Supplementary information for:

**Defining the current scope and limitations of dual noncanonical amino acid mutagenesis in mammalian cell**

Yunan Zheng, Partha Sarathi Addy, Raja Mukherjee and Abhishek Chatterjee*

Department of Chemistry, Boston College, 2609 Beacon Street, Chestnut Hill, MA 02467, USA.

*email: abhishek.chatterjee@bc.edu
Figure S1: Fluorescence images of cells associated with experiments described in Figure 3A. Each aaRS is cotransfected with three different nonsense suppressors of its cognate tRNA and the appropriate EGFP-mutant to evaluate how well each pair suppresses the three different nonsense codons.
Figure S2: Fluorescence images of cells associated with experiments described in Figure 3B. Each of the three nonsense suppressing mutants of tRNA\textsubscript{Pyl} is co-transfected with three different nonsense mutants of EGFP, as well as PylRS, to evaluate if any of the nonsense suppressors can charge a non-cognate nonsense codon.
Figure S3: Fluorescence images of cells associated with the following experiments related to those described in Figure 3B. Each of the three nonsense suppressing mutants of tRNA$^{\text{EcTyr}}$ is co-transfected with three different nonsense mutants of EGFP, as well as EcTyrRS, to evaluate if any of the nonsense suppressors can charge a non-cognate nonsense codon.
Figure S4: Fluorescence images of cells associated with experiments described in Figure 3C. Each of the three different aaRSs was cotransfected with the three different TAG-suppressing tRNAs, as well as EGFP-39-TAG, to identify potential aminoacylation of non-cognate tRNAs.
Figure S5: Maps for different plasmids constructed to facilitate site-specific dual ncAA incorporation.
Figure S6: Site-specific incorporation of different combinations of two distinct ncAAs into EGFP-39-TAG-151-TGA. For each ncAA combination, efficiency of full-length EGFP expression is reported, measured by its characteristic fluorescence in cell-free extract, upon transfecting HEK293T cells with pAcBac3-EcTyr\textsubscript{TGA}-EGFP** and pAcBac1-Pyl\textsubscript{TAG}, or pAcBac3-EcTyr\textsubscript{TGA}-EGFP** and pAcBac1-Pyl\textsubscript{TAG}. While the former suppression systems provides higher dual ncAA incorporation yields in general, the latter system was better when incorporating AzK+PrY.
Figure S7: Fluorescence images of cells associated with experiments described in Figure 5B. Upon transfecting HEK293T cells with pAcBac3-EcLeu$_{\text{TAG}}$-EGFP** and pAcBac1-Pyl$_{\text{TGA}}$, expression of EGFP-39-TAG-151-TGA is visualized in the presence or absence of the indicated ncAAs.
Figure S8: Fluorescence images of cells associated with experiments described in Figure 6A. Site-specific incorporation of two different bioconjugation handles into EGFP-39-TAG-151-TGA. A) Upon transfecting HEK293T cells with pAcBac3-EcTyr\textsubscript{TGA}-EGFP** and pAcBac1-Pyl\textsubscript{TAG} (for hPrK+AzF, AzK+AcF, and CpK+AzF), or pAcBac3-EcTyr\textsubscript{TAG}-EGFP** and pAcBac1-Pyl\textsubscript{TGA} (for AzK+PrY). A representative image of cells in the absence of the ncAAs was also included.
Figure S9: ESI-MS analysis of purified EGFP-39-TAG-151-TGA reporter incorporating indicated ncAAs. Lower panels show raw data, upper panel the deconvoluted mass. In all cases, the observed mass confirms incorporation of the desired ncAAs into indicated sites.
Figure S10: A) Structures of DBCO-Cy5 and AF488-PCA. Treating EGFP-39hPrK-151AzF (B) and EGFP-39PrY-151AzK (C) with DBCO-Cy5 or AF488-PCA (Cu(I)-catalyzed) leads to appropriate fluorescence-labeling, corroborating the presence of the alkyne and azide functionalities. Wild-type EGFP fails to undergo labeling under identical conditions.
Materials and Methods

General materials:

Oligonucleotides and gBlock gene fragments were obtained from IDT. Phusion Hot Start II DNA polymerase, restriction enzymes and T4 DNA ligase were purchased from Fisher Scientific. Various antibiotics were supplemented in media (bacterial culture) at following concentrations: 100 $\mu$g/mL ampicillin, 10 $\mu$g/mL gentamycin, 50 $\mu$g/mL kanamycin. TOP10 E. coli cells are used for cloning and plasmid propagation. For mammalian cell culture, Dulbecco’s modified Eagle’s medium (high glucose DMEM), trypsin 0.25% (1X) solution, Penicillin-Streptomycin (10,000 U/mL) and fetal bovine serum (FBS) were obtained from Fisher Scientific. For transient transfection, we used polyethyleneimine (PEI MAX 40000) from Polysciences (Warrington, PA). For protein purification, TALON His-tag purification resin is obtained from Takara (Clontech). DNA sequencing was performed by Eton Biosciences.

Construction of plasmids to express aaRS, tRNA or EGFP* individually:

The previously reported pAcBac1, pAcBac2 and pAcBac3 vectors were used to generate additional plasmids. To make a plasmid that only expresses *M. bakeri* pyrrolysyl synthetase (MbpYlRS), DNA sequence encoding for MbpYlRS was PCR amplified, digested with NheI/EcoRI and inserted into pAcBac1 vector. The same strategy was also used to generate a pAcBac1 plasmid that expresses *E. coli* leucyl synthetase (EcLeuRS-AK1). Standard site-directed mutagenesis was used on pAcBac1-EGFP-39TAG to create variants that encode EGFP-39TAA or EGFP-39TGA.

DNA sequence encoding one copy of *E. coli* tyrosyl tRNA$_{CUA}$ (tRNA$_{CUA}^{EcTy}$) driven by U6 promoter was PCR amplified from pIDT-Kan-2xtRV, and inserted into pIDT-Kan vector between NheI and AvrII to generate pIDT-Kan-1xU6tRNA$_{CUA}^{EcTy}$. Using site-directed mutagenesis of this plasmid, pIDT-Kan-1xU6tRNA$_{UCA}^{EcTy}$ and pIDT-Kan-1xU6tRNA$_{UUA}^{EcTy}$ were generated, which encodes the TGA- or TAA-suppressing *E. coli* tyrosyl tRNA, respectively. DNA sequences encoding *M. mazei* pyrrolysyl tRNA$_{CUA}$ (tRNA$_{CUA}^{Pyl}$) and *E. coli* leucyl tRNA$_{CUA}$ (tRNA$_{CUA}^{ELeu}$) were synthesized by IDT. gBlock fragments encoding these sequences were PCR amplified, digested with NheI/AvrII and inserted into pIDT-Kan vector to generate pIDT-Kan-1xU6tRNA$_{CUA}^{Pyl}$ and pIDT-Kan-1xU6tRNA$_{CUA}^{ELeu}$. Site-directed mutagenesis was later performed to generate plasmids harboring the corresponding anticodon variants: pIDT-Kan-1xU6tRNA$_{UCA}^{Pyl}$, pIDT-Kan-1xU6tRNA$_{UAA}^{Pyl}$, pIDT-Kan-1xU6tRNA$_{UCA}^{ELeu}$, and pIDT-Kan-1xU6tRNA$_{UUA}^{ELeu}$.

Construction of plasmids for incorporating two ncAAs into EGFP:

To generate plasmids that allow the expression of five essential components (two mutually orthogonal aaRS/tRNA pairs and EGFP harboring stop codons TAG and TGA) for incorporating two distinct ncAAs, we generated four plasmids: 1) For EcTyTGA+PylTAG, we made plasmids
pAcBac3-EcTyr<sub>TGA</sub>-EGFP** (EGFP** stands for EGFP-39TAG-151TGA) and pAcBac1-UbiC-MbPylRS-8xtRNA<sub>CUA<sub>Pyl</sub></sub>. 2) For EcTyr<sub>TAG</sub>+Pyl<sub>TGA</sub>, we made plasmids pAcBac3-EcTyr<sub>TAG</sub>-EGFP** and pAcBac1-CMV-MbPylRS-8xtRNA<sub>UCAPyl</sub>.

Construction of the pAcBac3-EcTyr<sub>TAG</sub>-EGFP** was performed as previously described<sup>1</sup>, except only 16 tRNA copies were incorporated (instead of 20 in the original pAcBac3). The analogous pAcBac3-EcTyr<sub>TGA</sub>-EGFP** plasmid was created in a similar fashion, except only 8 tRNA copies were incorporated. To generate the pAcBac1-CMV-MbPylRS-8xtRNA<sub>UCAPyl</sub> plasmid, a pIDT-Kan-8xtRNA<sub>UCAPyl</sub> plasmid was first prepared containing 8 copies of tandem tRNA<sub>UCAPyl</sub>, each driven by a U6 promoter, following a procedure reported previously<sup>1</sup>. The 8xtRNA cassette was excised out of this plasmid by NheI/AvrII digestion and inserted into the SpeI site of the previously reported pAcBac1-MbPylRS plasmid. An analogous cloning strategy was used to create the corresponding TAG-suppressing plasmid pAcBac1-UbiC-MbPylRS-8xtRNA<sub>UCAPyl</sub>.

For the construction of the pAcBac3-EcLeu<sub>TAG</sub>-EGFP** plasmid, DNA sequence encoding EcLeuRS-AK1<sup>2</sup> was PCR amplified to replace the AnapRS in the previously reported plasmid pAcBac2R-8xEcLtR-AnapRS<sup>3</sup> by NheI and EcoRI restriction enzyme sites. EGFP** was then PCR amplified and inserted into the SfiI sites to make the final plasmid.

**EGFP** fluorescence analysis, expression and purification:

For small-scale expression analysis, HEK293T cells were seeded at a density of 600,000 cells per well for a 12-well plate the day before transfection. A total amount of 1.2 μg DNA + 4 μl PEI + 20 μl DMEM was used for transfection of each well. For the modular three-plasmid transfection, 0.4 μg of each plasmid is used. For two-plasmid transfections, 0.6 μg of each plasmid is used. Fluorescence images and EGFP expression analysis were performed 48 hours post transfection. To obtain EGFP expression data, cells were harvested and lysed as described before<sup>1</sup>. EGFP fluorescence in lysate was collected in a 96-well plate using a SpectraMAX M5 (Molecular Devices) (ex=480 nm and em=530 nm). Mean of three independent experiments were reported, and error bars represent standard deviation. For larger scale protein expression incorporating two ncAAs, HEK293T cells were seeded in 100 mm cell culture dishes (5 million per dish) two days before transfection. A total amount of 12 μg DNA + 50 μl PEI + 200 μl DMEM was used to transfect cells that are between 80% - 90% confluency. The ncAAs were supplemented at 1 mM final concentration at the time of transfection. Cells were harvested 48 hours post-transfection, lysed with CelLytic M, and the C-terminally polyhistidine tagged protein was purified using cobalt-containing TALON metal affinity resin (Clontech) following manufacturer’s protocol. Purified proteins were subjected to SDS-PAGE gels and ESI-MS (Agilent TOF HPLC-MS) for further characterization.
Protein labeling, and MS analysis:

For click-labeling the EGFP** protein incorporating either AzF+hPrK or PrY+AzK, 1 μg of purified EGFP** was incubated with 20 μM DBCO-Cy5 (Sigma) or 50 μM Alexa Fluor 488 picolyl azide (Fisher Scientific), respectively, for 2 hours at room temperature. Identical reactions were set up with wild-type EGFP as control experiments. For labeling EGFP-39CpK-151AzF, 5 μM of protein was incubated with 50 μM DBCO-TAMRA (Click Chemistry Tools, AZ) followed by 200 μM tetrazine-fluorescein, at room temperature in PBS for 30 min each. Subsequently, the reaction product was subjected to SDS-PAGE and HPLC-coupled ESI-MS analysis.

List of Primers:

EcLtR-Nhe-R: aataatGCTAGCAAAAAATACCCGGAGCGGGACTTTGAACC
MmPytR-Nhe-R: aataatgtagcAAAAAACGGAAAAACCCCGGAATCTAACCCCG
EcYtR-Nhe-R: aataatGCTAGCAAAAATGGGTGGGGAAGGATTCG
BstYtR-Nhe-R: aataatGCTAGCAAAAAATGGAGGGGGACGGATTCG
U6-AvrII-F: aataatCCTAGGTGGCCGCAAGAAGGAGGC
H1-AvrII-F: aataatCCTAGGAATTCGACGTCAACATCAACC
EGFP-SfiI-F: aataatGAATTGGCCAAGGCCGGGCAACCATTG
12xHis-SfiI-R: AATTCGGCCTTAAATGATGGTGGTGGATGGTGGTGGATGGTGGTGGATGACC
LeuRS-Nhel-F: aataatGGCTAGCGTTTAAACTTAAGCTTGCCGCCACCATGGATGACG
LeuRS-EcoRI-R: attattaGAATTCTTAACCGGGCCCGCAACGAGATTGAG

References:

1. Y. Zheng, T. Lewis Jr, P. Igo, F. Polleux and A. Chatterjee, *ACS Synth. Biol.*, 2017, 6, 13-18.
2. H. W. Ai, W. Shen, E. Brustad and P. G. Schultz, *Angew. Chem. Int. Ed.*, 2010, 49, 935-937.
3. A. L. Mitchell, P. S. Addy, M. A. Chin and A. Chatterjee, *ChemBioChem*, 2017, 18, 511-514.
4. K. Lang, L. Davis, S. Wallace, M. Mahesh, D. J. Cox, M. L. Blackman, J. M. Fox and J. W. Chin, *J. Am. Chem. Soc.*, 2012, 134, 10317-10320.
Plasmids sequences:

pAcBac3-EcTyr\_TGA-EGFP**: EcTyrRS sequence is colored orange, EGFP** is colored green. U6 promoter is colored blue. tRNA\textsubscript{UCA}\textsuperscript{EcTyr} is colored red. H1 promoter is colored magenta. tRNA\textsubscript{UCA}\textsuperscript{BsrTyr} is colored cyan.
agcttgagcaagcactatacgagtaacactacagcgtgagcattgagaaagcgccacgcttcccgaagggagaaaggcggacaggtatc
cggtaagcggcagggtcggaacaggagagcgcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctga
ccttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggcctttttacggttcctggccttttgctggccttttg
catacatgttctttcctgatccctgatcttcttcgttaagttacatattatgcgagctgatcgtttggctagcttcttttttactttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
pAcBac1-UbiC-MbPylRS-8xRNA<sub>CUA</sub><sup>Pyl</sup>: MbPylRS sequence is colored orange. U6 promoter is colored blue. tRNA<sub>CUA</sub><sup>Pyl</sup> is colored red. H1 promoter is colored magenta.
pAcBac3-EcLeuTAG-EGFP**: EcLeuRS-AK1 is colored orange. EGFP** is colored green.
tRNA_{CUA}^{EcLeu} is colored red. H1 promoter is colored magenta.
