacI_yPheRS_naph/pREP4_$ytRNA^{Phe}$_AAA] cells were induced to express GFP6 (2UUU) in MM18_FW medium supplemented with 5 μM Phe, 50 μM Trp and 3 mM 2Nal. 80% and 20% of the UUU codon at the second position of GFP6 (2UUU) were decoded as 2Nal and Phe, respectively, but Trp was not detected at this position (Table 2). These results suggest that yPheRS_naph has a high selectivity toward 2Nal over Trp.

Table 2. Occupancy of UUU and UAG Codons by Various Amino Acids

		occupancy of codon (%)[a]			
codon	ytRNA	2Nal	Phe	Trp	Lys
UUU[b]	$ytRNA^{Phe}_{AAA}$	80	20	ND[c]	ND
UAG[d]	$ytRNA^{Phe}_{CUA\_UG}$	98	ND	ND	2

[a]UUU and UAG codon occupancy was determined by N-terminal protein sequencing and LC-MS analysis, respectively. [b]The second position in the amino acid sequence of GFP6. [c]Not detected. [d]The 38th position in the amino acid sequence of mDHFR (38Am).

Site-Specific Incorporation of 2Nal into a single site of mDHFR *In Vivo*. In order to further evaluate the substrate specificity of yPheRS, site-specific incorporation of 2Nal into a single site of mDHFR was investigated. An amber codon was introduced to position 38 of mDHFR (mDHFR (38Am)). Site-specific incorporation of 2Nal into a single site of mDHFR (38Am) was achieved by AFWK [pQE16_mDHFR_yPheRS_naph/pREP4_$ytRNA^{Phe}$_UG] in minimal medium supplemented with 25 μg/mL 17 amino acids (MM17_FWK), 50 μM Phe, 50 μM Trp, 50 μM Lys, and 3 mM 2Nal. MALDI-MS analysis of tryptic digests of mDHFR (38Am) revealed that 2Nal was dominant at the amber site (Figure 7a). Neither Trp nor Phe was detected, confirming the high selectivity of yPheRS_naph toward 2Nal. Furthermore, LC-MS analysis of peptide **2** variants showed that the fidelity of 2Nal at an amber site is greater than 98% (Table 2). In order to investigate whether yPheRS_naph is very selective toward 2Nal over natural amino acids, nutrient-rich 2× YT medium was used for the expression of mDHFR (38Am) in AFWK [pQE16_mDHFR_yPheRS_naph/pREP4_$ytRNA^{Phe}$_UG]. The incorpora-

dx.doi.org/10.1021/sb500309r | ACS Synth. Biol. 2015, 4, 634−643

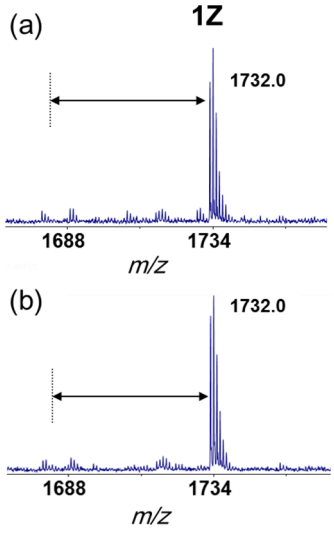


Figure 7. MALDI-TOF spectra of tryptic digests of mDHFR. Peptide **1Z** has 2Nal at position 38 (residues 26−39; NGDLPWPPLRNEZK). The modified mDHFR was synthesized in a Phe/Trp autoxtrophic expression host outfitted with ytRNA$^{Phe}_{CUA_UG}$ and yPheRS_naph in the minimal expression medium supplemented with 18 amino acids (25 μg/mL), 50 μM Phe, 50 μM Trp, 3 mM 2Nal (a) or 2× YT medium supplemented with 3 mM 2Nal (b).

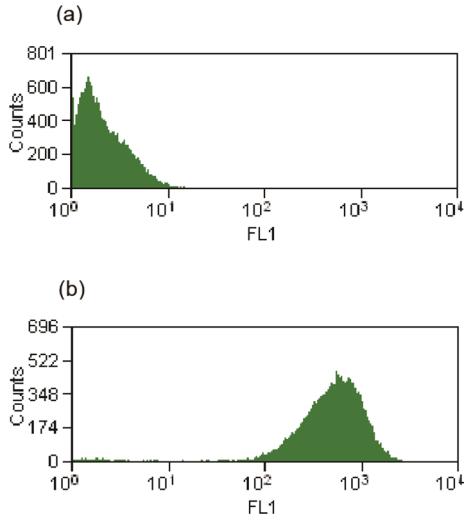


Figure 8. Fluorescence intensities of cells under uninduced conditions (a); cells expressing full length GFP6 (158Am) in the presence of 2Nal (b). The DHF *E. coli* strains were outfitted with ytRNA$^{Phe}_{CUA_UG}$ and yPheRS_naph. The expression medium was supplemented with 18 amino acids (25 μg/mL), 50 μM Phe, 50 μM Trp, and 3 mM 2Nal.

tion level of 2Nal at the amber site is comparable to that in minimal medium (Figure 7b). Furthermore, the yield of purified mDHFR (38Am) containing 2Nal in 2× YT medium was 4.3 mg/L, which is almost 3-fold greater than that in minimal medium (1.5 mg/L).

Therefore, the combination of *in vitro* activation assay results and 2Nal incorporation data in bacterial cells supports that yPheRS_naph possesses the great substrate specificity toward 2Nal over any natural amino acids. Furthermore, these results suggest that the high-throughput screening method was successfully developed to select yPheRS variants specific to a Phe analog from a large pool of yPheRS variants.

Site-Specific Incorporation of 2Nal into a Single Site of GFP6 *In Vivo*. Site-specific incorporation of 2Nal into a single site of GFP6 was investigated to determine whether a correctly folded fluorescent protein can be obtained upon the incorporation of a Phe analog. An AAG (Lys) codon was changed to an amber codon in position 158 in GFP6, which is known to be permissive to replacement of Lys, with other natural amino acids or to circular permutation.[33] DHF [pQE9_GFP6 (158Am)_lacI_yPheRS_naph/pREP4_ytRNA^{Phe}_UG] cells were used to express GFP6 (158Am) in MM17_FWK medium supplemented with 50 μM Phe, 50 μM Trp, 50 μM Lys, and 3 mM 2Nal. The mean fluorescence of cells expressing full length of GFP6 (158Am) was 280-fold higher than that in uninduced conditions, which clearly indicated that 2Nal can be inserted into GFP6 with a significant fluorescence (Figure 8).

Conclusions. The high-throughput screening of a yPheRS saturation library successfully led to an yPheRS variant (yPheRS_naph) with high substrate specificity toward 2Nal over any natural amino acids, strongly supporting that yPheRS can selectively recognize Phe analogs. The yPheRS_naph has mutations of all four positions (N412G, T415G, S418C, and S437F). The successful identification of yPheRS_naph variant validates the utility of the screening method based on the reduction of GFP fluorescence upon incorporation of a Phe analog into multiple sites. This will prove useful to screen the yPheRS libraries for other Phe analogs, and furthermore to screen other aminoacyl-tRNA synthetase libraries. Use of the yPheRS_naph variant resulted in the elimination of misincorporation of Trp, which was detected previously in experiments with the rationally designed yPheRS (T415G) variant. Combined use of yPheRS_naph and ytRNA$^{Phe}_{CUA_UG}$ realized high-fidelity (ca. 98%) incorporation of 2Nal into a protein in response to an amber codon. Since 2Nal has a hydrophobic side chain bulkier than any other natural amino acids, site-specific incorporation of 2Nal was successfully used to fine-tune the active site of mDHFR in order to modulate the binding affinity to an inhibitor.[40] The high-fidelity incorporation of 2Nal into a protein even in rich culture medium allows a high yield biosynthesis of a recombinant protein containing 2Nal in a site-specific manner.

METHODS

Materials. Unless otherwise noted, reagents were obtained from commercial suppliers and were used without further purification. 2Nal was purchased from Chem-Impex (Wood Dale, IL). Restriction enzymes were obtained from New England Biolabs (Beverly, MA). Quikchange mutagenesis kits were purchased from Stratagene (La Jolla, CA). Nickel-nitrilotriacetic acid affinity columns and plasmid pREP4 were obtained from Qiagen (Valencia, CA). DNA primers were obtained from Integrated DNA Technologies (Coralville, IA) and Operon Technologies (Huntsville, AL). Sequences of the DNA primers used in this research are listed in Supporting Information Table S1.

Preparation of *E. coli* Hosts. Preparation of the Phe/Trp double auxotrophic strain (AFW; (K10, Hfr(Cavalli), *pheS13, rel-1 tonA22, thi, T2*^{R}*, pheA18, trpB114*) and the Phe/Trp/Lys triple auxotrophic strain (AFWK; K10, Hfr(Cavalli), *pheS13, rel-1 tonA22, thi, T2*^{R}*, pheA18, trpB114, lysA*) was described previously.[20] A Phe auxotrophic derivative of DH10B (Stratagene) *E. coli* strain was prepared by chemical mutagenesis[34] and designated DHF (F$^-$, *endA1, recA1, galE15,*

dx.doi.org/10.1021/sb500309r | ACS Synth. Biol. 2015, 4, 634−643

galK16, *nupG*, *rpsL*, Δ*lacX74*, Φ*80lacZ*Δ*M15*, *araD139*, Δ(*ara,leu*)*7697*, *mcrA*, Δ(*mrr-hsdRMS-mcrBC*), λ^-, *pheA18*).

Plasmid Construction for Reporter Gene Expression. The GFP_{UV} gene was amplified from pGFPuv (Clontech, Palo Alto, CA) using two primers (GFP1 and GFP2) containing a *Pst*I restriction site. The amplified GFPuv gene was inserted into pQE9 (Qiagen) at the *Pst*I site to generate pQE9_GFP_{UV}. Since GFP_{UV} has excitation maxima at 475 and 395 nm,[35] it was mutated into an EGFP variant (GFP3)[35] with excitation maximum at 488 nm suitable for FACS analysis. A series of PCR mutagenesis reactions were performed at four residues (F64L, S65T, S99F, and T153M) using three pairs of complementary primers (F64LS65T_f/F64LS65T_r; S99F_f/S99F_r; T153M_f/T153M_r). The GFP3 gene expression cassette was inserted into pQE16_mDHFR_yPheRS (T415G)[21] between the *Aat*II and *Nhe*I restriction sites to generate pQE9_GFP3_yPheRS (T415G). GFP3 has 12 Phe residues, of which six are encoded as UUU (Supporting Information Figure S3a). A GFP variant (GFP6) was prepared by replacing all UUC Phe codons and one CUG Leu codon (at position 64) with UUU codons via gene fragment reassembly using 16 primers (primers 1−16) (Supporting Information Figure S3b). The GFP6 expression cassette was removed by restriction digestion with the *Aat*II and *Nhe*I, and inserted into pQE9_GFP3_yPheRS (T415G) between the *Aat*II and *Nhe*I sites to yield pQE9_GFP6_yPheRS (T415G). In order to introduce an extra copy of the *lac*I gene into pQE9_GFP6_yPheRS (T415G), the *lac*I gene expression cassette was amplified from pREP4 (Qiagen) with two primers (*lac*I_*Afl*II_f and *lac*I_*Sac*II_r). *Afl*II and *Sac*II restriction sites were introduced into pQE9_GFP6_yPheRS (T415G) by PCR reactions with two primers (pQE_*Afl*II_f and pQE_*Sac*II_r). The amplified *lac*I gene expression cassette was inserted into pQE9_GFP6_yPheRS (T415G) between the *Afl*II and *Sac*II sites to generate pQE9_GFP6_lacI_yPheRS (T415G). The DNA and amino acid sequences of the proteins used in this work are included in the Supporting Information.

Construction of a Homology Model of yPheRS. A homology model of yPheRS was constructed using a protein homology modeling server, SWISS-MODEL.[36] Briefly, the evolutionary related protein structures matching the α-subunit of yPheRS were searched in the SWISS-MODEL template library by Blast[37] and HHBlits.[38] Then, using the crystal structure of human cytosolic phenylalanyl-tRNA synthetase (PDB ID: 3L4G)[39] identified as the most appropriate template, a homology structural model of yPheRS was generated (Figure 2).

Construction of yPheRS Library. Four residues (N412, T415, S418, and S437) in the active site of yPheRS were saturated by two step PCR mutagenesis. First, NNK (N = A/U/G/C; K = G/U) degenerate codons were introduced into the 437th position in the α-subunit of yPheRS by PCR mutagenesis with two complementary primers (437_f and 437_r) using pQE9_GFP6_lacI_yPheRS (T415G) as a template. The PCR product was digested by *Dpn*I for 1 h at 37 °C and transformed into XL-1 blue (Stratagene) competent cells. The plasmids isolated from the culture of transformants were used as a template for the second PCR mutagenesis with one pair of complementary primers (412_415_418_f and 412_415_418_r) to saturate the 412th, 415th, and 418th positions in the α-subunit of yPheRS. The PCR product was digested by *Dpn*I for 1 h at 37 °C and desalted on a spin column. The eluted plasmids were transformed into ElectroTen-Blue electro-competent cells (Stratagene) according to the manufacturer's protocol, and plasmids were isolated from the culture of pooled transformants. The plasmid fragments encoding yPheRS were removed by digestion with *Nsi*I and *Bgl*II restriction enzymes and ligated to the large fragments of pQE9_GFP6_lacI_yPheRS (T415G) obtained by digestion with *Nsi*I and *Bgl*II. The ligated plasmids were desalted and transformed into ElectroTen-Blue electro-competent cells to generate six million transformants. The yPheRS library plasmids (P_yFS_20) were isolated from the culture of pooled transformants.

Construction of Expression Library. pREP4_ytRNA_AAA, prepared previously,[21] was transformed into DHF electro-competent cells to generate DHF_AAA cells. The P_yFS_20 library was transformed into DHF_AAA electro-competent cells to generate the yPheRS expression library (L_01) consisting of 10 million transformants. The transformants were transferred to 0.5 L 2× YT medium with 200 μg/mL ampicillin and 35 μg/mL kanamycin. When OD_{600} reached 1.0, glycerol stocks were prepared and kept in a −80 °C freezer.

Screening of yPheRS Library. Glycerol stock of the expression library (0.5 mL) was inoculated into 100 mL of M9 minimal medium supplemented with glucose, thiamine, $MgSO_4$, $CaCl_2$, 20 amino acids (at 25 μg/mL), antibiotics (35 μg/mL of kanamycin and 200 μg/mL of ampicillin). When the culture reached an OD_{600} of 0.6−0.8, cells were spun down, washed twice with ice-cold 0.9% NaCl, and shifted to expression medium supplemented with 18 amino acids (at 25 μg/mL) and the indicated concentrations of Phe, Trp, and 2Nal. Expression of GFP was induced by addition of 1 mM IPTG. After 3 h, 1 mL of the culture (based on OD of 1.0) was collected and washed twice with 0.5 mL of PBS (pH 7.4). Cells (300 μL) were diluted with 3 mL of distilled water and then subjected to cell sorting using a MoFlo cell sorter (DakoCytomation, Ft. Collins, CO). The excitation and emission wavelengths were 488 and 525 nm, respectively. The library was subjected to both positive and negative screening (Figure 3). FACS gates were set based upon FSC/SSC and FL. In order to enrich yPheRS variants capable of 2Nal activation and thus multiple-site-specific incorporation into GFP, 5×10^4 weakly fluorescent cells were collected by positive screening in the first round. The enriched cells (L_01_P01) were incubated in 2× YT medium containing 200 μg/mL of ampicillin and 35 μg/mL of kanamycin. When OD_{600} reached 1.0, glycerol stocks of the cells were prepared. 0.1 mL of the glycerol stocks was inoculated into 20 mL of M9 minimal medium containing 20 amino acids. In order to enrich yPheRS variants toward 2Nal, the top 1% of highly fluorescent cells was collected as a negative screening (L_01_P01_N01). The negative screening was applied to enrich the fluorescent cell population to exclude cells that cannot express GFP. After one more positive screening was performed, the collected cells (L_01_P02_N01) were spread onto a 2× YT agar plate containing 35 μg/mL of kanamycin and 200 μg/mL of ampicillin. After overnight incubation at 37 °C, ten single colonies were isolated and subjected to characterization.

Characterization of the Isolated Clones. The 10 single colonies (L_01_P02_N01_c1 ∼ c10) were transferred to 0.5 L 2× YT medium with 200 μg/mL ampicillin and 35 μg/mL kanamycin. When OD_{600} reached 1.0, glycerol stocks were prepared and kept in a −80 °C freezer. 200 μL of each glycerol stock was inoculated into minimal medium and incubated until

OD_{600} reached 0.6. The cultures were washed twice with 0.9% NaCl solution and resuspended with 20 mL of minimal medium supplemented with 18 amino acids, 50 μM Trp, and 5 μM Phe. The resuspended cells were divided into two fractions and transferred into two flasks. 2Nal (3 mM) was supplemented into one of them. GFP6 was expressed by adding 1 mM IPTG. After 3 h, 1 mL of the culture was collected, and washed twice with 0.5 mL of PBS (pH 7.4). 100 μL of cells were diluted with 3 mL of distilled water. Fluorescence intensities of the cells were analyzed by a MoFlo cell sorter. At least 20 000 events were collected in each measurement. Data were analyzed with Summit software (DakoCytomation). One clone (L_01_P02_N01_c2), showing dim fluorescence in the presence of 2Nal but high fluorescence in the absence of 2Nal, was identified and subjected to further characterization. The fluorescence images of the cell pellets of the clone (L_01_P02_N01_c2) were taken with a digital camera. The plasmid isolated from the positive clone was subjected to DNA sequencing, which revealed that the yPheRS variant selected contained four mutations (N412G, T415G, S418C, and S437F). The yPheRS variant and the plasmid encoding were designated yPheRS_naph and pQE9_GFP6_lacI_yPheRS_naph, respectively.

Amino Acid Activation Assay. The plasmid fragment encoding yPheRS was removed from pQE9_GFP6_lacI_yPheRS_naph by digestion with *Nsi*I and *Bgl*II, and then ligated with a large fragment of pQE32-yPheRS (T415G) generated by digestion with *Nsi*I and *Bgl*II to yield pQE32-yPheRS_naph. Expression and purification of yPheRS variants were described previously.[20] The kinetics of activation of amino acids by yPheRS (T415G) and yPheRS_naph were determined by the amino acid-dependent adenosine triphosphate (ATP)-[^{32}P]-pyrophosphate (PP_i) exchange assay. The assay buffer included 50 mM *N*-(2-hydroxethyl) piperazine-*N*′-(2-ethanesulfonic acid) (potassium-HEPES) (pH = 7.6), 20 mM $MgCl_2$, 1 mM dithiothreitol (DTT), 2 mM ATP, and 2 mM [^{32}P]- PP_i. The amino acid concentration varied from 100 nM to 10 mM, and the enzyme concentration varied from 100 nM to 800 nM. Detailed procedures of the assay were described previously.[20,22]

Construction of Plasmids and Expression Hosts for Incorporation of 2Nal at Phe Codons. The gene fragment containing four mutations in the binding pocket of yPheRS_naph was excised by *Nsi*I and *Bgl*II digestion of pQE9_GFP6_lacI_yPheRS_naph and inserted into pQE16_mDHFR_yPheRS (T415G) between the *Nsi*I and *Bgl*II sites to generate pQE16_mDHFR_yPheRS_naph. Both pQE16_mDHFR_yPheRS (T415G) and pQE16_mDHFR_yPheRS_naph were cotransformed with pREP4_ytRNA^{Phe}_AAA into AFW competent cells to generate AFW [pQE16_mDHFR_yPheRS (T415G)/pREP4_ytRNA^{Phe}_AAA] and AFW [pQE16_mDHFR_yPheRS_naph/pREP4_ytRNA^{Phe}_AAA], respectively. In order to express intact mDHFR, pQE16 (Qiagen) and pREP4 plasmids were cotransformed into AFW competent cells to generate AFW [pQE16/pREP4]. pQE9_GFP6_lacI_yPheRS_naph was transformed into DHF_AAA electro-competent cells to construct DHF [pQE9_GFP6_lacI_yPheRS_naph/pREP4_ytRNA^{Phe}_AAA]. In order to determine Phe codon occupancy by N-terminal amino acid sequencing, the AGA (Arg) codon in the second position of GFP6 was mutated to either a UUU codon by PCR mutagenesis. The PCR reactions were conducted with one pair of complementary primers (2F_UUU_f/2F_UUU_r) using pQE9_GFP6_lacI_yPheRS_naph as a template to generate pQE9_GFP6 (2UUU)_lacI_yPheRS_naph.

Site-Specific Incorporation of 2Nal into a Single Site of mDHFR and GFP6. pREP4_ytRNA^{Phe}_UG was prepared previously.[20] AFWK [pQE16_mDHFR (38Am)_yPheRS (T415G)/pREP4_ytRNA^{Phe}_UG] was prepared previously.[20] The yPheRS_naph gene fragment excised by *Nsi*I and *Bgl*II digestion of pQE9_GFP6_lacI_ PheRS_naph was inserted into pQE16_mDHFR (38Am)_yPheRS (T415G) between the *Nsi*I and *Bgl*II sites to generate pQE16_mDHFR (38Am)_yPheRS_naph. An AAG (Lys) codon was changed to a UAG codon at position 158 of GFP6 sequence in pQE9_GFP6_lacI_yPheRS_naph by PCR mutagenesis with one pair of complementary primers (K158_UAG_f/K158_UAG_r) to yield pQE9_GFP6 (158Am)_lacI_yPheRS_naph. Both pQE16_mDHFR (38Am)_yPheRS_naph and pQE9_GFP6 (158Am)_lacI_yPheRS_naph were cotransformed with pREP4_ytRNA^{Phe}_UG into AFWK and DHF competent cells to construct AFWK [pQE16_mDHFR (38Am)_yPheRS_naph/pREP4_ytRNA^{Phe}_UG] and DHF [pQE9_GFP6 (158Am)_lacI_yPheRS_naph/pREP4_ytRNA^{Phe}_UG], respectively.

Expression of mDHFR Variants and GFP Variants *In Vivo*. AFW expression strain cotransformed with pQE plasmid variants and pREP4 plasmid variants was grown in M9 minimal medium supplemented with 0.4 wt % glucose, 35 mg/L thiamine, 1 mM $MgSO_4$, 1 mM $CaCl_2$, 20 amino acids (at 25 μg/mL), 35 μg/mL kanamycin, and 200 μg/mL ampicillin. The overnight cultures of expression strains were diluted 20-fold in fresh M9 minimal medium and incubated at 37 °C. When cells reached OD of 0.8−1.0, cells were spun down and washed twice with cold 0.9% NaCl. The cultures were resuspended in fresh M9 minimal medium supplemented with 18 amino acids (25 μg/mL) and the indicated concentrations of Phe, Trp, and 2Nal. After 10 min incubation, 1 mM of isopropyl-β-D-thiogalactopyranoside (IPTG) was added to induce expression of GFP at 30 °C. After 4 h, cells were harvested and either kept at −80 °C or subjected to fluorescence measurement, according to the procedures described earlier (see Characterization of the Isolated Clones). Whole cell lysates were analyzed by SDS-PAGE. Due to slow growth of DHF and MPC390 expression hosts cotransformed with pQE plasmid variants and pREP4 plasmid variants, transformants were grown in 2× YT medium to prepare glycerol stocks first. Then glycerol stocks were inoculated into minimal medium supplemented with 20 amino acids (at 25 μg/mL) and incubated overnight at 37 °C. The remaining steps are similar to those for AF expression host.

Quantitative Analysis of Codon Occupancy. Quantitative analysis of codon occupancy was performed using either N-terminal protein sequencing or LC-MS analysis of tryptic digests. The GFP6 (2UUU) variants were expressed in minimal medium and purified by Ni-NTA affinity chromatography according to the manufacturer's protocol (Qiagen) under denaturing conditions. The purified GFP variants were subjected to N-terminal protein sequencing using a 492 cLC Procise protein microsequencer (Applied Biosystems, Foster City, CA). Occupancy of Phe codons in mDHFR was determined by LC-MS analysis. The mDHFR variants expressed in minimal medium were subjected to purification via Ni-NTA affinity chromatography according to the manufacturer's protocol (Qiagen) under denaturing conditions. After purification, expression levels of GFP and mDHFR were determined by UV absorbance at 280 nm using a calculated extinction coefficient of 20 010 cm^{-1} M^{-1} and 24 750 cm^{-1}

dx.doi.org/10.1021/sb500309r | ACS Synth. Biol. 2015, 4, 634−643

M^{-1}, respectively. The purified proteins were concentrated by ultrafiltration (Millipore); 10 μL of the concentrate was diluted into 90 μL of 75 mM $(NH_4)_2CO_3$ solution, and then, 1 μL of modified trypsin (Promega, 0.2 μg/μL) was added. Reaction was carried out for 2−4 h at 37 °C and quenched by adding 13 μL of 5% trifluoroacetic acid (TFA) solution. The solution was then directly subjected to LC-MS analysis conducted on a LCT Premier XE MICROMASS MS system (MS Technologies, Montgomery Village, MD) with Acquity UPLCTM system (Waters, Milford, MA). Tryptic digests were separated by Acquity BEH300 C18 column (1.7 μm, 300 Å, 2.1 × 50 mm) using a gradient of 5−95% of solvent B (90% of acetonitrile/10% of 0.1% formic acid solution) and solvent A (2% of acetonitrile/98% of 0.1% formic acid solution) in 10 min. The column eluent was transferred to the electrospray source and mass spectra were recorded.

■ ASSOCIATED CONTENT

Supporting Information

Names and sequences of the PCR primers, DNA and amino acid sequences of the proteins; chemical structures of the amino acids; four mutations in the yPheRS homology model; and structures of GFP3 and GFP6. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +82 62-715-2312. Fax: +82 62-715-2304. E-mail: inchan@gist.ac.kr.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Dr. M. Shahgholi and Dr. J. Zhou for mass spectrometry studies. We also thank Dr. F. Rusnak for help with N-terminal sequencing. We are grateful to Dr. A. J. Link for help with the cell sorter. We appreciate Dr. I. C. Tanrikulu and Dr. P. M. Kekenes-Huskey for help with analysis of the PheRS crystal structure. We also thank Dr. D. A. Tirrell for thoughtful discussion on the results. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea Government (MSIP) (2014R1A2A1A11050322), the Korea CCS R&D Center (KCRC) grant (2013M1A8A1038187), and an NIH grant (GM R01 GM062523).

■ REFERENCES

(1) Jones, D. H., Cellitti, S. E., Hao, X. S., Zhang, Q., Jahnz, M., Summerer, D., Schultz, P. G., Uno, T., and Geierstanger, B. H. (2010) Site-specific labeling of proteins with NMR-active unnatural amino acids. *J. Biomol. NMR 46*, 89−100.

(2) Cho, H., Daniel, T., Buechler, Y. J., Litzinger, D. C., Maio, Z., Putnam, A.-M. H., Kraynov, V. S., Sim, B.-C., Bussell, S., Javahishvili, T., Kaphle, S., Viramontes, G., Ong, M., Chu, S., GC, B., Lieu, R., Knudsen, N., Castiglioni, P., Norman, T. C., Axelrod, D. W., Hoffman, A. R., Schultz, P. G., DiMarchi, R. D., and Kimmel, B. E. (2011) Optimized clinical performance of growth hormone with an expanded genetic code. *Proc. Natl. Acad. Sci. U.S.A. 108*, 9060−9065.

(3) Bae, J. H., Alefelder, S., Kaiser, J. T., Friedrich, R., Moroder, L., Huber, R., and Budisa, N. (2001) Incorporation of *β*-selenolo[3,2-*b*]pyrrolyl-alanine into proteins for phase determination in protein X-ray crystallography. *J. Mol. Biol. 309*, 925−936.

(4) Bae, J. H., Rubini, M., Jung, G., Wiegand, G., Seifert, M. H. J., Azim, M. K., Kim, J. S., Zumbusch, A., Holak, T. A., Moroder, L., Huber, R., and Budisa, N. (2003) Expansion of the genetic code enables design of a novel "gold" class of green fluorescent proteins. *J. Mol. Biol. 328*, 1071−1081.

(5) Beatty, K. E., Xie, F., Wang, Q., and Tirrell, D. A. (2005) Selective dye-labeling of newly synthesized proteins in bacterial cells. *J. Am. Chem. Soc. 127*, 14150−14151.

(6) Dieterich, D. C., Link, A. J., Graumann, J., Tirrell, D. A., and Schuman, E. M. (2006) Selective identification of newly synthesized proteins in mammalian cells using bioorthogonal noncanonical amino acid tagging (BONCAT). *Proc. Natl. Acad. Sci. U.S.A. 103*, 9482−9487.

(7) Hong, S. H., Ntai, I., Haimovich, A. D., Kelleher, N. L., Isaacs, F. J., and Jewett, M. C. (2013) Cell-free protein synthesis from a release factor 1 deficient *Escherichia coli* activates efficient and multiple site-specific nonstandard amino acid incorporation. *ACS Synth. Biol. 3*, 398−409.

(8) Hoesl, M. G., and Budisa, N. (2011) *In vivo* incorporation of multiple noncanonical amino acids into proteins. *Angew. Chem., Int. Ed. 50*, 2896−2902.

(9) Lim, S. I., Mizuta, Y., Takasu, A., Hahn, Y. S., Kim, Y. H., and Kwon, I. (2013) Site-specific fatty acid-conjugation to prolong protein half-life *in vivo*. *J. Controlled Release 170*, 219−225.

(10) Wang, N., Ju, T., Niu, W., and Guo, J. (2014) Fine-tuning interaction between aminoacyl-tRNA synthetase and tRNA for efficient synthesis of proteins containing unnatural amino acids. *ACS Synth. Biol.*, DOI: 10.1021/sb500195w.

(11) Anderson, J. C., and Schultz, P. G. (2003) Adaptation of an orthogonal archaeal leucyl-tRNA and synthetase pair for four-base, amber, and opal suppression. *Biochemistry 42*, 9598−9608.

(12) Pastrnak, M., Magliery, T. J., and Schultz, P. G. (2000) A new orthogonal suppressor tRNA/aminoacyl-tRNA synthetase pair for evolving an organism with an expanded genetic code. *Helv. Chim. Acta 83*, 2277−2286.

(13) Wang, L., and Schultz, P. G. (2001) A general approach for the generation of orthogonal tRNAs. *Chem. Biol. 8*, 883−890.

(14) Bianco, A., Townsley, F. M., Greiss, S., Lang, K., and Chin, J. W. (2012) *Expanding the genetic code of* Drosophila melanogaster. *Nat. Chem. Biol. 8*, 748−750.

(15) Greiss, S., and Chin, J. W. (2011) Expanding the genetic code of an animal. *J. Am. Chem. Soc. 133*, 14196−14199.

(16) Liu, C. C., and Schultz, P. G. (2010) Adding new chemistries to the genetic code. *Annu. Rev. Biochem. 79*, 413−444.

(17) Bose, M., Groff, D., Xie, J. M., Brustad, E., and Schultz, P. G. (2006) The incorporation of a photoisomerizable amino acid into proteins in *E. coli*. *J. Am. Chem. Soc. 128*, 388−389.

(18) Santoro, S. W., Wang, L., Herberich, B., King, D. S., and Schultz, P. G. (2002) An efficient system for the evolution of aminoacyl-tRNA synthetase specificity. *Nat. Biotechnol. 20*, 1044−1048.

(19) Chin, J. W., Martin, A. B., King, D. S., Wang, L., and Schultz, P. G. (2002) *Addition of a photocrosslinking amino acid to the genetic code of* Escherichia coli. *Proc. Natl. Aca. Sci. U.S.A 99*, 11020−11024.

(20) Kwon, I., Wang, P., and Tirrell, D. A. (2006) Design of a bacterial host for site-specific incorporation of *p*-bromophenylalanine into recombinant proteins. *J. Am. Chem. Soc. 128*, 11778−11783.

(21) Kwon, I., Kirshenbaum, K., and Tirrell, D. A. (2003) Breaking the degeneracy of the genetic code. *J. Am. Chem. Soc. 125*, 7512−7513.

(22) Wang, P. (2003) Expanding the biosynthetic capacity of the aminoacyl-tRNA synthetases. PhD thesis, California Institute of Technology, Pasadena.

(23) Kwon, I., and Tirrell, D. A. (2007) Site-specific incorporation of tryptophan analogues into recombinant proteins in bacterial cells. *J. Am. Chem. Soc. 129*, 10431−10437.

(24) Chin, J. W., Cropp, T. A., Anderson, J. C., Mukherji, M., Zhang, Z. W., and Schultz, P. G. (2003) An expanded eukaryotic genetic code. *Science 301*, 964−967.

(25) Wang, L., Brock, A., Herberich, B., and Schultz, P. G. (2001) Expanding the genetic code of *Escherichia coli*. *Science 292*, 498−500.

(26) Neumann, H., Wang, K., Davis, L., Garcia-Alai, M., and Chin, J. W. (2010) Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature 464*, 441−444.

(27) Yoo, T. H., Link, A. J., and Tirrell, D. A. (2007) Evolution of a fluorinated green fluorescent protein. *Proc. Natl. Acad. Sci. U.S.A. 104*, 13887−13890.
(28) Pedelacq, J. D., Cabantous, S., Tran, T., Terwilliger, T. C., and Waldo, G. S. (2006) Engineering and characterization of a superfolder green fluorescent protein. *Nat. Biotechnol. 24*, 79−88.
(29) Waldo, G. S., Standish, B. M., Berendzen, J., and Terwilliger, T. C. (1999) Rapid protein-folding assay using green fluorescent protein. *Nat. Biotechnol. 17*, 691−695.
(30) Gregoire, S., and Kwon, I. (2012) A revisited folding reporter for quantitative assay of protein misfolding and aggregation in mammalian cells. *Biotechnol. J. 7*, 1297−1307.
(31) Fishman, R., Ankilova, V., Moor, N., and Safro, M. (2001) Structure at 2.6 A resolution of phenylalanyl-tRNA synthetase complexed with phenylalanyl-adenylate in the presence of manganese. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 57*, 1534−1544.
(32) Hamano-Takaku, F., Iwama, T., Saito-Yano, S., Takaku, K., Monden, Y., Kitabatake, M., Soll, D., and Nishimura, S. (2000) A mutant *Escherichia coli* tyrosyl-tRNA synthetase utilizes the unnatural amino acid azatyrosine more efficiently than tyrosine. *J. Biol. Chem. 275*, 40324−40328.
(33) Ghosh, I., Hamilton, A. D., and Regan, L. (2000) Antiparallel leucine zipper-directed protein reassembly: Application to the green fluorescent protein. *J. Am. Chem. Soc. 122*, 5658−5659.
(34) Ishii, Y., and Kondo, S. (1975) Comparative analysis of deletion and base-change mutabilities of *Escherichia coli* B-strains differing in DNA-repair capacity (wild type, uvra-, pola-, reca-) by various mutagens. *Mutat. Res. 27*, 27−44.
(35) Tsien, R. Y. (1998) The green fluorescent protein. *Annu. Rev. Biochem. 67*, 509−544.
(36) Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006) The SWISS-MODEL workspace: A web-based environment for protein structure homology modelling. *Bioinformatics 22*, 195−201.
(37) Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res. 25*, 3389−3402.
(38) Remmert, M., Biegert, A., Hauser, A., and Soding, J. (2012) HHblits: Lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nat. Methods 9*, 173−175.
(39) Finarov, I., Moor, N., Kessler, N., Klipcan, L., Safro, M. G. Structure of human cytosolic phenylalanyl-tRNA synthetase: Evidence for Kingdom-specific design of the active sites and tRNA binding patterns, *Structure 18*, 343−353.
(40) Zheng, S., and Kwon, I. (2013) Controlling enzyme inhibition using an expanded set of genetically encoded amino acids. *Biotechnol. Bioeng. 110*, 2361−2370.