

UC Berkeley

UC Berkeley Previously Published Works

Title

Genetic Encoding of Three Distinct Noncanonical Amino Acids Using Reprogrammed Initiator and Nonsense Codons

Permalink

<https://escholarship.org/uc/item/9s91b0zb>

Journal

ACS Chemical Biology, 16(4)

ISSN

1554-8929

Authors

Tharp, Jeffery M
Vargas-Rodriguez, Oscar
Schepartz, Alanna
et al.

Publication Date

2021-04-16

DOI

10.1021/acscchembio.1c00120

Peer reviewed



HHS Public Access

Author manuscript

ACS Chem Biol. Author manuscript; available in PMC 2021 August 04.

Published in final edited form as:

ACS Chem Biol. 2021 April 16; 16(4): 766–774. doi:10.1021/acscchembio.1c00120.

Genetic Encoding of Three Distinct Noncanonical Amino Acids Using Reprogrammed Initiator and Nonsense Codons

Jeffery M. Tharp,

Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06520, United States

Oscar Vargas-Rodriguez,

Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06520, United States

Alanna Schepartz,

Department of Chemistry and Department of Molecular and Cell Biology, University of California, Berkeley, California 94705, United States

Dieter Söll

Department of Molecular Biophysics and Biochemistry and Department of Chemistry, Yale University, New Haven, Connecticut 06520, United States

Abstract

We recently described an orthogonal initiator tRNA ($\hat{t}RNA^{Ty2}$) that can initiate protein synthesis with noncanonical amino acids (ncAAs) in response to the UAG nonsense codon. Here, we report that a mutant of $\hat{t}RNA^{Ty2}$ ($\hat{t}RNA^{Ty2}_{AUA}$) can efficiently initiate translation in response to the UAU tyrosine codon, giving rise to proteins with an ncAA at their N-terminus. We show that, in cells expressing $\hat{t}RNA^{Ty2}_{AUA}$, UAU can function as a dual-use codon that selectively encodes ncAAs at the initiating position and predominantly tyrosine at elongating positions. Using $\hat{t}RNA^{Ty2}_{AUA}$, in conjunction with its cognate tyrosyl-tRNA synthetase and two mutually orthogonal pyrrolysyl-tRNA synthetases, we demonstrate that UAU can be reassigned along with UAG or UAA to encode two distinct ncAAs in the same protein. Furthermore, by engineering the substrate specificity of one of the pyrrolysyl-tRNA synthetases, we developed a triply orthogonal system that enables simultaneous reassignment of UAU, UAG, and UAA to produce proteins containing three distinct ncAAs at precisely defined sites. To showcase the utility of this system, we produced proteins containing two or three ncAAs, with unique bioorthogonal functional groups, and demonstrate that these proteins can be separately modified with multiple fluorescent probes.

Corresponding Authors Alanna Schepartz – Department of Chemistry and Department of Molecular and Cell Biology, University of California, Berkeley, California 94705, United States; schepartz@berkeley.edu; Dieter Söll – Department of Molecular Biophysics and Biochemistry and Department of Chemistry, Yale University, New Haven, Connecticut 06520, United States; dieter.soll@yale.edu.

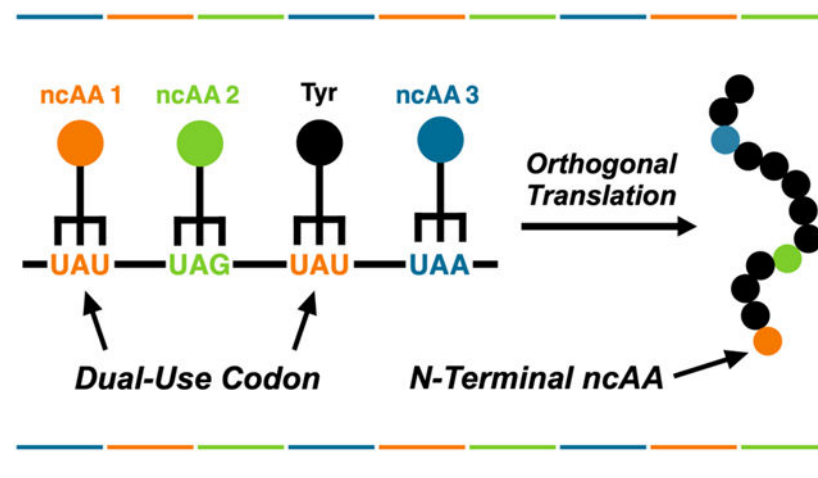
Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acscchembio.1c00120>.

Materials and Methods, Supplementary Figures and Tables, DNA Sequences (PDF)

The authors declare no competing financial interest.

Graphical Abstract



INTRODUCTION

Most lifeforms on Earth utilize a “universal” genetic code in which 64 triplet codons encode the 20 canonical amino acids and three stop signals. In the laboratory, however, artificial genetic code expansion (GCE) has enabled the biosynthesis of proteins containing diverse noncanonical amino acids (ncAAs) at precisely defined sites. This methodology employs orthogonal aminoacyl-tRNA synthetase (o-aaRS) and tRNA (o-tRNA) pairs to co-translationally install ncAAs into proteins, typically in response to a redefined nonsense codon. Laboratory GCE is not limited to the incorporation of a single ncAA. Through the combined action of multiple, mutually orthogonal o-aaRS•o-tRNA pairs, proteins with multiple distinct ncAAs can be produced in living cells.¹⁻⁵ This technology has broad-ranging applications, from introducing reactive moieties into proteins for site-specific bioconjugation at multiple sites^{4,6} to producing homogeneously modified proteins featuring genetically encoded post-translational modifications.^{7,8} An exciting prospect of GCE is the ability one day to produce completely unnatural polypeptides with new-to-nature functions. However, further expansion of the genetic code is limited by the lack of “blank” codons to encode additional unnatural moieties.

Most commonly, GCE relies on stop codon suppression, wherein a nonsense codon (i.e., UAG, UAA, or UGA) is reassigned to encode an ncAA. Recently, mutually orthogonal o-aaRS•o-tRNA pairs were used to simultaneously reassign all three nonsense codons and direct site-specific installation of three distinct ncAAs for the first time *in vivo*.⁴ However, nonsense suppression can only provide three blank codons for GCE. Moreover, reassigning all three nonsense codons requires modified expression systems that assist in accurately terminating translation.⁴ Four-base codons such as UAGA or AGGA represent a promising alternative with the possibility of providing 256 unique blank codons.⁹ Indeed, four-base codons have been used alone, and in combination with nonsense codons, to encode two and three distinct ncAAs in the same protein gene.^{1,5,10} However, translation of four-base codons is very inefficient on wildtype ribosomes.^{3,11,12} It is likely for this reason that four-base codons have not been widely adopted. A third strategy to further expand the genetic

code is to reassign one of the 61 sense codons that normally encode a canonical amino acid.^{13,14} This strategy is particularly challenging due to competition with endogenous aminoacyl-tRNAs for suppression of sense codons.¹⁵ Because of this competition, sense codon reassignment most often affords a heterogeneous product containing a mixture of the canonical and non-canonical amino acids.^{14,16-24}

In a recent study, we reported the development of an orthogonal initiator tRNA ($tRNA^{Ty2}$) that is a substrate for the *Methanocaldococcus jannaschii* tyrosyl-tRNA synthetase (*MjTyRS*)—an o-aaRS that is commonly used for GCE.²⁵ We demonstrated that $tRNA^{Ty2}$ can initiate translation at UAG codons with a variety of ncAAs. The unique ability of $tRNA^{Ty2}$ to initiate translation is afforded by a conserved set of sequence motifs that are exclusive to initiator tRNAs.^{25,26} Here, we asked whether $tRNA^{Ty2}$ could be engineered to initiate translation at a reassigned sense codon. We hypothesized that endogenous elongator tRNAs, which lack structural motifs required for initiation, would be unable to compete with $tRNA^{Ty2}$ for suppression of a reassigned initiator codon, thus abrogating a major hindrance to sense codon reassignment. We demonstrate that an anticodon mutant of $tRNA^{Ty2}$ can efficiently initiate translation with ncAAs in response to the UAU tyrosine codon giving rise to proteins with an ncAA at their N-terminus. Using this mutant initiator tRNA, alongside two mutually orthogonal pyrrolysyl-tRNA synthetase (*PylRS*) and tRNA pairs, we demonstrate that UAU can be reassigned along with UAA and UAG to simultaneously encode two and three distinct ncAAs in the same protein. Finally, we demonstrate that proteins containing two and three reactive ncAAs can be separately modified with multiple fluorescent probes.

RESULTS AND DISCUSSION

Reassigning Sense Codons for Translation Initiation with Noncanonical Amino Acids.

To investigate whether anticodon mutants of $tRNA^{Ty2}$ can initiate translation at reassigned sense codons, we employed a fluorescence-based assay that we previously used for measuring initiation at UAG.²⁵ This assay relies on a superfolder green fluorescent protein (sfGFP)²⁷ reporter in which the initiating AUG codon is replaced with UAG (sfGFP[1UAG]). The reporter is co-expressed with $tRNA^{Ty2}$ and a polyspecific *MjTyRS* variant (pCNFRS²⁸ or AzFRS.2.t1²⁹). Under these conditions, sfGFP production relies on the ability of *MjTyRS* to charge $tRNA^{Ty2}$ with an ncAA (provided in the growth media, Figure 1), and on the ability of aminoacyl- $tRNA^{Ty2}$ to initiate translation at UAG. To modify this assay for measuring initiation at sense codons, we constructed a series of reporter plasmids in which the initiating codon of sfGFP was replaced with one of eight sense codons (Figure 2A). Because *MjTyRS* interacts with the anticodon of its cognate tRNA, we limited our search to codons that required only one base substitution in $tRNA^{Ty2}$.³⁰ First, to assess whether endogenous *E. coli* tRNAs can compete for initiation at these codons, we measured the expression of each sfGFP reporter in the absence of an $tRNA^{Ty2}$ mutant. Without co-expression of a mutant initiator tRNA, very little sfGFP expression was detected (Figure S1), supporting our assumption that endogenous *E. coli* tRNAs cannot initiate translation at these sense codons.

Next, we measured the expression of each sfGFP mutant in *E. coli* cells that were simultaneously expressing $\hat{t}RNA^{Ty2}$ (with complementary anticodon mutations) and pCNFRS. With four of the initiating sense codons (UAU, UUC, AAC, CAC), we observed significant sfGFP expression when the ncAA *para*-methyl-L-phenylalanine (pMeF) was added to the growth media (Figures 1 and 2B). This indicates that the $\hat{t}RNA^{Ty2}$ mutants were successfully charged with pMeF and that the aminoacyl-tRNA can initiate translation at the reassigned sense codon. Of the four codons that showed significant sfGFP expression with pMeF, the tyrosine codon UAU was most efficient. The expression efficiency of sfGFP[1UAU] with pMeF was ~41% that of sfGFP[1UAG] expressed with the same ncAA, and ~14% that of wildtype sfGFP (Figure 2C and S2). We measured sfGFP[1UAU] expression in wildtype *E. coli* DH10B and in an engineered strain in which redundant copies of the methionine initiator tRNA gene were deleted from the genome (strain DH10B *metZWW*). We have shown that initiation with ncAAs is far more efficient in DH10B *metZWW* than in wildtype cells.²⁵ Consistent with our assumption that the mutant initiator tRNA ($\hat{t}RNA^{Ty2}_{AUA}$) is initiating translation at UAU, sfGFP[1UAU] expression was nearly 5-fold greater in DH10B *metZWW* than in the wildtype strain (Figure S3).

To confirm pMeF incorporation in response to UAU, we expressed and purified sfGFP[1UAU] using pCNFRS and $\hat{t}RNA^{Ty2}_{AUA}$. Consistent with in-cell fluorescence measurements, robust expression of the reporter only occurred in the presence of pMeF (Figures 2D and S4A). LC-MS of the purified protein revealed a major peak matching the expected mass of sfGFP with pMeF in place of the initiating methionine (sfGFP-1pMeF, Figure 2E). We also observed a peak corresponding to *N*-formyl-sfGFP-1pMeF (Figure 2E). We have shown that when ncAAs initiate translation in DH10B *metZWW*, a significant fraction of sfGFP retains an N-terminal formyl modification.²⁵ This result further supports the conclusion that pMeF-charged $\hat{t}RNA^{Ty2}_{AUA}$ is initiating translation at UAU. Importantly, no peaks corresponding to tyrosine incorporation at the initiating UAU codon were detected, further demonstrating that the endogenous *E. coli* tyrosine tRNA ($EctRNA^{Tyr}$) is unable to initiate translation at UAU.

In a previous study, we showed that $\hat{t}RNA^{Ty2}$ can both initiate and elongate translation at UAG codons.²⁵ In addition to the initiating UAU, our sfGFP reporter has five elongating UAU codons; however, pMeF was not detected at these positions. This could be due to the low mass difference between pMeF and tyrosine (mass = 1.97 Da) which makes it difficult to distinguish between these residues based on mass alone. Therefore, we expressed sfGFP[1UAU] with *para*-iodo-L-phenylalanine (pIF; Figure 1), which has a larger mass than tyrosine (mass = 109.9 Da). Intact LC-MS of the purified protein revealed major peaks corresponding to sfGFP-1pIF and *N*-formyl-sfGFP-1pIF (Figure S5). No peaks corresponding to tyrosine incorporation at the initiating UAU were detected. Likewise, no peaks corresponding to pIF incorporation at elongating UAU codons were detected. We further analyzed this protein by tandem mass spectrometry (MS/MS) following proteolysis. In this analysis, we identified peptides with masses and fragmentation patterns consistent with pIF incorporation at the N-terminus (Figure S6); we were unable to detect pIF incorporation in response to any elongating UAU codon. Thus, our data suggest that while $\hat{t}RNA^{Ty2}_{AUA}$ can initiate translation with ncAAs at UAU, under these conditions, it is unable to efficiently compete with endogenous $EctRNA^{Tyr}$ for suppression of elongating UAU codons. Taken

together, our results indicate that, in cells expressing $\dot{t}RNA^{Ty2}_{AUA}$, UAU can function as a dual-use codon that selectively encodes ncAAs at the initiating position and predominantly tyrosine at elongating positions. Similar dual-use codons have been reported *in vitro*³¹ and partial dual reassignment of the methionine AUG codon has been reported *in vivo*;¹⁸ however, to our knowledge, this is the first time that reassignment of a dual-use codon has been achieved in living cells.

Although we were unable to detect ncAA incorporation at elongating UAU codons via mass spectrometry, it is likely that ncAAs are still being incorporated at these positions at low levels. To investigate the degree of ncAA incorporation in response to elongating UAU codons, we grew cells expressing AzFRS.2.t1 along with $\dot{t}RNA^{Ty2}_{AUA}$, $\dot{t}RNA^{Ty2}_{CUA}$, or no $\dot{t}RNA$ in the presence of 2 mM *para*-azido-L-phenylalanine (pAzF) and labeled the resulting cell lysates with an alkyne dye. Lysates of cells expressing $\dot{t}RNA^{Ty2}_{AUA}$ showed a greater degree of labeling compared to those from cells expressing $\dot{t}RNA^{Ty2}_{CUA}$, or no $\dot{t}RNA$, suggesting that pAzF is integrated throughout the proteome by $\dot{t}RNA^{Ty2}_{AUA}$ (Figure S7). To evaluate the effect of low-level proteome-wide UAU reassignment on cellular fitness, we monitored the growth of cells expressing AzFRS.2.t1 and $\dot{t}RNA^{Ty2}_{AUA}$, $\dot{t}RNA^{Ty2}_{CUA}$, or no $\dot{t}RNA$ in the presence of 2 mM pMeF. Cells expressing $\dot{t}RNA^{Ty2}_{AUA}$ and $\dot{t}RNA^{Ty2}_{CUA}$ exhibited nearly identical growth characteristics; both showed higher doubling times and lower carrying capacities than cells expressing no $\dot{t}RNA$ (Figure S8). Thus, UAU reassignment with $\dot{t}RNA^{Ty2}_{AUA}$ does not impact cellular fitness to a greater degree than UAG reassignment with $\dot{t}RNA^{Ty2}_{CUA}$. This result is not surprising given the apparently low level of ncAA incorporation in response to elongating UAU codons. To avoid possible complications from elongating UAU codons in subsequent experiments, we synthesized a codon-optimized sfGFP in which all elongating UAU codons were replaced with the synonymous codon UAC. As with our previous reporter, robust expression of the codon-optimized variant occurred only when an ncAA was provided in the growth media (Figures 2G and S4B). When this optimized gene was expressed with *para*-acetyl-L-phenylalanine (pAcF; Figure 1), sfGFP-1pAcF was obtained with a yield of ~260 mg per liter of culture.

Several of the codons that we tested showed an increase in sfGFP expression irrespective of whether an ncAA was provided in the growth media (Figure 2B). One possible explanation for this observation is that mutating the anticodon of $\dot{t}RNA^{Ty2}$ converted the tRNA into a substrate for an endogenous *E. coli* aaRS. For example, we observed an increase in sfGFP expression when the anticodon of $\dot{t}RNA^{Ty2}$ was mutated to GUC (for initiating at GAC). In *E. coli*, GAC encodes aspartate, and the aspartyl-tRNA synthetase uses anticodon bases to recognize its cognate tRNA.³² Similar mis-aminoacylation was observed when the anticodon of pyrrolysine tRNA ($tRNA^{Pyl}$) was mutated to reassign an arginine codon in *Mycoplasma capricolum*.¹⁶ While no tyrosine incorporation at the initiating UAU codon was detected, it is possible that $\dot{t}RNA^{Ty2}_{AUA}$ is still being mis-charged with tyrosine by *E. coli* TyrRS (*EcTyrRS*). Proteins that contain an N-terminal tyrosine are extremely unstable in *E. coli*,³³ thus, mis-charging of $\dot{t}RNA^{Ty2}_{AUA}$ with tyrosine might go undetected if the reporter protein is rapidly degraded *in vivo*. To investigate this possibility, we performed *in vitro* aminoacylation assays using purified *EcTyrRS* and an $\dot{t}RNA^{Ty2}_{AUA}$ transcript. While *EcTyrRS* charged its cognate tRNA with radiolabeled tyrosine, no aminoacylation of

$\dot{\iota}$ RNA^{Ty2}_{AUA} was detected (Figure 2F), indicating that the mutant tRNA is orthogonal to *Ec*TyrRS.

Dual Incorporation of Noncanonical Amino Acids Using Reprogrammed Initiator and Nonsense Codons.

After demonstrating that $\dot{\iota}$ RNA^{Ty2}_{AUA} can initiate translation with ncAAs at UAU codons, we next asked whether UAU could be used together with nonsense codons to encode two distinct ncAAs in the same protein. The two most commonly used codons for GCE are UAG and UAA; however, $\dot{\iota}$ RNA^{Ty2}_{AUA} can base pair with these triplets at all but the wobble position. To encode distinct ncAAs using a combination of UAU and UAG/UAA, $\dot{\iota}$ RNA^{Ty2}_{AUA} must be able to recognize UAU and discriminate against UAG and UAA. Therefore, to evaluate whether $\dot{\iota}$ RNA^{Ty2}_{AUA} recognizes these nonsense codons, we constructed two reporters in which the second position of sfGFP was mutated to UAG or UAA (sfGFP[2UAG] and sfGFP[2UAA], respectively; Figure S9A). We co-expressed these mutant reporters, along with $\dot{\iota}$ RNA^{Ty2}_{AUA} and AzFRS.2.t1, in *E. coli* DH10B *metZWW*, and we monitored sfGFP production in the presence of pMeF. As expected, sfGFP[1UAU] afforded a strong fluorescence signal upon pMeF addition, whereas no significant increase in fluorescence was observed when pMeF was added to the growth media of cells expressing sfGFP[2UAG] or sfGFP-[2UAA] (Figure S9B). This result demonstrates that $\dot{\iota}$ RNA^{Ty2}_{AUA} is orthogonal to the UAG and UAA nonsense codons.

After confirming orthogonality of $\dot{\iota}$ RNA^{Ty2}_{AUA} toward UAG and UAA, we next explored whether these codons could be used, together with a UAU, to encode two different ncAAs. For introducing a second ncAA, we conscripted the PylRS•tRNA^{Pyl} pair, which is orthogonal in bacteria and eukaryotes and has been engineered to recognize numerous ncAA substrates.^{34,35} The most widely used PylRS•tRNA^{Pyl} pairs for GCE are those originating from the *Methanosarcina* species *mazei* (*Mm*PylRS•*Mmt*RNA^{Pyl}) and *barkeri*. Importantly, the PylRS and *Mj*TyrRS systems are mutually orthogonal and have been used to install distinct ncAAs in response to UAA and UAG codons within the same gene.² Here, we asked whether the *Mj*TyrRS and *Mm*PylRS can be used to introduce two unique ncAAs in response to UAU and UAA, respectively. With this objective, we constructed a plasmid system for simultaneous expression of the two o-aaRS•o-tRNA pairs, together with a sfGFP reporter containing a UAU mutation at position 1 and a UAA mutation at position 151 (sfGFP-[1UAU-151UAA]; Figure S10A). In this system, we employed the *Mj*TyrRS variant AzFRS.2.t1, together with $\dot{\iota}$ RNA^{Ty2}_{AUA}, to incorporate pMeF in response to the initiating UAU and wildtype *Mm*PylRS, together with a mutant opal-suppressor *Mmt*RNA^{Pyl} (*Mmt*RNA^{Pyl}_{UUA}),³⁶ to introduce *N*^ε-boc-L-lysine (BocK; Figure 1) in response to the elongating UAA. We measured sfGFP production in *E. coli* DH10B *metZWW* in the presence of both pMeF and BocK. Under these conditions, robust expression occurred only when both ncAAs were provided in the growth media (Figure 3A), suggesting that the UAU and UAA codons were simultaneously suppressed to afford full-length sfGFP with the desired ncAAs.

Next, we evaluated dual ncAA incorporation using a combination of UAU and UAG. To introduce a second ncAA in response to UAG, we employed a recently characterized

Author Manuscript

PyIRS•tRNA^{Py1} pair from the methanogenic archaeon “*Candidatus* Methanomethylphilus alvus” (*Ma*PyIRS•*Mat*RNA^{Py1}).³⁷ Like the *Mm*PyIRS pair, the *Ma*PyIRS pair is orthogonal in bacteria and eukaryotes.^{37,38} Furthermore, we have shown that this pair is mutually orthogonal with the engineered *Mj*TyrRS•tRNA^{Ty2} pair.²⁵ We employed the same plasmid-based system to co-express these two o-aARS•o pairs along with a sfGFP mutant containing a UAU mutation at position 1 and a UAG mutation at position 135 (sfGFP-[1UAU-135UAG]; Figure S10B). AzFRS.2.t1 and tRNA^{Ty2}_{AUA} were used to incorporate pMeF in response to the initiating UAU, while wildtype *Ma*PyIRS and a variant of *Mat*RNA^{Py1} (*Mat*RNA(6)^{Py1}), which was engineered to be orthogonal to *Mm*PyIRS, were used to incorporate BocK in response to UAG.³⁷ Again, robust sfGFP expression occurred only when both pMeF and BocK were added to the growth media, demonstrating dual suppression of UAU and UAG and successful incorporation of both ncAAs (Figure 3B).

Author Manuscript

AzFRS.2.t1 and the wildtype PyIRSSs are polyspecific synthetases that recognize a number of substrates.^{29,34} Among these are ncAAs with bioorthogonal functional groups that enable site-specific protein bioconjugation. For example, in addition to pMeF, AzFRS.2.t1 recognizes the ketone-containing ncAA pAcF, which can undergo oxime ligation with hydroxylamine probes.³⁹ Furthermore, in addition to BocK, wildtype PyIRS recognizes *N*^ε-propargyl-L-lysine (PrK; Figure 1), which can undergo Cu(I)-catalyzed cycloaddition with azides.⁴⁰ We tested whether these o-aARS•o-tRNA pairs could be used to produce proteins containing ketone and alkyne reactive moieties at precisely defined sites by expressing sfGFP[1UAU-135UAG] and sfGFP[1UAU-151UAA] in the presence of pAcF and PrK. Again, robust expression of these reporter proteins occurred only when both ncAAs were provided in the growth media (Figure S10C,D). Next, we overexpressed and purified these proteins to confirmed pAcF and PrK incorporation by mass spectrometry. Purified sfGFP-1pAcF-135PrK and sfGFP-1pAcF-151PrK were obtained in yields of ~18 and 6 mg per liter of culture, respectively. We analyzed these purified proteins by LC-MS; in both cases, we observed mass peaks consistent with sfGFP containing the expected ncAA substitutions (Figure 3C,D). We also confirmed the site-specificity of pAcF and PrK incorporation by MS/MS analysis (Figures S11 and S12). Finally, to demonstrate the utility of this system, we showed that these proteins, which contain ketone and alkyne bioconjugation handles, can be readily labeled with both hydroxylamine- and azide-based fluorescent dyes (Figures 3E and S13).

Co-translational Installation of Three Distinct Non-canonical Amino Acids.

Author Manuscript

The above data demonstrate that *Mj*TyrRS•tRNA^{Ty2}_{AUA} can be used, in conjunction with two different PyIRS•tRNA^{Py1} pairs, to incorporate distinct ncAAs in response to UAU and UAG or UAA. Next, we asked whether these mutually orthogonal o-aARS•o-tRNA pairs can be combined to direct site-specific installation of three unique ncAAs. As a prerequisite, we first sought to identify three o-aARS variants that recognize distinct ncAA substrates. The *Mj*TyrRS variant AzFRS.2.t1 primarily recognizes *para*-substituted phenylalanine derivatives, whereas wildtype *Mm*PyIRS and *Ma*PyIRS both recognize *N*^ε-substituted lysine derivatives and have overlapping substrate specificities.⁴¹ A major contributor to the substrate specificity of PyIRS is the so-called *gatekeeper residue*, a conserved asparagine located in the enzyme’s substrate binding pocket. This residue is involved in several

interactions with the substrate amino acid, including a hydrogen bond with the side chain amide oxygen of pyrrolysine, and other N^ϵ -substituted lysine derivatives (Figure S14A).^{42,43} Mutating this asparagine in *MmPylRS* to alanine (N346A) obliterates recognition of N^ϵ -substituted lysine derivatives.⁴⁴ Introducing a second alanine mutation (C348A) affords an *MmPylRS* variant (*MmPylRS*(N346A/C348A)) that recognizes more than 30 *ortho*-, *para*-, and *meta*-substituted phenylalanine derivatives.⁴⁴⁻⁴⁶ However, this variant poorly recognizes derivatives with small *para* substitutions, such as those recognized by AzFRS.2.t1.⁴⁴ Therefore, we hypothesized that *MmPylRS*(N346A/C348A) and AzFRS.2.t1 could be used together in the same cell to selectively install *meta*- and *para*-substituted phenylalanine derivatives, respectively. In addition, we hypothesized that wildtype *MaPylRS* could serve as a third o-aaRS to install N^ϵ -substituted lysine derivatives.

To confirm the substrate specificity of these o-aaRSs, we measured sfGFP[2UAG] production in cells expressing either AzFRS.2.t1, *MmPylRS*(N346A/C348A), or wildtype *MaPylRS* (as well as their cognate amber-suppressor tRNAs) in the presence of three substrates: pAcF, PrK, and *meta*-iodo-L-phenylalanine (mIF; Figure 1). Consistent with our hypothesis, we found that AzFRS.2.t1 selectively recognizes the *para*-substituted ncAA pAcF, while *MmPylRS*(N346A/C348A) selectively recognizes the *meta*-substituted ncAA mIF (Figure 4A,B). However, surprisingly, we found that mIF is also recognized by wildtype *MaPylRS* (Figure 4C). Indeed, *MaPylRS* enabled robust sfGFP expression in the presence of both mIF and PrK. In contrast, wildtype *MmPylRS* only afforded robust reporter expression in the presence of PrK (Figure 4B).

In light of this observation, we decided instead to use *MaPylRS* to install *meta*-substituted ncAAs and wildtype *MmPylRS* to install N^ϵ -substituted lysine derivatives. However, this required that we engineer an *MaPylRS* variant that accepts mIF and rejects PrK. Toward this end, we constructed a two-site *MaPylRS* library randomizing residues N166 and V168 (corresponding to N346 and C348 in *MmPylRS*, S14B). We subjected this library to positive selection using chloramphenicol acetyltransferase with a UAG codon at position 112 (cat[112UAG]).⁴⁷ Cells expressing the *MaPylRS* N166/V168 library, *MatRNA*(6)^{Py1}, and cat[112UAG] were challenged to grow on media containing 1 mM mIF and 50 μ g/mL chloramphenicol. We also performed a parallel selection using *ortho*-methyl-L-phenylalanine (oMeF; Figure 1). Following selection, surviving clones were screened for ncAA-dependent growth (Figure S15), and five clones from each selection were sequenced to reveal mutations in *MaPylRS* that altered its specificity. In total, four unique mutants with similar amino acid substitutions were identified (Figure 4D). We screened these mutants for selective incorporation of mIF using sfGFP[2UAG]. Gratifyingly, we found that robust expression occurred only in the presence of mIF, indicating that these newly identified *MaPylRS* mutants are selective for the *meta*-substituted ncAA (Figures 4C and S16). In subsequent experiments, we employed the N166S mutant, *MaPylRS*(N166S). In addition to mIF and oMeF, we screened *MaPylRS*(N166S) for the ability to recognize various substituted phenylalanine derivatives. Similar to *MmPylRS*-(N346A/C348A), *MaPylRS*(N166S) recognizes a number of substrates (Figure S17) including those with bioorthogonal handles such as *meta*-azido-L-phenylalanine (mAzF; Figure 1). Like mIF, we found that mAzF is rejected by AzFRS.2.t1 and wildtype *MmPylRS* (Figure 4A-C). To test if this *ortho/para/meta* substrate selectivity is a general feature of *MaPylRS*-(N166S) and AzFRS.2.t1 we

compared sfGFP[2UAG] expression with these enzymes using a panel of phenylalanine derivatives containing the same substituent at different positions of the aromatic ring. In all cases we found that AzFRS.2.t1 selectively recognizes *para*-substituted derivatives, while *Ma*PyIRS(N166S) selectively recognizes derivatives with *ortho* or *meta* substitutions (Figure S18).

Next, we devised a plasmid-based system for simultaneous incorporation of three distinct ncAAs. This three-plasmid system consisted of (1) accessory plasmids encoding AzFRS.2.t1, *Ma*PyIRS(N166S), and *Mat*RNA(6)^{Py1}; (2) a second accessory plasmid encoding *Mm*PyIRS and *Mm*tRNA^{Py1}_{UUA}; and (3) a reporter plasmid encoding tRNA^{Ty2}_{AUA} and a triple mutant sfGFP with a UAU mutation at position 1, a UAG mutation at position 135, and a UAA mutation at position 151 (Figure S19A). We measured sfGFP production *E. coli* DH10B *metZ*WV in the presence of pMeF, BocK, and mIF. With this initial system, we observed significant sfGFP production only in the presence of all three ncAAs; however, the signal was relatively low (Figure S19B). We hypothesized that this was due to poor suppression of the UAA codon by *Mm*tRNA^{Py1}_{UUA} since *Mm*PyIRS is the least active o-aaRS in this system.⁴¹ Therefore, an additional copy of *Mm*tRNA^{Py1}_{UUA} was added to the reporter plasmid resulting in a ~2-fold increase in the fluorescence signal (Figure S19B). Next, we overexpressed and purified sfGFP containing pAcF, PrK, and mIF and analyzed the protein by LC-MS. This analysis revealed a major peak consistent with sfGFP containing three ncAAs at the desired positions; however, a second peak was also present (Figure S20). This second peak was consistent with misincorporation of PrK at the UAG codon. Further analysis by MS/MS confirmed a mixture of mIF and PrK at position 135 (Figures S21 and S22). We surmised that mis-incorporation of PrK likely results from an excess of *Mm*tRNA^{Py1}_{UUA}, since ochre-suppressor tRNAs can suppress both UAA and UAG.⁴⁸ Therefore, we constructed a third reporter plasmid in which the added copy of *Mm*tRNA^{Py1}_{UUA} was replaced with *Mat*RNA(6)^{Py1} (Figure S19A). Using this new reporter plasmid, we again found that sfGFP production occurred only when the growth medium was supplemented with a *para*-substituted phenylalanine, a *meta*-substituted phenylalanine, and an *N*^ε-substituted lysine derivative, i.e., pMeF, mIF, and BocK or pAcF, mIF, and PrK (Figure 5A,B). LC-MS analysis of sfGFP expressed with pAcF, mIF, and PrK revealed a major peak consistent with the incorporation of all three amino acids; a peak corresponding to mis-incorporation of PrK was not detected (Figure 5C). Furthermore, MS/MS analysis confirmed the incorporation of each ncAA at the desired positions (Figure S23). The yield of this purified sfGFP-1pAcF-135mIF-151PrK was ~1.5 mg per liter of culture. Next, we used this optimized expression system to site-specifically install three reactive ncAAs (pAcF, mAzF, and PrK) into sfGFP, affording a protein with three unique bioorthogonal functional groups (ketone, azide, and alkyne; Figure 5D). Finally, we demonstrated that this single protein could be modified with multiple fluorescent probes bearing hydroxylamine, alkyne, or azide reactive handles (Figures 5D and S24).

In summary, we have engineered an orthogonal initiator tRNA that enables efficient reassignment of the UAU codon to initiate translation with ncAAs. We demonstrated that this system enables double and triple incorporation of distinct ncAAs encoded by UAU, UAG, and UAA codons. Several of the ncAAs that we tested for simultaneous incorporation (e.g., pAcF, mIF, mAzF, and PrK) contain unique bioorthogonal reaction handles that enable

double and triple labeling of proteins produced from this system. Due to the polyspecificity of the o-aaRSs used in this study, this system should be compatible with several other reactive ncAAs. For example, the wildtype *MmPylRS* used in this study also recognizes cyclopropene-containing ncAAs which can undergo strain-promoted inverse electron-demand Diels–Alder cycloaddition with tetrazine probes.^{4,49} Furthermore, in addition to the demonstrated reactivity of pAcF, mAzF, and PrK, mIF can also be used as a reaction handle for protein conjugation with boronic acids via palladium-catalyzed Suzuki–Miyaura cross-coupling.⁵⁰

Safety Statement.

No unexpected, new, or significant hazards or risks were encountered during the course of this work.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We thank K. Hoffman and S. Santiago for help with mass spectrometry data collection and analysis and E. Vatansever, J. Fischer, C. Chung, and N. Krahn for providing critical feedback on the manuscript. J. M. Tharp was supported by the Center for Genetically Encoded Materials, an NSF Center for Chemical Innovation (NSF CHE-2021739 to A. Schepartz). O. Vargas-Rodriguez was supported by the National Institute of General Medical Sciences (R35GM122560 to D. Söll). The genetic experiments were supported by the Division of Chemical Sciences, Geosciences, and Biosciences, Office of Basic Energy Sciences of the Department of Energy (DE-FG0298ER20311 to D. Söll).

REFERENCES

- (1). Anderson JC, Wu N, Santoro SW, Lakshman V, King DS, and Schultz PG (2004) An expanded genetic code with a functional quadruplet codon. *Proc. Natl. Acad. Sci. U. S. A* 101, 7566–7571. [PubMed: 15138302]
- (2). Wan W, Huang Y, Wang Z, Russell WK, Pai PJ, Russell DH, and Liu WR (2010) A facile system for genetic incorporation of two different noncanonical amino acids into one protein in *Escherichia coli*. *Angew. Chem., Int. Ed* 49, 3211–3214.
- (3). Neumann H, Wang K, Davis L, Garcia-Alai M, and Chin JW (2010) Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature* 464, 441–444. [PubMed: 20154731]
- (4). Italia JS, Addy PS, Erickson SB, Peeler JC, Weerapana E, and Chatterjee A (2019) Mutually orthogonal nonsense-suppression systems and conjugation chemistries for precise protein labeling at up to three distinct sites. *J. Am. Chem. Soc* 141, 6204–6212. [PubMed: 30909694]
- (5). Dunkelmann DL, Willis JCW, Beattie AT, and Chin JW (2020) Engineered triply orthogonal pyrrolysyl-tRNA synthetase/tRNA pairs enable the genetic encoding of three distinct non-canonical amino acids. *Nat. Chem* 12, 535–544. [PubMed: 32472101]
- (6). Wu B, Wang Z, Huang Y, and Liu WR (2012) Catalyst-free and site-specific one-pot dual-labeling of a protein directed by two genetically incorporated noncanonical amino acids. *ChemBioChem* 13, 1405–1408. [PubMed: 22628069]
- (7). Venkat S, Sturges J, Stahman A, Gregory C, Gan Q, and Fan C (2018) Genetically incorporating two distinct post-translational modifications into one protein simultaneously. *ACS Synth. Biol* 7, 689–695. [PubMed: 29301074]

- (8). Zheng Y, Gilgenast MJ, Hauc S, and Chatterjee A (2018) Capturing post-translational modification-triggered protein-protein interactions using dual noncanonical amino acid mutagenesis. *ACS Chem. Biol* 13, 1137–1141. [PubMed: 29544052]
- (9). Wang K, Schmied WH, and Chin JW (2012) Reprogramming the genetic code: from triplet to quadruplet codes. *Angew. Chem., Int. Ed* 51, 2288–2297.
- (10). Hankore ED, Zhang LY, Chen Y, Liu K, Niu W, and Guo JT (2019) Genetic incorporation of noncanonical amino acids using two mutually orthogonal quadruplet codons. *ACS Synth. Biol* 8, 1168–1174. [PubMed: 30995842]
- (11). Wang N, Shang X, Cerny R, Niu W, and Guo J (2016) Systematic evolution and study of UAGN decoding tRNAs in a genomically recoded bacteria. *Sci. Rep* 6, 21898. [PubMed: 26906548]
- (12). Chatterjee A, Lajoie MJ, Xiao H, Church GM, and Schultz PG (2014) A bacterial strain with a unique quadruplet codon specifying non-native amino acids. *ChemBioChem* 15, 1782–1786. [PubMed: 24867343]
- (13). Link AJ, Mock ML, and Tirrell DA (2003) Non-canonical amino acids in protein engineering. *Curr. Opin. Biotechnol* 14, 603–609. [PubMed: 14662389]
- (14). Bohlke N, and Budisa N (2014) Sense codon emancipation for proteome-wide incorporation of noncanonical amino acids: rare isoleucine codon AUA as a target for genetic code expansion. *FEMS Microbiol. Lett* 351, 133–144. [PubMed: 24433543]
- (15). Krishnakumar R, and Ling J (2014) Experimental challenges of sense codon reassignment: an innovative approach to genetic code expansion. *FEBS Lett.* 588, 383–388. [PubMed: 24333334]
- (16). Krishnakumar R, Prat L, Aerni HR, Ling JQ, Merryman C, Glass JI, Rinehart J, and Söll D (2013) Transfer RNA misidentification scrambles sense codon recoding. *ChemBioChem* 14, 1967–1972. [PubMed: 24000185]
- (17). Zeng Y, Wang W, and Liu WR (2014) Towards reassigning the rare AGG codon in *Escherichia coli*. *ChemBioChem* 15, 1750–1754. [PubMed: 25044341]
- (18). De Simone A, Acevedo-Rocha CG, Hoesl MG, and Budisa N (2016) Towards reassignment of the methionine codon AUG to two different noncanonical amino acids in bacterial translation. *Croat. Chem. Acta* 89, 243–253.
- (19). Mukai T, Yamaguchi A, Ohtake K, Takahashi M, Hayashi A, Iraha F, Kira S, Yanagisawa T, Yokoyama S, Hoshi H, Kobayashi T, and Sakamoto K (2015) Reassignment of a rare sense codon to a non-canonical amino acid in *Escherichia coli*. *Nucleic Acids Res.* 43, 8111–8122. [PubMed: 26240376]
- (20). Biddle W, Schmitt MA, and Fisk JD (2015) Evaluating sense codon reassignment with a simple fluorescence screen. *Biochemistry* 54, 7355–7364. [PubMed: 26536053]
- (21). Lee BS, Shin S, Jeon JY, Jang KS, Lee BY, Choi S, and Yoo TH (2015) Incorporation of unnatural amino acids in response to the AGG codon. *ACS Chem. Biol* 10, 1648–1653. [PubMed: 25946114]
- (22). Wang Y, and Tsao ML (2016) Reassigning sense codon AGA to encode noncanonical amino acids in *Escherichia coli*. *ChemBioChem* 17, 2234–2239. [PubMed: 27647777]
- (23). Kwon I, Kirshenbaum K, and Tirrell DA (2003) Breaking the degeneracy of the genetic code. *J. Am. Chem. Soc* 125, 7512–7513. [PubMed: 12812480]
- (24). Ho JM, Reynolds NM, Rivera K, Connolly M, Guo LT, Ling J, Pappin DJ, Church GM, and Söll D (2016) Efficient reassignment of a frequent serine codon in wild-type *Escherichia coli*. *ACS Synth. Biol* 5, 163–171. [PubMed: 26544153]
- (25). Tharp JM, Ad O, Amikura K, Ward FR, Garcia EM, Cate JHD, Schepartz A, and Söll D (2020) Initiation of protein synthesis with non-canonical amino acids in vivo. *Angew. Chem., Int. Ed* 59, 3122–3126.
- (26). Tharp JM, Krahn N, Varshney U, and Söll D (2020) Hijacking translation initiation for synthetic biology. *ChemBioChem* 21, 1387–1396. [PubMed: 32023356]
- (27). Pédelacq JD, Cabantous S, Tran T, Terwilliger TC, and Waldo GS (2006) Engineering and characterization of a superfolder green fluorescent protein. *Nat. Biotechnol* 24, 79–88. [PubMed: 16369541]

- (28). Young DD, Young TS, Jahnz M, Ahmad I, Spraggon G, and Schultz PG (2011) An evolved aminoacyl-tRNA synthetase with atypical polysubstrate specificity. *Biochemistry* 50, 1894–1900. [PubMed: 21280675]
- (29). Amiram M, Haimovich AD, Fan C, Wang YS, Aerni HR, Ntai I, Moonan DW, Ma NJ, Rovner AJ, Hong SH, Kelleher NL, Goodman AL, Jewett MC, Söll D, Rinehart J, and Isaacs FJ (2015) Evolution of translation machinery in recoded bacteria enables multi-site incorporation of nonstandard amino acids. *Nat. Biotechnol* 33, 1272–1279. [PubMed: 26571098]
- (30). 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. *Nat. Struct. Mol. Biol* 10, 425–432.
- (31). Goto Y, Iseki M, Hitomi A, Murakami H, and Suga H (2013) Nonstandard peptide expression under the genetic code consisting of reprogrammed dual sense codons. *ACS Chem. Biol* 8, 2630–2634. [PubMed: 24099623]
- (32). Giegé R, Sissler M, and Florentz C (1998) Universal rules and idiosyncratic features in tRNA identity. *Nucleic Acids Res.* 26, 5017–5035. [PubMed: 9801296]
- (33). Tobias JW, Shrader TE, Rocap G, and Varshavsky A (1991) The N-end rule in bacteria. *Science* 254, 1374–1377. [PubMed: 1962196]
- (34). Wan W, Tharp JM, and Liu WR (2014) Pyrrolysyl-tRNA synthetase: an ordinary enzyme but an outstanding genetic code expansion tool. *Biochim. Biophys. Acta, Proteins Proteomics* 1844, 1059–1070.
- (35). Dumas A, Lercher L, Spicer CD, and Davis BG (2015) Designing logical codon reassignment—Expanding the chemistry in biology. *Chem. Sci* 6, 50–69. [PubMed: 28553457]
- (36). Chatterjee A, Sun SB, Furman JL, Xiao H, and Schultz PG (2013) A versatile platform for single- and multiple-unnatural amino acid mutagenesis in *Escherichia coli*. *Biochemistry* 52, 1828–1837. [PubMed: 23379331]
- (37). Willis JCW, and Chin JW (2018) Mutually orthogonal pyrrolysyl-tRNA synthetase/tRNA pairs. *Nat. Chem* 10, 831–837. [PubMed: 29807989]
- (38). Meineke B, Heimgärtner J, Lafranchi L, and Elsässer SJ (2018) *Methanomethylphilus alvus* Mx1201 provides basis for mutual orthogonal pyrrolysyl tRNA/aminoacyl-tRNA synthetase pairs in mammalian cells. *ACS Chem. Biol* 13, 3087–3096. [PubMed: 30260624]
- (39). 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. U. S. A* 100, 56–61. [PubMed: 12518054]
- (40). Nguyen DP, Lusic H, Neumann H, Kapadnis PB, Deiters A, and Chin JW (2009) Genetic encoding and labeling of aliphatic azides and alkynes in recombinant proteins via a pyrrolysyl-tRNA synthetase/tRNA(CUA) pair and click chemistry. *J. Am. Chem. Soc* 131, 8720–8721. [PubMed: 19514718]
- (41). Yamaguchi A, Iraha F, Ohtake K, and Sakamoto K (2018) Pyrrolysyl-tRNA synthetase with a unique architecture enhances the availability of lysine derivatives in synthetic genetic codes. *Molecules* 23, 2460.
- (42). Kavran JM, Gundllapalli S, O'Donoghue P, Englert M, Söll D, and Steitz TA (2007) Structure of pyrrolysyl-tRNA synthetase, an archaeal enzyme for genetic code innovation. *Proc. Natl. Acad. Sci. U. S. A* 104, 11268–11273. [PubMed: 17592110]
- (43). Flügel V, Vrabel M, and Schneider S (2014) Structural basis for the site-specific incorporation of lysine derivatives into proteins. *PLoS One* 9, e96198. [PubMed: 24760130]
- (44). Wang YS, Fang X, Wallace AL, Wu B, and Liu WR (2012) A rationally designed pyrrolysyl-tRNA synthetase mutant with a broad substrate spectrum. *J. Am. Chem. Soc* 134, 2950–2953. [PubMed: 22289053]
- (45). Wang YS, Fang X, Chen HY, Wu B, Wang ZU, Hilty C, and Liu WR (2013) Genetic incorporation of twelve *meta*-substituted phenylalanine derivatives using a single pyrrolysyl-tRNA synthetase mutant. *ACS Chem. Biol* 8, 405–415. [PubMed: 23138887]
- (46). Tharp JM, Wang YS, Lee YJ, Yang Y, and Liu WR (2014) Genetic incorporation of seven *ortho*-substituted phenylalanine derivatives. *ACS Chem. Biol* 9, 884–890. [PubMed: 24451054]
- (47). Wang L, Brock A, Herberich B, and Schultz PG (2001) Expanding the genetic code of *Escherichia coli*. *Science* 292, 498–500. [PubMed: 11313494]

- (48). Eggertsson G, and Söll D (1988) Transfer ribonucleic acid-mediated suppression of termination codons in *Escherichia coli*. *Microbiol. Rev* 52, 354–374. [PubMed: 3054467]
- (49). Patterson DM, Nazarova LA, Xie B, Kamber DN, and Prescher JA (2012) Functionalized cyclopropenes as bioorthogonal chemical reporters. *J. Am. Chem. Soc* 134, 18638–18643. [PubMed: 23072583]
- (50). Chalker JM, Wood CS, and Davis BG (2009) A convenient catalyst for aqueous and protein Suzuki-Miyaura cross-coupling. *J. Am. Chem. Soc* 131, 16346–16347. [PubMed: 19852502]

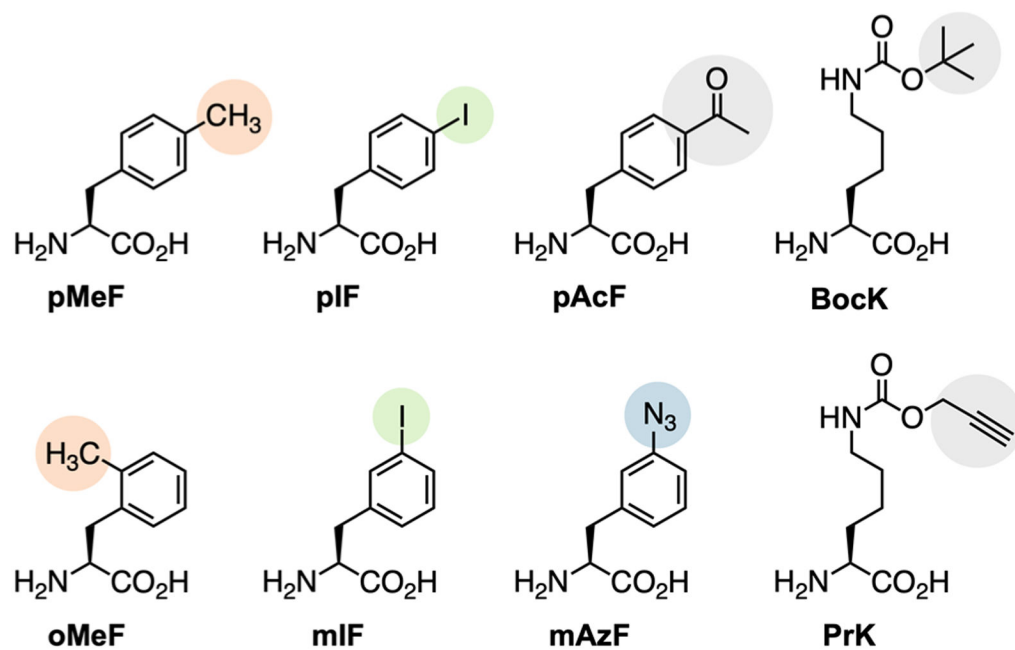
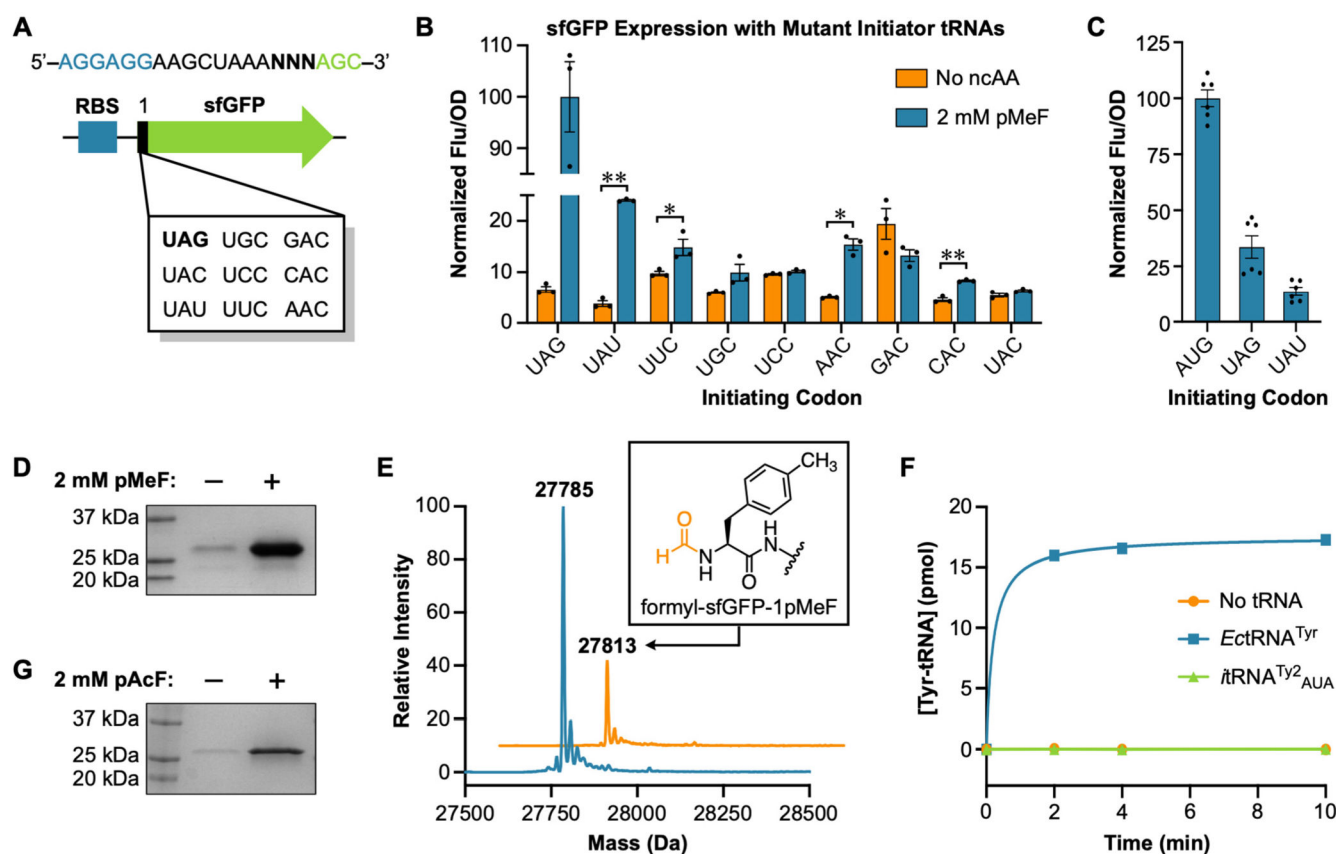


Figure 1.
Structures of ncAAs used in this study.

**Figure 2.**

Reassigning sense codons for initiation with ncAAs. (A) A series of sfGFP reporters was constructed in which the initiating methionine codon was replaced with UAG or one of eight sense codons. The ribosome binding site and initiating codon are shown in blue and bold text, respectively. (B) Co-expression of a mutant initiator tRNA enables translation initiation with ncAAs at select codons. Data are displayed as the mean \pm SEM of three biological replicates (* $p < 0.05$, ** $p < 0.005$, paired t test). (C) Relative expression of sfGFP initiating at AUG, UAG, and UAU. Data are displayed as the mean \pm SEM of six biological replicates. (D) Expression of sfGFP[1UAU] is dependent on the addition of pMeF to the growth media. (E) LC-MS of panel D. Peaks corresponding to sfGFP-1pMeF (theoretical mass = 27786 Da) and N-formyl-sfGFP-1pMeF (theoretical mass = 27814 Da) were detected. (F) tRNA^{Ty2}_{AUA} is not aminoacylated by the *E. coli* TyrRS. Time points represent the mean of three independent experiments. Error bars are contained within each data point. (G) ncAA-dependent expression of codon-optimized sfGFP in which all elongating UAU codons were replaced with UAC.

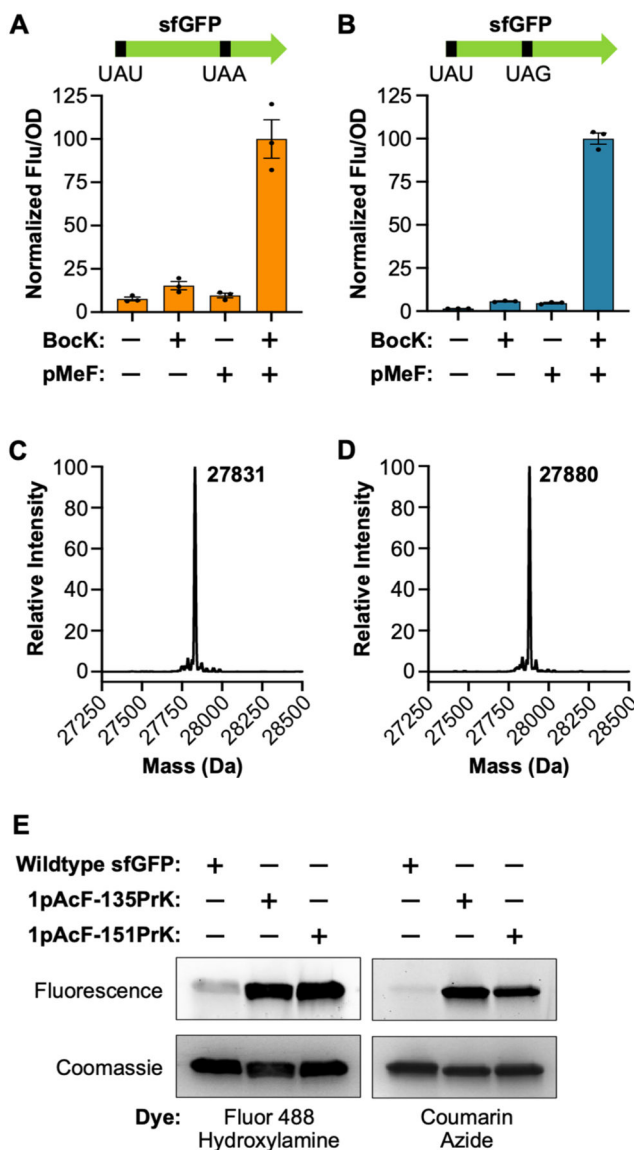


Figure 3. Co-translational installation of two distinct ncAAs encoded by UAU/UAA and UAU/UAG. (A,B) Expression of sfGFP[1UAU-151UAA] (A) and sfGFP[1UAU-135UAG] (B) in the presence of pMeF and BocK. Data are displayed as the mean \pm SEM of three biological replicates. (C) LC-MS of sfGFP-1pAcF-135PrK (theoretical mass = 27833 Da). (D) LC-MS of sfGFP-1pAcF-151PrK (theoretical mass = 27882 Da). (E) Labeling of sfGFP-1pAcF-135PrK and sfGFP-1pAcF-151PrK with Fluor 488-hydroxylamine and coumarin azide. Proteins were resolved by SDS-PAGE and visualized by in-gel fluorescence and Coomassie staining.

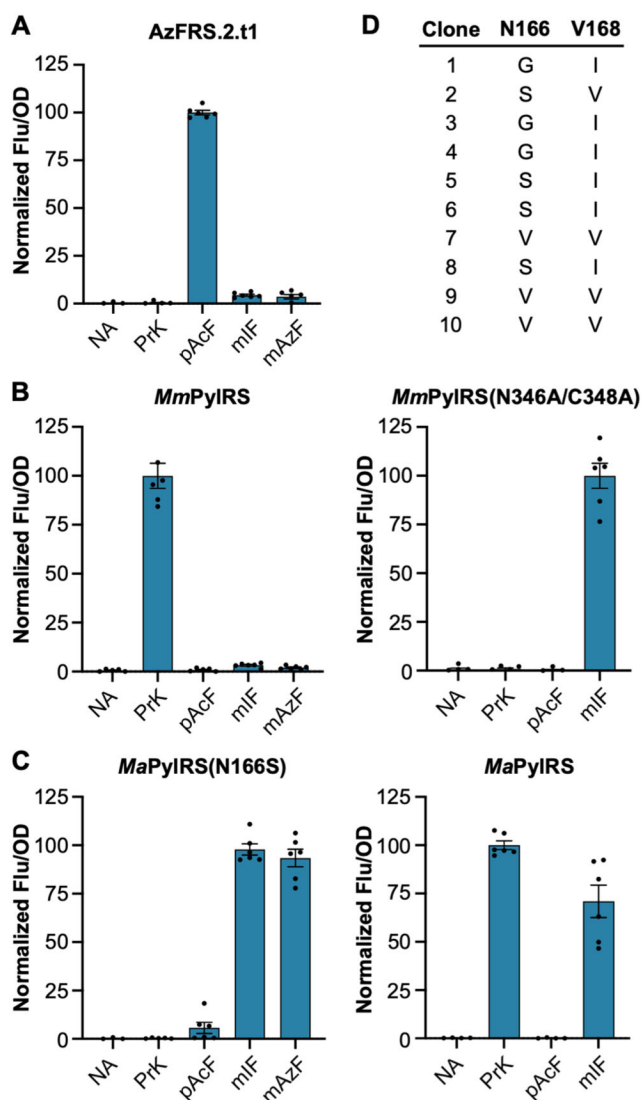


Figure 4. Substrate specificities of three mutually orthogonal aminoacyl-tRNA synthetases. (A–C) Expression of sfGFP[2UAG] in the presence of PrK, pAcF, mIF, or mAzF facilitated by (A) the *Mj*TyrRS variant AzFRS.2.t1, (B) wildtype *Mm*PyIRS and variant *Mm*PyIRS(N346A/C348A), and (C) wildtype *Ma*PyIRS and variant *Ma*PyIRS(N166S). Data are displayed as the mean \pm SEM of six biological replicates. NA = None Added. (D) *Ma*PyIRS mutants identified from an N166/V168 library selected on media containing mIF (clones 1–5) or oMeF (clones 6–10).

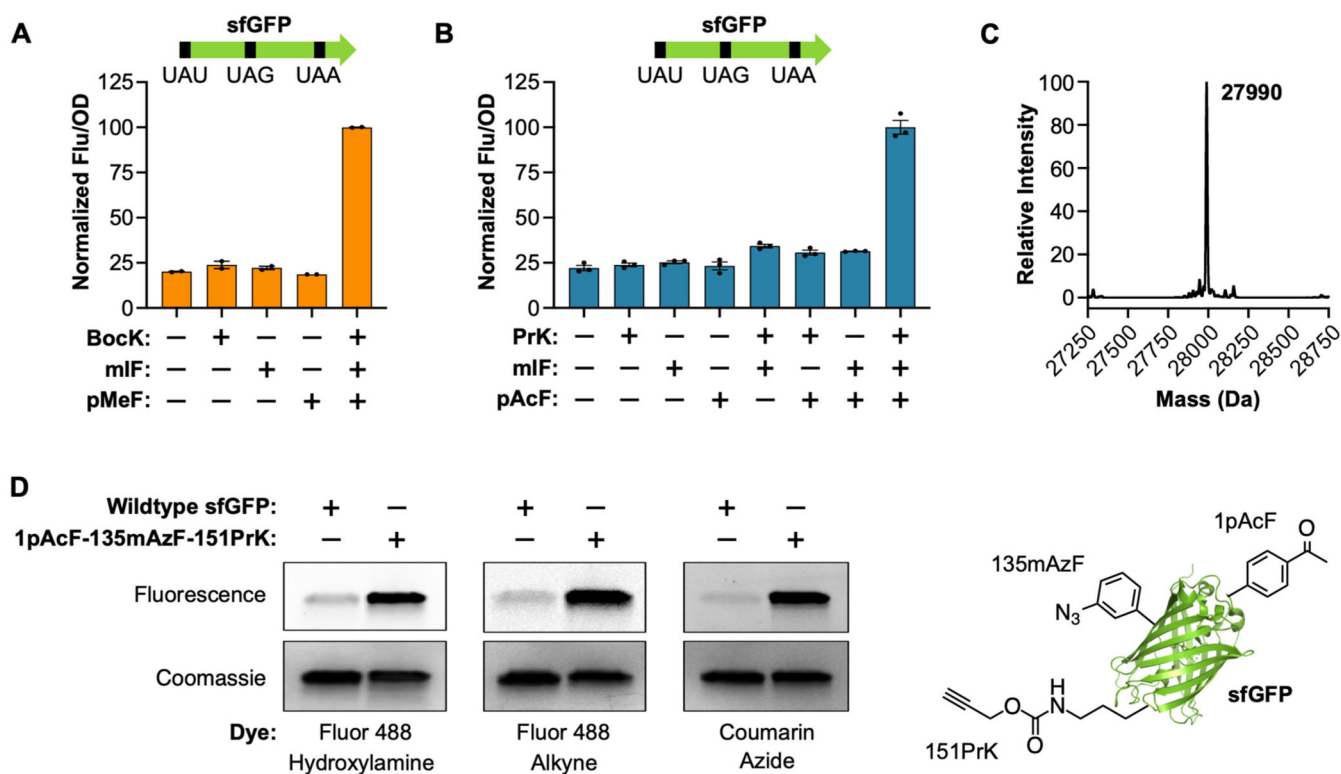


Figure 5. Co-translational installation of three distinct noncanonical amino acids encoded by UAU, UAG, and UAA. (A) Expression of sfGFP[1UAU-135UAG-151UAA] in the presence of Bock, mIF, and pMeF. Data are displayed as the mean \pm SEM of two biological replicates. (B) Expression of sfGFP[1UAU-135UAG-151UAA] in the presence of PrK, mIF, and pAcF. Data are displayed as the mean \pm SEM of three biological replicates. (C) LC-MS of sfGFP-1pAcF-135mIF-151PrK (theoretical mass = 27992 Da). (D) Labeling of sfGFP-1pAcF-135 mAzF-151PrK with Fluor 488-hydroxylamine, Fluor 488-alkyne, and coumarin azide. Proteins were resolved by SDS-PAGE and visualized by in-gel fluorescence and Coomassie staining. A cartoon representation of sfGFP-1pAcF-135 mAzF-151PrK is also shown (PDB: 2B3P).