Mutually orthogonal pyrrolysyl-tRNA synthetase/tRNA pairs

Supplementary Information

Julian C. W. Willis¹, Jason W. Chin^{1,2}

¹Medical Research Council Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge, England, UK.

²Correspondence: chin@mrc-lmb.cam.ac.uk

Supplementary Information: Index

Page 3-8, Methods

Page 9, Supplementary Figure 1	Cladogram of PylRS classes
Page 10, Supplementary Figure 2	Alignment of PyIRS sequences
Page 11, Supplementary Figure 3	^{Pyl} tRNA sequences and predicted secondary structures
Page 12, Supplementary Figure 4	Analysis of purified GFPHis ₆ for ncAA incorporation using orthogonal PylRS/ ^{Pyl} tRNA pairs
Page 13, Supplementary Figure 5	<i>Ma</i> PylRS/ <i>Ma</i> ^{Pyl} tRNA decoding with altered anticodons
Page 14, Supplementary Figure 6	The substrate selectivity of <i>Mm</i> PylRS-MutRS2
Page 15, Supplementary Figure 7	ncAA incorporation using mutually orthogonal PylRS/ ^{Pyl} tRNA pairs
Page 16, Supplementary Figure 8	Analysis of distinct ncAA incorporation using mutually orthogonal PyIRS/ ^{PyI} tRNA pairs
Page 17, Supplementary Figure 9	Analysis dual ncAA incorporation using mutually orthogonal PylRS/ ^{Pyl} tRNA pairs

Methods

Identification of PyIRS/^{PyI}tRNA sequences

We identified protein sequences homologous to the C-terminal region of MmPyIRS by protein BLAST¹ using $MmPyIRS\Delta184$ as the query sequence and filtering for Expect values below 1×10^{-10} . We next identified protein sequences homologous to DhPyISn by protein BLAST using DhPyISn as the query sequence and filtering for Expect values below 1×10^{-10} . We sorted identified protein sequences which are homologous to the C-terminal region of MmPyIRS according to whether we found sequence homology to DhPyISn within the same protein sequence (pyISn-pyISc fusion class), within the same genome in a separate protein sequence (pyISn class), or not within the same genome (Δ pyISn class). We generated protein sequence alignments and percentage identity scores using Clustal Omega², discarding duplicate sequences with a percentage identity score above 98%. Protein sequence alignments were visualised using Clustal X³ and phylogenetic trees generated by phyloT were visualised using ITOL⁴.

DNA constructs

PylRS and ^{Pyl}tRNA genes were synthesised by IDT and we inserted the genes into pKW vectors by Gibson cloning to be expressed from constitutively active *glmS* and *lpp* promoters respectively. We appended the gene for *Mm*PylRS with a sequence encoding a C-terminal Ser(Gly₄Ser)₄FLAG-tag, while all other PylRS genes were appended with a sequence encoding a C-terminal Ser(Gly₄Ser)₄His₆SerGlyStrep-tag II. We used these plasmids together with pBAD GFP(150TAG)His₆ (in which sfGFP containing an amber stop codon at position 150 and a C-terminal His₆ tag is expressed from the arabinose promoter of pBAD; GFP refers to sfGFP throughout). We also created versions of pKW *Ma*PylRS/^{Pyl}tRNA_{CUA} and pKW *GI*PylRS/^{Pyl}tRNA_{CUA} in which the PylRS gene contains no C-terminally appended sequence. *Mm*PylRS-AF, *Ma*PylRS-AF, *Mm*PylRS-MutRS2 and *Ma*PylRS-MutRS1 genes were synthesised by IDT and we inserted the genes into pKW vectors⁵ by Gibson cloning.

For library generation of Ma^{Pyl} tRNA with randomised variable loop sequences, Ma^{Pyl} tRNA genes containing nucleotide positions 41, 42 and 43 replaced with either NNN, NNNN, NNNNN or NNNNNN were synthesised by IDT and we inserted the genes into an empty pKW vector by Gibson cloning. We transformed each library separately into competent *E. coli* DH10B cells to give a library diversity of 1×10^6 ,

exceeding the theoretical diversity of 4×10^4 required for complete library coverage. The libraries of Ma^{Pyl} tRNA plasmids were used together with pBAD MaPylRS GFP(150TAG)His₆ CAT(111TAG) and pBAD MmPylRS GFP(150TAG)His₆ CAT(111TAG). We created these plasmids by transferring the relevant PylRS cassette into the pBAD GFP(150TAG)His₆ vector together with a CAT(111TAG) gene under constitutive expression (chloramphenicol acetyl transferase containing an amber stop codon at position 111).

We created pKW-DUAL plasmids to assess the mutual orthogonality of the PylRS/^{Pyl}tRNA pairs in the same cell and to enable incorporation of distinct ncAAs into a single polypeptide. We created these plasmids by transferring the relevant *Ma*PylRS/*Ma*^{Pyl}tRNA cassette into pKW *Mm*PylRS/*Mm*^{Pyl}tRNA vectors by Gibson cloning. The pRSF vector containing ribo-Q1 rRNA with o-GST-CaM(1TAG, 40AGGA) was previously reported.

Measuring the activity and specificity of PyIRS/^{PyI}tRNA_{CUA} pairs

To measure the activity and specificity of cognate and non-cognate PyIRS/^{PyI}tRNA combinations we transformed pKW PyIRS/^{PyI}tRNA plasmids into competent *E. coli* DH10B cells bearing pBAD GFP(150TAG)His₆. We recovered the transformed cells for 1 h at 37°C in 0.5 ml super optimal broth (SOB) medium. The transformation was used to inoculate 6 ml 2XTY-ST (2XTY medium with 37.5 μ g ml⁻¹ spectinomycin and 12.5 μ g ml⁻¹ tetracycline) and incubated overnight (37°C, 16h, 220 rpm). The overnight culture was used to inoculate 190 μ l 2XTY-STA (2XTY medium with 37.5 μ g ml⁻¹ spectinomycin, 12.5 μ g ml⁻¹ tetracycline and 0.2% arabinose) in 96-well microtitre plate format, supplemented with or without the indicated ncAA (8 mM BocK (1), 8 mM BCNK (2), 3 mM CbzK (3), or 2 mM CypK (4)). OD₆₀₀ and GFP fluorescence (λ_{ex} 485 nm, λ_{em} 520 nm) measurements were recorded over a 20 h time course at 10 min intervals using a Tecan Infinite M200 Pro at 37°C, 220 rpm. At the maximum OD₆₀₀ for each sample the GFP fluorescence measurement was normalised by the OD₆₀₀ measurement to give a value for the amber suppression activity in normalised fluorescence units.

To express GFP incorporating BocK (1) for mass spectrometry we transformed pKW $PyIRS^{Pyl}tRNA$ plasmids in which PyIRS contains no C-terminally appended sequence into competent *E. coli* DH10B cells bearing pBAD GFP(150TAG)His₆. We recovered the transformed cells for 1 h at 37°C in 0.5 ml SOB medium. The

transformation was used to inoculate 2XTY-STA (2XTY medium with 37.5 μ g ml⁻¹ spectinomycin, 12.5 μ g ml⁻¹ tetracycline and 0.2% arabinose) supplemented with 8 mM BocK (1) and incubated overnight (37°C, 16 h, 220 rpm). 50 ml culture was pelleted, washed with 800 μ l PBS and flash frozen. The cell pellets were resuspended in 1 ml of lysis buffer (1X BugBuster Protein Extraction Reagent supplemented with 1X cOmplete protease inhibitor cocktail, 1 mg ml⁻¹ lysozyme and 1 mg ml⁻¹ DNase I) and lysed for 1 h at 25°C, 1,400 rpm. The lysate was clarified by centrifugation (21,000 g, 30 min, 4°C) and imidazole added to a final concentration of 10 mM. GFP was purified by its C-terminal His₆ tag using 100 μ l Ni-NTA agarose beads and left to bind for 1 h at 4°C. The beads were washed four times with 800 μ l PBS supplemented with 250 mM imidazole. The three eluting fractions were pooled together to reach a combined volume of 180 μ l. The imidazole was removed using 7K MWCO Zeba spin desalting columns and the samples were then concentrated using an Eppendorf Concentrator Plus and analysed by electrospray ionisation mass spectrometry.

To express GFP incorporating CbzK (2) or CypK (3) we transformed pKW-DUAL PylRS/^{Pyl}tRNA plasmids into competent E. coli DH10B cells bearing pBAD GFP(150TAG)His₆. We recovered the transformed cells for 1 h at 37°C in 0.5 ml SOB medium. The transformation was used to inoculate 2XTY-STA (2XTY medium with 37.5 μ g ml⁻¹ spectinomycin, 12.5 μ g ml⁻¹ tetracycline and 0.2% arabinose) supplemented with or without the indicated ncAAs (3 mM CbzK (3), or 2 mM CypK (4)) and incubated overnight (37°C, 16 h, 220 rpm). 50 ml culture was pelleted, washed with 800 µl PBS and flash frozen. The cell pellets were resuspended in 1 ml of lysis buffer (1X BugBuster Protein Extraction Reagent supplemented with 1X cOmplete protease inhibitor cocktail, 1 mg ml⁻¹ lysozyme and 1 mg ml⁻¹ DNase I) and lysed for 1 h at 25°C, 1,400 rpm. The lysate was clarified by centrifugation (21,000 g, 30 min, 4°C). PylRS was removed by its C-terminal Strep-tag II using 100 µl StrepTactin sepharose beads and left to bind for 1 h at 4°C. Imidazole was added to the recovered supernatant to a final concentration of 10 mM. GFP was purified by its C-terminal His₆ tag using 100 µl Ni-NTA agarose beads and left to bind for 1 h at 4°C. The beads were washed four times with 800 µl PBS supplemented with 10 mM imidazole and eluted in three volumes of 60 µl PBS supplemented with 250 mM imidazole. The three eluting fractions were pooled together to reach a combined volume of 180 µl. Samples were both analysed by electrospray ionisation mass

spectrometry and analysed on 4-12% Bis-Tris SDS-PAGE gels, visualised with InstantBlue Coomassie stain and imaged using a ChemiDoc XRS+.

Discovering orthogonal Ma^{Pyl}tRNA variants

To discover Ma^{Pyl}tRNA variants orthogonal to MmPylRS, we transformed the Ma^{Pyl}tRNA libraries into competent E. coli DH10B cells bearing pBAD MaPylRS CAT(111TAG) GFP(150TAG)His₆. We recovered the transformed cells for 1 h at 37°C in 0.5 ml SOB medium supplemented with 8 mM BocK (1). The transformation was plated on LB agar containing 37.5 µg ml⁻¹ spectinomycin, 12.5 µg ml⁻¹ tetracycline and 100 µg ml⁻¹ chloramphenicol. The plates were incubated at 37°C for 40 h. After incubation, colonies on the plates were washed off and collected in 2XTY buffer and the plasmids were extracted by DNA midiprep. To remove the pBAD MaPyIRS CAT(111TAG) GFP(150TAG)His₆ plasmid, the extracted DNA was digested with both NcoI restriction endonuclease and T5 exonuclease and re-purified using a PCR purification column. The remaining pKW plasmids were transformed into competent E. coli DH10B cells bearing pBAD MmPylRS CAT(111TAG) GFP(150TAG)His₆. The transformed cells were recovered for 1 h at 37°C in 0.5 ml SOB medium. The transformation was plated on LB agar containing 37.5 µg ml⁻¹ spectinomycin and 12.5 µg ml⁻¹ tetracycline. The plates were incubated at 37°C for 20 h. For each library, 376 colonies were picked from the plates using a QPix 420 Colony Picking System and inoculated into 190 µl 2XTY-STA in 96-well microtitre plate format supplemented with 8 mM BocK (1). The plates were incubated at 37°C, 220 rpm and OD₆₀₀ and GFP fluorescence (λ_{ex} 485 nm, λ_{em} 520 nm) measurements were recorded after 20 h using a SpectraMax i3. Cells from 96 wells with the lowest GFP/OD₆₀₀ ratios were used to inoculate 2XTY medium with 75 μ g ml⁻¹ spectinomycin, and the pKW plasmids containing Ma^{Pyl} tRNA variants were extracted by DNA miniprep and then sequenced. Each hit corresponding to a distinct Ma^{Pyl}tRNA sequence was cloned into a pKW MaPylRS vector and a pKW MmPylRS vector and re-phenotyped with pBAD GFP(150TAG)His₆.

Incorporating distinct ncAAs

To express GST-CaM proteins incorporating two distinct ncAAs we co-transformed competent *E. coli* DH10B cells with pKW-DUAL *Mm*PylRS/*Mm*^{Pyl}tRNA_{UCCU} *Ma*PylRS-MutRS1/*Ma*^{Pyl}tRNA(6)_{CUA} and pRSF ribo-Q1 o-GST-CaM(1TAG,

40AGGA). We recovered the transformed cells for 1 h at 37°C in 0.5 ml SOB medium. The transformation was used to inoculate 6 ml LB-KS (LB medium with 25 μ g ml⁻¹ kanamycin and 37.5 μ g ml⁻¹ spectinomycin) and incubated overnight (37°C, 16 h, 220 rpm). The overnight culture was diluted to OD₆₀₀ 0.1 in LB-KS and incubated at 37°C, 220 rpm. At OD₆₀₀ 0.5, IPTG was added to a final concentration of 1 mM together with or without the indicated ncAAs (3 mM CbzK (3) or 2 mM CypK (4)). After 4 h incubation at 37°C, 220 rpm, 50 ml culture was pelleted, washed with 800 µl PBS and flash frozen. The cell pellets were resuspended in 1 ml of lysis buffer (1X BugBuster Protein Extraction Reagent supplemented with 1X cOmplete protease inhibitor cocktail, 1 mg ml⁻¹ lysozyme and 1 mg ml⁻¹ DNase I) and lysed for 1 h at 25°C, 1,400 rpm. The lysate was clarified by centrifugation (21,000 g, 30 min, 4°C). GST-containing proteins from the lysate supernatant were left to bind to 70 µl glutathione sepharose beads for 1 h at 4°C. The beads were washed four times with 800 µl PBS before eluting in 180 µl 1X NuPAGE LDS sample buffer supplemented with 100 mM DTT at 95°C for 5 min. Samples were analysed on 4-12% Bis-Tris SDS-PAGE gels, visualised with InstantBlue Coomassie stain and imaged using a ChemiDoc XRS+.

To express GFPHis₆Strep proteins incorporating two distinct ncAAs we transformed competent *E. coli* DH10B cells with pKW-DUAL *Mm*PylRS/*Mm*^{Pyl}tRNA_{UCCU} *Ma*PylRS-MutRS1/*Ma*^{Pyl}tRNA(6)_{CUA} and pRSF ribo-Q1 and po-GFP(40AGGA, 150TAG)His₆Strep. We recovered the transformed cells for 1 h at 37°C in 0.5 ml SOB medium. The transformation was used to inoculate 20 ml 2XTY-KST (2XTY medium with 20 μ g ml⁻¹ kanamycin, 30 μ g ml⁻¹ spectinomycin and 10 μ g ml⁻¹ tetracycline) and incubated overnight (37°C, 16 h, 220 rpm). The overnight culture was diluted 1:50 in LB-KST and IPTG was added to a final concentration of 1 mM together with or without the indicated ncAAs (3 mM CbzK (**3**) or 2 mM CypK (**4**)). After 16 h incubation at 37°C, 220 rpm, 50 ml culture was lysed and used to purify GFPHis₆Strep as described above but instead using 70 μ l StrepTactin sepharose beads and eluting in three volumes of 60 μ l PBS supplemented with 2.50 mM desthiobiotin. The three eluted fractions were pooled together into a combined volume of 180 μ l and analysed by electrospray ionisation mass spectrometry.

Electrospray ionization mass spectrometry

Mass spectra for protein samples were acquired using an Agilent 1200 LC-MS with a 6130 Quadrupole spectrometer. The solvent system used for liquid chromatography was 0.2% formic acid in water as buffer A and 0.2% formic acid in acetonitrile as buffer B. Samples were injected into a Phenomenex Jupiter C4 column (150 x 2 mm, 5 μ m) and subsequently into the mass spectrometer. Spectra were acquired in the positive mode and analyzed using OpenLAB CDS Chemstation software. The deconvolution program provided in the software was used to obtain the mass spectra. The theoretical molecular weights of proteins with ncAAs was calculated by first computing the theoretical molecular weight of wild-type protein using an online tool (http://web.expasy.org/protparam/) and then manually correcting for the theoretical molecular weight of ncAAs.

References

- 1. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. Basic local alignment search tool. *J Mol Biol* **215**, 403-410 (1990).
- 2. Sievers, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539 (2011).
- 3. Larkin, M.A. et al. Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948 (2007).
- 4. Letunic, I. & Bork, P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res* **44**, W242-245 (2016).
- Wang, K. et al. Optimized orthogonal translation of unnatural amino acids enables spontaneous protein double-labelling and FRET. *Nature chemistry* 6, 393-403 (2014).



Supplementary Figure 1. Cladogram derived from the phylogenetic analysis of identified PylRS-containing species using phyloT, visualised using iTOL. PylRS sequences are classified and coloured according to the presence of a domain homologous to *D. hafniense* PylSn either within the same gene (*pylSn-pylSc* fusion class, green), present within a separate gene (*pylSn* class, black), or absent entirely ($\Delta pylSn$ class, blue). Labels are of the form: Species name (Taxonomy ID) [Protein sequence accession number(s)].

MmPylRS	MDKKPLNTLISATGLWM <mark>SRTGTIHKIKHHEVSRSKIYIEMACGDH</mark> LVVNNSRSSRTARALRHHKYRKTCKRCRVSDEDLN	80
MbPylRS	MDKKPLDVLISATGLWMSRTGTLHKIKHHEVSRSKIYIEMACGDHLVVNNSRSCRTARAFRHHKYRKTCKRCRVSDEDIN	80
MaPylRS		
G1Py1RS		
H5PylRS		
MtPylRS		
MlPylRS		
1		
_		
MmPylRS	KFLTKANEDQTSVKVKVVSAPTRTKKAMPKSVARAPKPLENTEAAQAQPSGSKFSPAIPVSTQESVSVPASVSTSISSIS	160
MbPylRS	NFL <mark>TRSTES</mark> KN <mark>SVKVRVVSAP</mark> K-V <mark>KK</mark> AMPKSVSRAPKPLENSVSAKASTNTSRSVPSP	137
MaPylRS		
G1PylRS		
H5PylRS		
MtPylRS		
MlPylRS		
	* *	
MmD1DC		225
MMPYIRS	TGATASALVKGNINPITSMSAPVQASAPALTKSOTORLEVLLNPKDEISLNSGKPFRELESELLSRKKKDLQQIT	235
MDPYIRS	KISLMARPRESVPASAPAPSLIRSOLDRVEALLSPEDKISLMARPFRELEPELVIRKKNDFORL	200
MAPYIRS		55
GIPYIRS	WVVKFTDSQ10HLMEYGDNDWSEAEFEDAAARDKEFSSQFSKLKSANDKGLKDV1	55
HSPYIRS	TCKLTDP010RLREYGHEPKNESEFETEEERDKAFTKMMSKLQRENEKGIRDM1	55
MEPYIRS	MSIGFTPSQIQKLREFGEDPRDNSTYQNVEQRDKAFSKLMSDLVSSNEKEIAGML	55
MIPyIRS	PADTRLTPAQAQRIREMGGTVDPSLAFSSEAERESAFQRISADLQGANLAKIRRCA	55
	· * · · **·* · · · * * · · · * * · * ·	
Mmpylps	AFFERING CHERETTEREVISED FT FT SEAT FOR THAT FT SEAT FRUNKNE'T DOMLA DNI VNVT BEL DRA	315
Mhpylpg	THEFT I AND AND THE PROPERTY AND	280
Mapulpg	AND SOUCH TOT MAN TANAT VERST FUT OF AN A DAME TO TANA TANAY AND	135
C1Dw1PS	AND - NDI TOI ENVIROPET A DEFENSION OF TEVERSAL AVMETTENDED TO TEVENTONE ADMIT DOWN AND TONE	134
UED UIDS	AND AND IDLENNI KEKLAANGI ISVALE IFVANDALANI I EDINE LENGVENI DDARALAFMAANA I KUMATALAF	125
MEDUIDE		125
MUPYIKS		135
MIPYIKS	EAPERNPIGSLENTLACALAARGFIEVRTPMMIPADGLVRMGIDESNPLWNQVFWVGPRRALRPMLAPNLIFIMRALRRS	135
	▼ ** :**:* ** . *******: : * :: * .:. : : : ******: .	
MmPylRS	LPDPIKIFEIGPCYRKESDGKEHLEEFTMLNFCOMGS-GCTRENLESIITDFLNHLGI-DFKIVGDSCMVYGDTLDVMHG	393
MbPvlRS	LPGPIKIFEVGPCYRKESDGKEHLEEFTMYNFCOMGS-GCTRENLEALIKEFLDYLEI-DFETVGDSCMVYGDTDTMHG	358
MaPvlRS	TOGPVKIFEMGSCFRKESHSGMHLEEFTMLNLUDMGPRGDATEVLKNYISVVMKAAGI.PDYDLUGESDVYKETIDVEIN	215
GIPVIRS	TKGPVKIFEIGSCFRKESKSSTHLEEFTMLNLVEMGPDGDPMEHLKMYIGDIMDAVGV-EYTTSREESDVYVETLDVEIN	213
H5PvlRS	TO GPVKFFE I GSOF BAESHSNDHI, BEFTMI, NI, VDMGPOGDTTEK I KHY I DI VMKTIGI, - DVELVHEESDVYKET I DVEVD	214
M+PvlRS	TO GPVKIFEIGSCERKESHSGSHIEEFTMINI, VELGPEGDATEALKDHIGAVMKVTGI. EVTLVREESDVVVETLDVEID	214
MIDVIRS	VDADILIEFT GOOFDWESD GSNHLEEFTMINT VELADOADATEDIWEHTATWNAVGI. DVELVVEGSEVVCTTIDVEVD	214
MIFYIKS	VERTILLE ELOFOF KREAKSANDEEF TALKEVELAF GADATEKENENTAT VARAVGE FIELVVEGSEVIGI TIVVEVD	214
	. *:.*. :* .: :** * *****: : . :: . * * **	
MmPylRS	DLELSSAVVGPIPLDREWGIDKPWIGAGFGLERLLKVKHDFKNIKRAARSESYYNGISTNL 454	
MbPylRS	DLELSSAVVGPVSLDREWGIDKPWIGAGFGLERLLKVMHGFKNIKRASRSESYYNGISTNL 419	
MaPylRS	GQEVCSAAVGPHYLDAAHDVHEPWSGAGFGLERLLTIREKYSTVKKGGASISYLNGAKIN- 275	
G1PylRS	GTEVASGAVGPHKLDPAHDVHEPWAGIGFGLERLLMLKNGKSNARKTGKSITYLNGYKLD- 273	
H5PylRS	GEEVCSAAVGPHYLDKAHNINEPWCGAGFGLERLIMMRDGDGSVKKTGKSVNYLNGYKIN- 274	
MtPylRS	GGEVASGAVGPHVLDNAHDIHEPWSGIGFGLERLLMIMNEKSTVKKTGRSLSVLNGAKIN- 274	
MlPylRS	GVELASGAVGPLPMDKPHGITEPWAGVGFGLERIALMRTKEQNIKKVGRSLVYVNGARIDI 275	

Supplementary Figure 2. Alignment of the indicated PyIRS sequences using Clustal Omega, visualised using Clustal X. Active site residues referenced in the main text are indicated by triangles.



Supplementary Figure 3. ^{Pyl}tRNA sequences from the indicated genomes, and their predicted secondary structures.



Supplementary Figure 4. Raw ESI-MS spectra for the data shown in Fig. 2 c,d,e are shown in panels a,b,c respectively. ESI-MS experiments in (a-c) were each performed once.



Supplementary Figure 5. *Ma*PylRS/*Ma*^{Pyl}tRNA decoding with altered anticodons. *o-GST-CaM(1UAG)*, *o-GST-CaM(1UAGA)* and *o-GST-CaM(1AGUA)* genes were decoded on ribo-Q by *Ma*PylRS/*Ma*^{Pyl}tRNA_{CUA}, *Ma*PylRS/*Ma*^{Pyl}tRNA_{UCUA}, *Ma*PylRS/*Ma*^{Pyl}tRNA_{UACU} respectively. Two independent experiments were performed with similar results.



Supplementary Figure 6. Engineering the active site of MmPyIRS for selective ncAA incorporation. *In vivo* amber suppression activity assay using *E. coli* DH10B bearing pBAD GFP(150TAG)His₆ and the corresponding pKW PyIRS/^{PyI}tRNA_{CUA} plasmid in the presence and absence of CbzK (**3**) and CypK (**4**) demonstrates the selective incorporation of CypK by MmPyIRS-MutRS2 is comparable to that with MmPyIRS. Each data point shows the mean of three technical replicates that form one biological replicate; the error bars show the mean and SEM of three independent biological replicates.



Supplementary Figure 7. Encoding distinct ncAAs using mutually orthogonal PylRS/^{Pyl}tRNA pairs. Full gels for **Fig. 6 a, c, e**. are shown in panels **a**, **b** and **c** respectively. Densitometry quantification of the four lanes in panel **a** gives the ratio of relative band intensities as 1 : 17 : 6 : 22. For each panel, two independent experiments were performed with similar results.



Supplementary Figure 8. Raw ESI-MS spectra for the data shown in **Fig 6b**, **d** are shown in panels **a** and **b** respectively. ESI-MS experiments in panels **a** and **b** were each performed once.



Supplementary Figure 9. GFPHis₆Strep purified from *E. coli* DH10B bearing pBAD GFP(40AGGA,150TAG)His₆Strep and pKW-DUAL *Mm*PylRS/^{Pyl}tRNA_{UCCU} *Ma*PylRS-MutRS1/^{Pyl}tRNA(6)_{CUA} in the presence of both CbzK (**3**) and CypK (**4**) leads to selective incorporation of distinct ncAAs at distinct codons. Electrospray ionisation mass spectrometry shows a single peak corresponding to CbzK and CypK incorporation (predicted mass 29,285 Da, observed mass 29,285 Da). Raw ESI-MS spectrum before de-convolution is shown in panel **b**. ESI-MS experiments in panels **a** and **b** were each performed once.