



An Evolved *Methanomethylophilus alvus* Pyrrolysyl-tRNA Synthetase/tRNA Pair Is Highly Active and Orthogonal in Mammalian Cells

Václav Beránek, Julian C. W. Willis, and Jason W. Chin*¹

Medical Research Council Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge CB2 0QH, England, U.K.

Supporting Information

ABSTRACT: We recently characterized a new class of pyrrolysyl-tRNA synthetase (PylRS)/^{Pyl}tRNA pairs from *Methanomassiliicocales* that are active and orthogonal in *Escherichia coli*. The aminoacyl-tRNA synthetases (aaRSs) of these pairs lack the N-terminal domain that is essential for tRNA recognition and *in vivo* activity in the *Methanosarcina mazei* (*Mm*) PylRS but share a homologous active site with *Mm*PylRS; this facilitates the transplantation of mutations discovered with existing PylRS systems into the new PylRS systems to reprogram their substrate specificity for the incorporation of noncanonical amino acids (ncAAs). Several of the new PylRS/^{Pyl}tRNA pairs, or their evolved variants [including *Methanomethylophilus alvus* (*Ma*) PylRS/*Ma*^{Pyl}tRNA-(6)_{CUA}], are mutually orthogonal to the *Mm*PylRS/*Mm*^{Pyl}tRNA pair, and the active sites of the *Mm* pair and *Ma* pair can be diverged to enable the incorporation of distinct ncAAs in response to distinct codons via orthogonal translation in *E. coli*. Here we demonstrate that *Ma*PylRS/*Ma*^{Pyl}tRNA(6)_{CUA} is orthogonal to the aaRSs and tRNAs in mammalian cells and directs efficient incorporation of ncAAs into proteins. Moreover, we confirm that the *Ma*PylRS/*Ma*^{Pyl}tRNA(6) and *Mm*PylRS/*Mm*^{Pyl}tRNA pairs are mutually orthogonal in mammalian cells and demonstrates that these pairs can be used to encode distinct ncAAs into a protein in mammalian cells. Thus, the *Ma*PylRS/*Ma*^{Pyl}tRNA(6)_{CUA} pair provides an additional pair that is orthogonal in both *E. coli* and mammalian systems and is mutually orthogonal to the most widely used system for genetic code expansion. Our results provide a foundation for expanding the scope of genetic code expansion and may also facilitate strategies for proteome-wide ncAA tagging with mutually orthogonal systems.

Genetically encoding the site-specific co-translational incorporation of noncanonical amino acids into proteins in eukaryotic cells and animals has provided numerous strategies for imaging and controlling the functions of proteins in their native environment.¹ Extensions of these approaches have enabled the tagging and labeling of cell-specific proteomes via stochastic orthogonal recoding of translation (SORT).^{2–5}

The incorporation of ncAAs into proteins relies on the development of orthogonal aminoacyl-tRNA synthetase

(aaRS)/tRNA pairs: the orthogonal aaRS selectively recognizes its cognate orthogonal tRNA over endogenous tRNAs, and the orthogonal tRNA is a substrate for the orthogonal aaRS but a poor substrate for endogenous synthetases. Because the sets of endogenous synthetases and tRNAs differ between organisms, aaRS pairs that are orthogonal in one system are commonly not orthogonal in another.¹ For example, the *Methanocaldococcus janaschii* (*Mj*) TyrRS pair that has been extensively used for genetic code expansion in *Escherichia coli* (*Ec*) cannot be used for genetic code expansion in eukaryotic cells because it is not orthogonal with respect to endogenous eukaryotic aaRS/tRNA pairs.

The *Ec*TyrRS/^{Tyr}tRNA^{6,7} pair and *Ec*LeuRS/^{Leu}tRNA pair⁸ are orthogonal in eukaryotic cells, and variants of these pairs have been discovered, primarily by directed evolution in yeast or subsequent screening, that enable the incorporation of a range of ncAAs in eukaryotic systems. The PylRS/^{Pyl}tRNA pair from *Mm* is commonly considered an ideal pair for genetic code expansion because it is orthogonal in both *E. coli* and eukaryotic cells and animals.¹ This has facilitated the discovery and characterization of *Mm*PylRS variants that incorporate ncAAs in *E. coli* and the transfer of these variants to eukaryotic systems, thereby facilitating genetic code expansion in eukaryotic cells and animals. Recent work has demonstrated that an evolved SepRS/^{v1.0/Sep}tRNA^{v1.0} pair enables the efficient incorporation of phosphoserine and its nonhydrolyzable analogue,⁹ can be further evolved to incorporate phosphothreonine in *E. coli*,¹⁰ and is also orthogonal in mammalian cells.¹¹ The ability to incorporate ncAAs into proteins in mammalian cells has been further expanded by strategies that replace the genomically encoded *Ec*TrpRS/^{Trp}tRNA pair in *E. coli* with the *Saccharomyces cerevisiae* (*Sc*) TrpRS/^{Trp}tRNA pair.^{12–14} Because the *Sc*TrpRS/^{Trp}tRNA is orthogonal in *E. coli*, suppressor derivatives of the *Ec*TrpRS/^{Trp}tRNA pair can be introduced into the resulting *E. coli* strains and evolved for ncAA incorporation. The resulting *Ec*TrpRS/^{Trp}tRNA pairs can then be used for genetic code expansion in mammalian cells, where they are orthogonal. Recent work has extended this strategy to the *Ec*TyrRS/^{Tyr}tRNA pair.¹⁵ In several cases, aaRS/tRNA pairs that are active and orthogonal with respect to the endogenous aaRSs and tRNAs in mammalian cells have

Special Issue: Regulating the Central Dogma

Received: July 31, 2018

Revised: September 7, 2018

Published: September 27, 2018

also been shown to be orthogonal with respect to other orthogonal pairs, creating “mutually orthogonal” pairs in mammalian cells.^{16–18}

We recently discovered that a new class of PylRS/^{Pyl}tRNA pairs from *Methanomassiliicocales* are active and orthogonal in *E. coli*.^{19,20} These pairs, unlike the *Mm*PylRS/*Mm*^{Pyl}tRNA pair, lack the N-terminal domain of PylRS, which was previously thought to be essential for tRNA recognition and aminoacylation.^{21,22} We showed that certain orthogonal pairs from this class are naturally mutually orthogonal to the *Mm*PylRS;¹⁹ this surprising result demonstrated that there is sufficient divergence between archeal PylRS/^{Pyl}tRNA pairs to generate mutually orthogonal pairs for the same amino acid within a domain of life. We developed a number of exceptionally active and orthogonal pairs, including the *Ma*PylRS/*Ma*^{Pyl}tRNA(6) pair, that are mutually orthogonal to the *Mm*PylRS/*Mm*^{Pyl}tRNA pair by virtue of mutations introduced into the body of *Ma*^{Pyl}tRNA.¹⁹ Moreover, we showed that, because the active sites of *Mm*PylRS and *Ma*PylRS share common substrate recognition determinants, we could transplant mutations that direct the selective incorporation of specific ncAAs from the *Mm*PylRS active site to the *Ma*PylRS active site to reprogram its substrate specificity.¹⁹ Finally, we showed that by diverging the active sites of *Ma*PylRS and *Mm*PylRS to selectively recognize distinct substrates and altering the anticodons of *Ma*^{Pyl}tRNA(6) and *Mm*^{Pyl}tRNA to decode distinct codons in orthogonal translation, we could use these pairs in the same cell to direct the incorporation of two distinct ncAAs into a single polypeptide.¹⁹ Here we show that the *Ma*PylRS/*Ma*^{Pyl}tRNA(6) pair is highly active and orthogonal in mammalian cells, where it is also mutually orthogonal to the *Mm*PylRS/*Mm*^{Pyl}tRNA pair (Figure 1), and that derivatives of the two pairs can be used together to incorporate distinct amino acids into a protein in mammalian cells.

We first demonstrated that the *Ma*PylRS/*Ma*^{Pyl}tRNA(6)_{CUA} pair is active in mammalian cells and that both *Ma*PylRS and *Ma*^{Pyl}tRNA(6)_{CUA} are orthogonal with respect to the endogenous tRNAs and the aaRSs in human cells. To this aim, we cloned coding sequences of *Ma*PylRS and *Ma*^{Pyl}tRNA(6)_{CUA} into a vector for mammalian expression;²³ we cloned *Ma*PylRS under the EF1a promoter and four copies of the *Ma*^{Pyl}tRNA(6)_{CUA} under the human U6 promoter into the same vector. We transiently co-transfected HEK293 cells with the *Ma*PylRS/*Ma*^{Pyl}tRNA(6)_{CUA} vector and an mCherry-TAG-GFP reporter¹¹ and cultured the resulting cells in the presence or absence of *N*ε-[(*tert*-butoxy)carbonyl]-L-lysine (BocK), a known substrate for this pair.¹⁹ We measured the ratio of GFP to mCherry fluorescence by flow cytometry (Figure 2a and Figure S1) and fluorescence microscopy (Figure S2) and performed control experiments with the well-characterized *Mm*PylRS/*Mm*^{Pyl}tRNA_{CUA} pair.

We observed minimal readthrough of the amber codon by the *Ma*PylRS/*Ma*^{Pyl}tRNA(6)_{CUA} pair in the absence of BocK (Figure 2a). This demonstrates that *Ma*^{Pyl}tRNA(6)_{CUA} is orthogonal with respect to the aminoacyl-tRNA synthetases that are endogenous in human cells. Upon addition of 0.5 and 1 mM BocK, we observed substantial readthrough of the TAG codon by the *Ma*PylRS/*Ma*^{Pyl}tRNA(6)_{CUA} pair. The level of amber codon readthrough mediated by the *Ma*PylRS/*Ma*^{Pyl}tRNA(6)_{CUA} pair is comparable to that mediated by the highly active *Mm*PylRS/*Mm*^{Pyl}tRNA_{CUA} pair (Figure 2a); this demonstrates that the *Ma*PylRS/*Ma*^{Pyl}tRNA(6)_{CUA} pair is highly active in mammalian cells.

E. coli

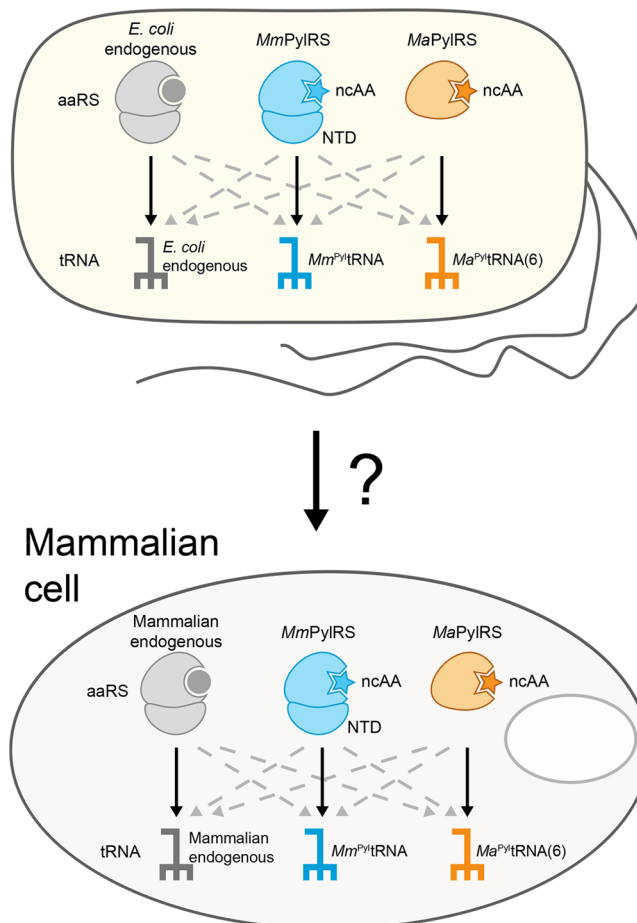


Figure 1. The evolved *Ma*PylRS/*Ma*^{Pyl}tRNA(6) is orthogonal with respect to the endogenous *E. coli* aaRS/tRNA pairs as well as mutually orthogonal to the *Mm*PylRS/*Mm*^{Pyl}tRNA(6) pair. Here we ask if this pair is also orthogonal and mutually orthogonal in mammalian cells. NTD (N-terminal domain).

To demonstrate that *Ma*PylRS is functionally orthogonal in human cells, we co-transfected plasmids encoding the *Ma*PylRS/*Ma*^{Pyl}tRNA(6)_{CUA} pair and *GFP(150TAG)His₆* and cultured the cells in the presence of 1 mM BocK. ESI-MS of the resulting GFP gives the expected mass (Figure 2b) and is indistinguishable from a control in which we used the *Mm*PylRS/*Mm*^{Pyl}tRNA_{CUA} pair to incorporate BocK into *GFP(150TAG)His₆* (Figure 2c). Taken together, our experiments reveal that the *Ma*PylRS/*Ma*^{Pyl}tRNA(6)_{CUA} pair is a highly active and orthogonal pair in mammalian cells.

Next, we aimed to demonstrate that the *Mm*PylRS/*Mm*^{Pyl}tRNA_{CUA} pair and *Ma*PylRS/*Ma*^{Pyl}tRNA(6)_{CUA} pair are mutually orthogonal in their aminoacylation specificity when expressed in mammalian cells. We swapped the 4× U6-tRNA cassette between the *Mm* and *Ma* expression vectors producing a plasmid containing the *Mm*PylRS/*Ma*^{Pyl}tRNA(6)_{CUA} pair and a plasmid containing the *Ma*PylRS/*Mm*^{Pyl}tRNA_{CUA} pair. We co-transfected the mCherry-TAG-GFP reporter with each aaRS/tRNA combinations and cultured the cells in the presence and absence of BocK. We compared the readthrough of the amber stop codon by these noncognate pairs to that mediated by the *Mm*- and *Ma*-derived cognate pairs (Figure 3). Our data show that the noncognate *Mm*PylRS/*Ma*^{Pyl}tRNA(6)_{CUA} and *Ma*PylRS/*Mm*^{Pyl}tRNA_{CUA}

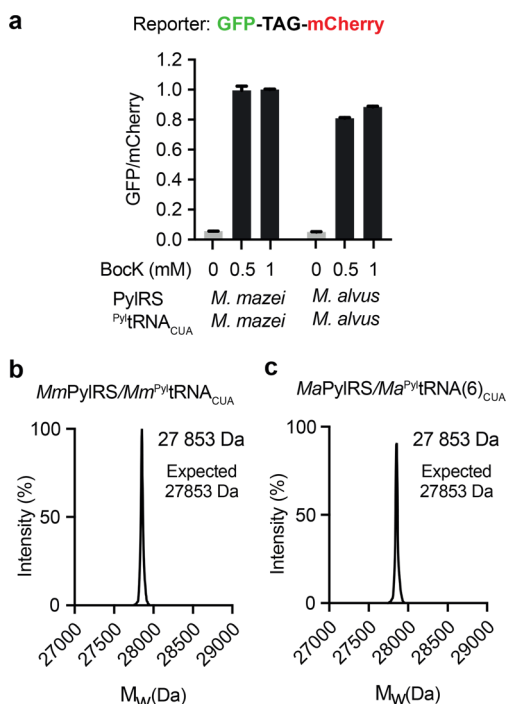


Figure 2. The *MaPylRS/MaPyltRNA(6)_{CUA}* pair is active and orthogonal in mammalian cells. (a) *MaPylRS* and *MmPylRS* show comparable, BocK-dependent, readthrough of the amber stop codon with their cognate tRNAs. Data represent means ± the standard deviation from two biological replicates. (b) ESI-MS of purified sfGFP confirms quantitative incorporation of BocK via the *MmPylRS/MmPyltRNA_{CUA}* pair. (c) ESI-MS of purified sfGFP confirms quantitative incorporation of BocK via the *MaPylRS/MaPyltRNA(6)_{CUA}* pair.

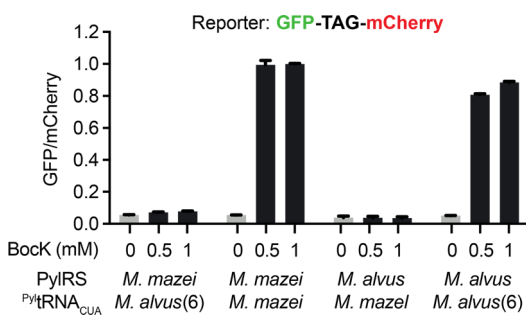


Figure 3. The *MaPylRS/MaPyltRNA(6)_{CUA}* pair is mutually orthogonal to the *MmPylRS/MmPyltRNA_{CUA}* pair in mammalian cells. *MaPylRS* and *MmPylRS* show comparable, BocK-dependent aminoacylation of their cognate tRNA and minimal cross-aminoacylation of the noncognate PyltRNA_{CUA}. Data represent means ± the standard deviation from two biological replicates.

pairs lead to minimal readthrough of the amber stop codon (Figure 3a), while the cognate pairs lead to efficient amber suppression (Figure 2). This demonstrates that these two PylRS/PyltRNA_{CUA} pairs are mutually orthogonal in mammalian cells.

Next we differentiated the active sites of *MaPylRS* and *MmPylRS* such that they selectively recognize distinct substrates. By screening a collection of *MaPylRS* mutants for non-natural substrate specificity in *E. coli*, we discovered a variant of *MaPylRS*, *MaPylRS(mut)*, that incorporates 3-methyl-L-histidine (Me-His) but not BocK; this synthetase contains L121M, L125I, Y126F, M129A, and V168F

mutations. We find that *MmPylRS* directs the incorporation of BocK but not Me-His; these specificities are maintained in mammalian cells (Figure S3a). Finally, we performed a double incorporation of BocK and Me-His into GFP-(101TGA,150TAG) using the *MmPylRS/MmPyltRNA_{CAU}* and *MaPylRS(mut)/MaPyltRNA(6)_{CUA}* pairs. Production of full-length protein was dependent on addition of both ncAAs (Figure S3b), consistent with the site-specific incorporation of both amino acids into GFP in mammalian cells.

We have demonstrated that the *MaPylRS/MaPyltRNA(6)_{CUA}* pair is orthogonal with respect to the synthetases and tRNAs present in mammalian cells and is highly active in mammalian cells. As the *MaPylRS/MaPyltRNA(6)_{CUA}* pair is also orthogonal in *E. coli* and we have demonstrated that mutations discovered in the active site of *MmPylRS* or *MbPylRS* can be transplanted into *MaPylRS* to reprogram its substrate specificity,¹⁹ it will be possible to rapidly expand the substrate scope of the *MaPylRS/MaPyltRNA(6)_{CUA}* pair. Moreover, as *MaPylRS* is a single-domain protein that lacks the poorly soluble N-terminal domain of *MmPylRS*, it may be even more amenable to directed evolution in *E. coli* than the *MmPylRS* system. Our results demonstrate that *MaPylRS* mutants discovered and characterized in *E. coli* will be of direct utility in mammalian cells.

Finally, we have confirmed that the mutual orthogonality of the *MmPylRS/MmPyltRNA* pair and *MaPylRS/MaPyltRNA(6)* pair, which we have characterized in *E. coli*, is maintained in mammalian cells and shown that these pairs can be used together for (unoptimized) double incorporation. These are therefore the first mutually orthogonal pairs in which each pair is itself orthogonal in both *E. coli* and eukaryotic systems. We anticipate that this foundational advance will facilitate multiplexed proteome labeling.^{2–5} Combinations of the advances reported herein, together with strategies for creating additional blank codons and increasing the efficiency of multisite ncAA incorporation in mammalian cells, may facilitate the site-specific incorporation of diverse, and currently inaccessible, combinations of ncAAs into proteins in mammalian cells.^{16–18}

■ ASSOCIATED CONTENT

📄 Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.8b00808.

Experimental methods and supplementary figures (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: chin@mrc-lmb.cam.ac.uk.

ORCID

Jason W. Chin: 0000-0003-1219-4757

Funding

This work was supported by the Medical Research Council, UK (MC_U105181009 and MC_UP_A024_1008), and an ERC Advanced Grant (SGCR), all to J.W.C. V.B. was supported by a MRC Case Studentship (Nikon UK).

Notes

The authors declare no competing financial interest.

■ ABBREVIATIONS

aaRS, aminoacyl-tRNA synthetase; *Ec*, *E. coli*; *Sc*, *S. cerevisiae*; *Ma*, *M. alvus*; *Mb*, *Methanosarcina barkeri*; *Mj*, *Me. janaschii*; *Mm*, *Methanosarcina mazei*; ncAA, noncanonical amino acid; Pul, pyrrolysyl; ESI-MS, electrospray ionization mass spectrometry.

■ REFERENCES

- (1) Chin, J. W. (2014) Expanding and reprogramming the genetic code of cells and animals. *Annu. Rev. Biochem.* 83, 379–408.
- (2) Elliott, T. S., Bianco, A., Townsley, F. M., Fried, S. D., and Chin, J. W. (2016) Tagging and Enriching Proteins Enables Cell-Specific Proteomics. *Cell Chem. Biol.* 23, 805–815.
- (3) Elliott, T. S., Townsley, F. M., Bianco, A., Ernst, R. J., Sachdeva, A., Elsässer, S. J., Davis, L., Lang, K., Pisa, R., Greiss, S., Lilley, K. S., and Chin, J. W. (2014) Proteome labeling and protein identification in specific tissues and at specific developmental stages in an animal. *Nat. Biotechnol.* 32, 465–472.
- (4) Krogager, T. P., Ernst, R. J., Elliott, T. S., Calo, L., Beránek, V., Ciabatti, E., Spillantini, M. G., Tripodi, M., Hastings, M. H., and Chin, J. W. (2017) Labeling and identifying cell-specific proteomes in the mouse brain. *Nat. Biotechnol.* 36, 156–159.
- (5) Hoffmann, J. E., Dziuba, D., Stein, F., and Schultz, C. (2018) A Bifunctional Noncanonical Amino Acid: Synthesis, Expression, and Residue-Specific Proteome-wide Incorporation. *Biochemistry* 57, 4747–4752.
- (6) Sakamoto, K., Hayashi, A., Sakamoto, A., Kiga, D., Nakayama, H., Soma, A., Kobayashi, T., Kitabatake, M., Takio, K., and Saito, K. (2002) Site-specific incorporation of an unnatural amino acid into proteins in mammalian cells. *Nucleic Acids Res.* 30, 4692–4699.
- (7) Chin, J. W., Cropp, T. A., Anderson, J. C., Mukherji, M., Zhang, Z., and Schultz, P. G. (2003) An expanded eukaryotic genetic code. *Science* 301, 964–967.
- (8) Wu, N., Deiters, A., Cropp, T. A., King, D., and Schultz, P. G. (2004) A genetically encoded photocaged amino acid. *J. Am. Chem. Soc.* 126, 14306–14307.
- (9) Rogerson, D. T., Sachdeva, A., Wang, K., Haq, T., Kazlauskaitė, A., Hancock, S. M., Huguenin-Dezot, N., Muqit, M. M., Fry, A. M., Bayliss, R., and Chin, J. W. (2015) Efficient genetic encoding of phosphoserine and its nonhydrolyzable analog. *Nat. Chem. Biol.* 11, 496–503.
- (10) Zhang, M. S., Brunner, S. F., Huguenin-Dezot, N., Liang, A. D., Schmied, W. H., Rogerson, D. T., and Chin, J. W. (2017) Biosynthesis and genetic encoding of phosphothreonine through parallel selection and deep sequencing. *Nat. Methods* 14, 729–736.
- (11) Beránek, V., Reinkemeier, C. D., Zhang, M. S., Liang, A. D., Kym, G., and Chin, J. W. (2018) Genetically Encoded Protein Phosphorylation in Mammalian Cells. *Cell Chem. Biol.* 25, 1–8.
- (12) Iraha, F., Oki, K., Kobayashi, T., Ohno, S., Yokogawa, T., Nishikawa, K., Yokoyama, S., and Sakamoto, K. (2010) Functional replacement of the endogenous tyrosyl-tRNA synthetase-tRNA^{Tyr} pair by the archaeal tyrosine pair in *Escherichia coli* for genetic code expansion. *Nucleic Acids Res.* 38, 3682–3691.
- (13) Hughes, R. A., and Ellington, A. D. (2010) Rational design of an orthogonal tryptophanyl nonsense suppressor tRNA. *Nucleic Acids Res.* 38, 6813–6830.
- (14) Italia, J. S., Addy, P. S., Wrobel, C. J., Crawford, L. A., Lajoie, M. J., Zheng, Y., and Chatterjee, A. (2017) An orthogonalized platform for genetic code expansion in both bacteria and eukaryotes. *Nat. Chem. Biol.* 13, 446–450.
- (15) Italia, J. S., Latour, C., Wrobel, C. J., and Chatterjee, A. (2018) Resurrecting the Bacterial Tyrosyl-tRNA Synthetase/tRNA Pair for Expanding the Genetic Code of Both *E. coli* and Eukaryotes. *Cell Chem. Biol.* 25, 1–9.
- (16) Xiao, H., Chatterjee, A., Choi, S. h., Bajjuri, K. M., Sinha, S. C., and Schultz, P. G. (2013) Genetic incorporation of multiple unnatural amino acids into proteins in mammalian cells. *Angew. Chem., Int. Ed.* 52, 14080–14083.
- (17) Zheng, Y., Addy, P. S., Mukherjee, R., and Chatterjee, A. (2017) Defining the current scope and limitations of dual noncanonical amino acid mutagenesis in mammalian cells. *Chem. Sci.* 8, 7211–7217.
- (18) Zheng, Y., Mukherjee, R., Chin, M. A., Igo, P., Gilgenast, M. J., and Chatterjee, A. (2018) Expanding the Scope of Single- and Double-Noncanonical Amino Acid Mutagenesis in Mammalian Cells Using Orthogonal Polyspecific Leucyl-tRNA Synthetases. *Biochemistry* 57, 441–445.
- (19) Willis, J. C. W., and Chin, J. W. (2018) Mutually orthogonal pyrrolysyl-tRNA synthetase/tRNA pairs. *Nat. Chem.* 10, 831–837.
- (20) Borrel, G., Gaci, N., Peyret, P., O'Toole, P. W., Gribaldo, S., and Brugere, J. F. (2014) Unique characteristics of the pyrrolysine system in the 7th order of methanogens: implications for the evolution of a genetic code expansion cassette. *Archaea* 2014, 374146.
- (21) Jiang, R., and Krzycki, J. A. (2012) PylSn and the homologous N-terminal domain of pyrrolysyl-tRNA synthetase bind the tRNA that is essential for the genetic encoding of pyrrolysine. *J. Biol. Chem.* 287, 32738.
- (22) Herring, S., Ambrogelly, A., Gundllapalli, S., O'Donoghue, P., Polycarpo, C. R., and Söll, D. (2007) The amino-terminal domain of pyrrolysyl-tRNA synthetase is dispensable in vitro but required for in vivo activity. *FEBS Lett.* 581, 3197–3203.
- (23) Schmied, W. H., Elsässer, S. J., Uttamapinant, C., and Chin, J. W. (2014) Efficient multisite unnatural amino acid incorporation in mammalian cells via optimized pyrrolysyl tRNA synthetase/tRNA expression and engineered eRF1. *J. Am. Chem. Soc.* 136, 15577–15583.