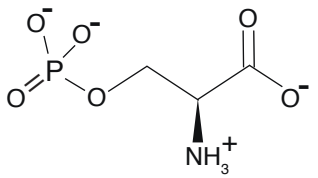
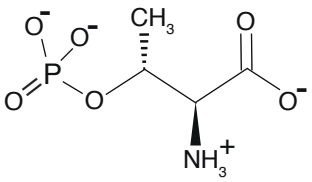
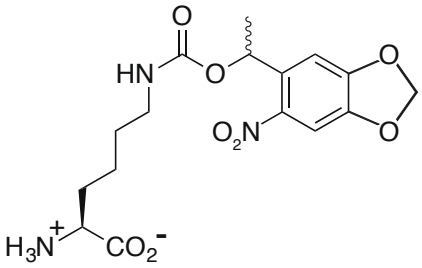


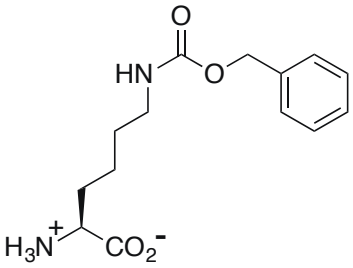
1



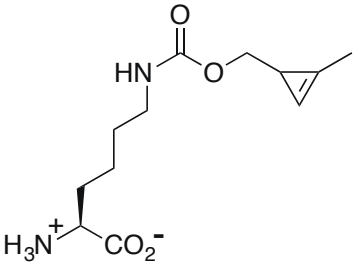
2



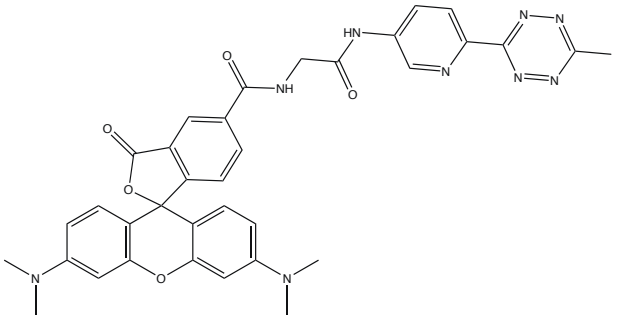
3



4

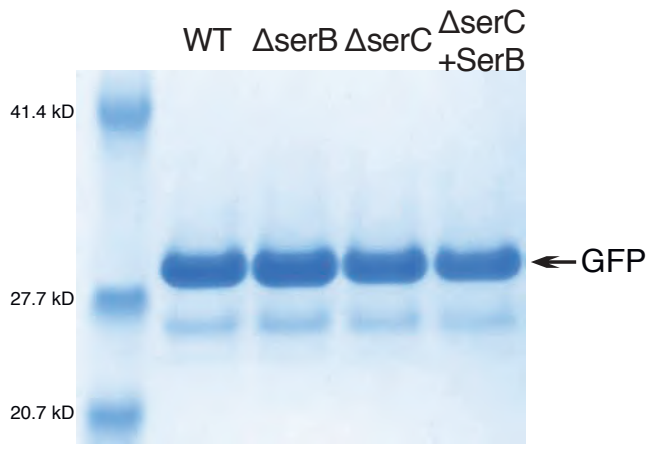


5

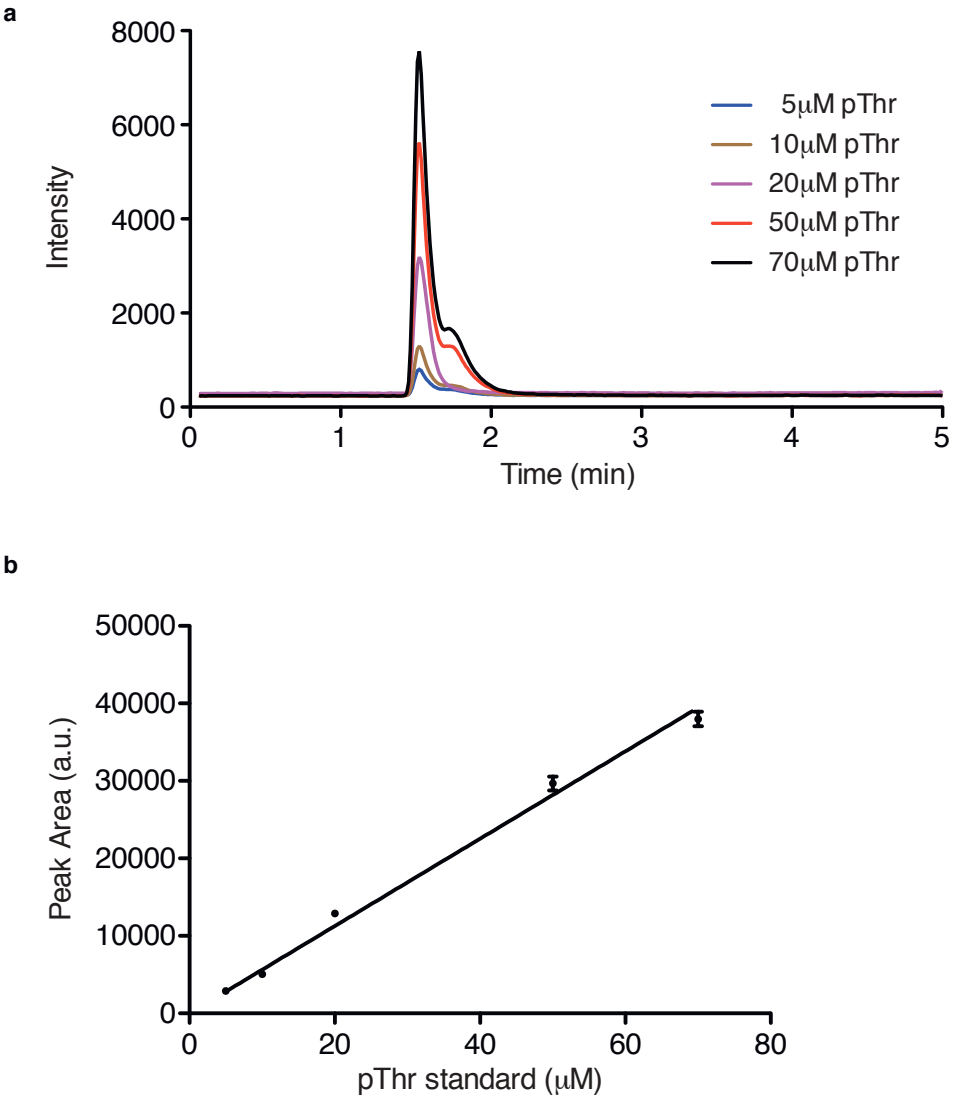


6

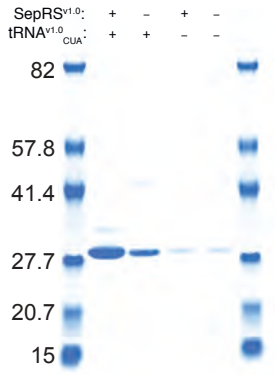
Supplementary Figure 1. Molecular structures of unnatural amino acids used in this study. O-phospho-L-serine (pSer, **1**); O-phospho-L-threonine (pThr, **2**); Nε-((1-(6-nitrobenzo[*d*][1,3]dioxol-5-yl)ethoxy)carbonyl)-L-lysine(PCK,**3**); Nε-(carbobenzoyloxy) -L-lysine (CbzK, **4**); Nε-(((2-methylcycloprop-2-en-1-yl) methoxy)carbonyl)-L-lysine (CypK, **5**); Tet1-TAMRA, **6**.



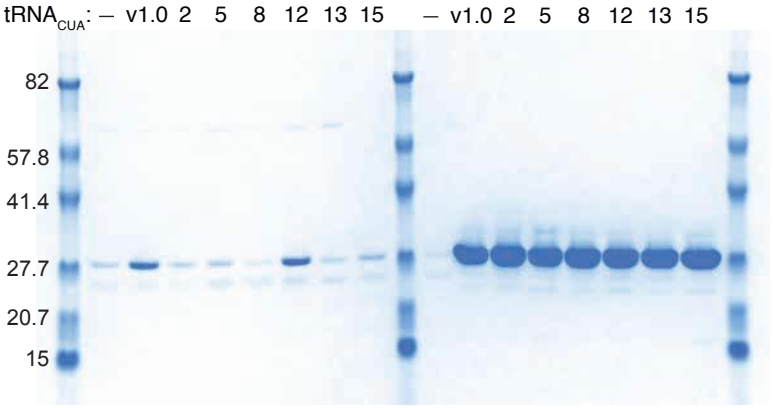
Supplementary Figure 2. pSer is present in *E.coli* at sufficient concentration to be incorporated into proteins with the SepRS$^{V1.0}$/tRNA$^{V1.0}_{CUA}$ pair. sfGFP(150TAG)-His6 was transformed into DH10B wild type, DH10B Δ*serB*, or DH10B Δ*serC* cells with an pKW vector containing SepRS$^{V1.0}$/tRNA$^{V1.0}_{CUA}$ pair and a pCDF vector containing EF-Sep. In DH10B Δ*serC* cells, a pCDF vector containing the gene coding for SerB and EF-Sep was also transformed and indicated as Δ*serC*+SerB. Recombinant sfGFP(150TAG)-His6 was expressed, purified by Ni-NTA resin, analyzed by SDS-PAGE, and visualized by InstantBlue staining.



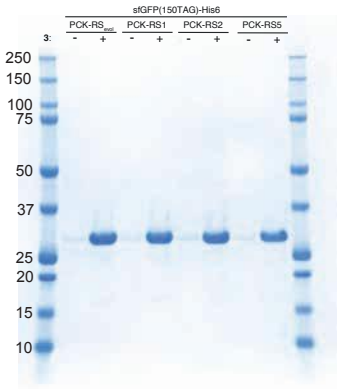
Supplementary Figure 3. Phosphothreonine standards were analyzed by ESI-MS. a, pThr is dissolved in H_2O to final concentration of 5μM (blue), 10μM (brown), 20μM (magenta), 50μM (red), or 70μM (black), and was subjected to ESI-MS analysis. Intensity representing ion counts in SIM negative mode (m/z=198) was plotted against elution time. **b,** Trace curves from **a** were integrated and the peak areas were calculated using Agilent software. The peak areas were then plotted against the concentration of pThr, and a linear regression was constructed using the software Prism. The data show the mean of three cell cultures and the error bars represent the standard deviation.



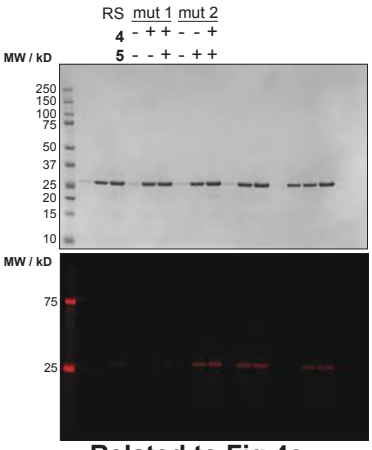
Related to Fig 2a



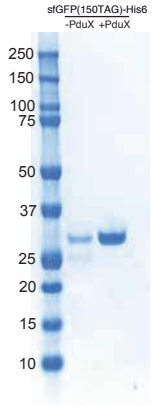
Related to Fig 2e



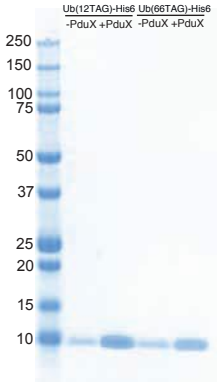
Related to Fig 3e



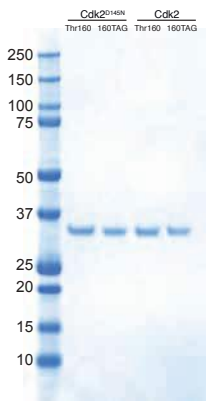
Related to Fig 4c



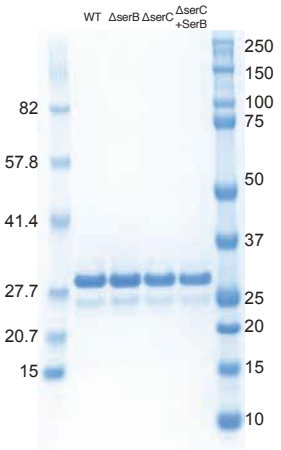
Related to Fig 5d



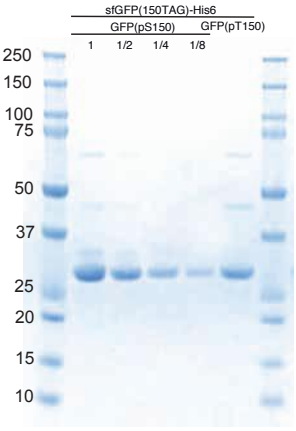
Related to Fig 5e



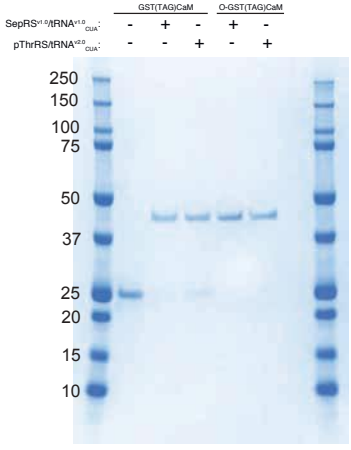
Related to Fig 5g



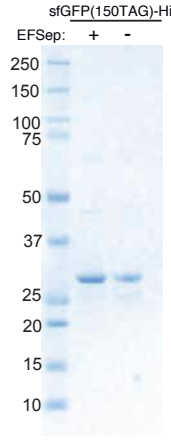
Related to Fig S2



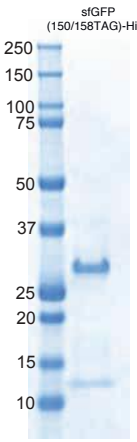
Related to Fig S15a



Related to Fig S15d

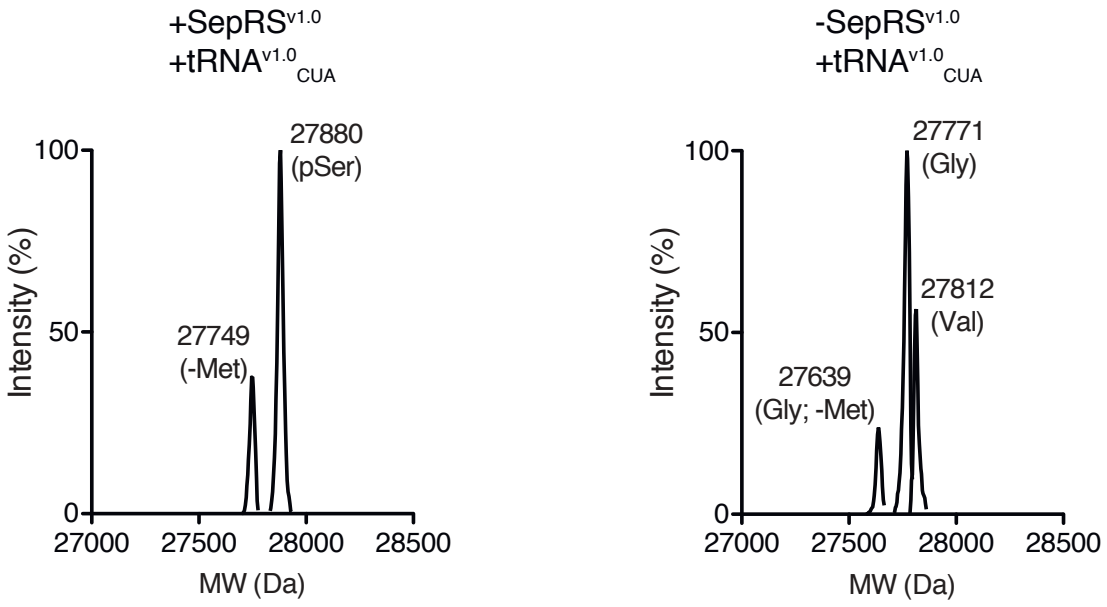


Related to Fig S16



Related to Fig S18

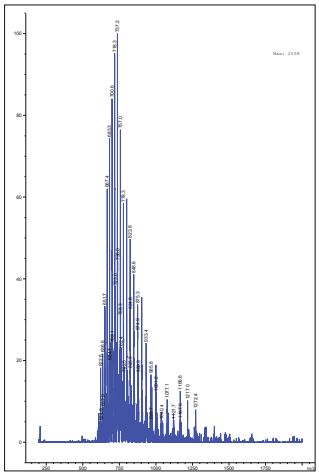
Supplementary Figure 4. Full-length gel images corresponding to cropped images in this study.



Supplementary Figure 5. In the absence of SepRS$^{V1.0}$, tRNA$^{V1.0}_{CUA}$ can be acylated by Glycyl-tRNA synthetase and Valyl-tRNA synthetase in *E.coli*. DH10B cells bearing sfGFP(150TAG)-His6 were transformed with an accessory plasmid configured as ±SepRS$^{V1.0}$/+tRNA$^{V1.0}_{CUA}$. Recombinant sfGFP(150TAG)-His6 was expressed, purified by Ni-NTA resin, and analyzed by ESI-MS. The observed mass is indicated and used to identify amino acid incorporated at a.a. position 150.

Fig.4d

Mut RS 1



Mut RS 2

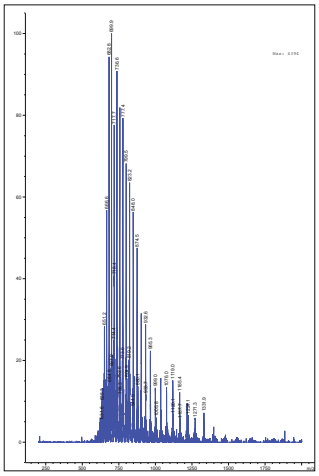


Fig.5d

sfGFP(pThr150)-His6

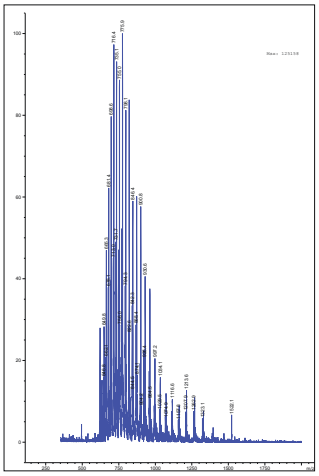
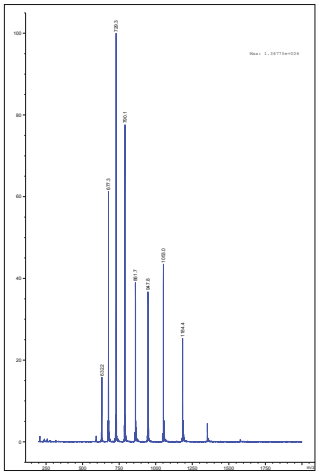


Fig.5e

Ub(pThr12)-His6



Ub(pThr66)-His6

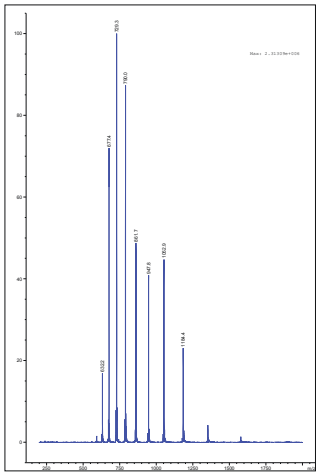
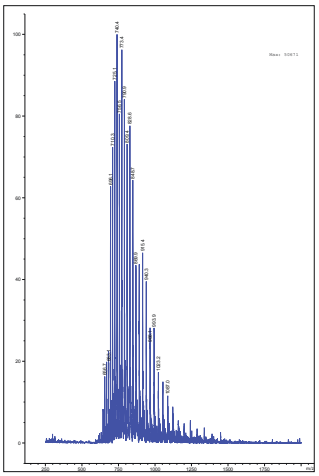


Fig.5g

Cdk2D145N(Thr160)-His6



Cdk2D145N(Thr160)-His6

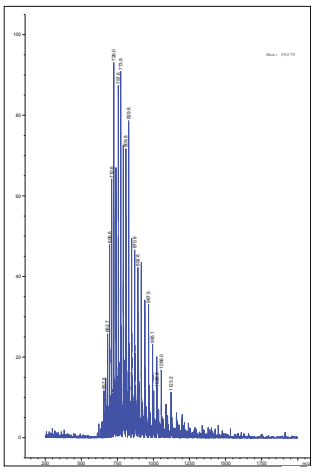
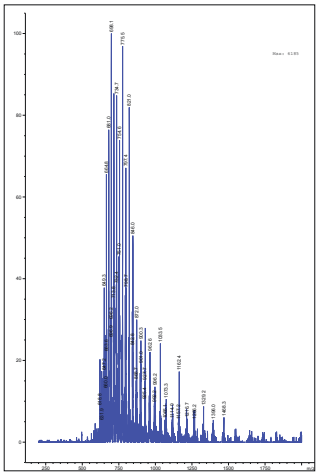


Fig.S5

+SepRS$^{v1.0}$

+tRNA$^{v1.0}_{CUA}$



-SepRS$^{v1.0}$

+tRNA$^{v1.0}_{CUA}$

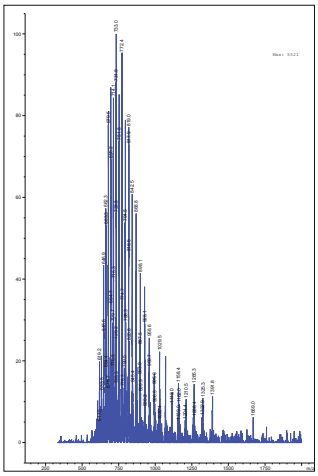
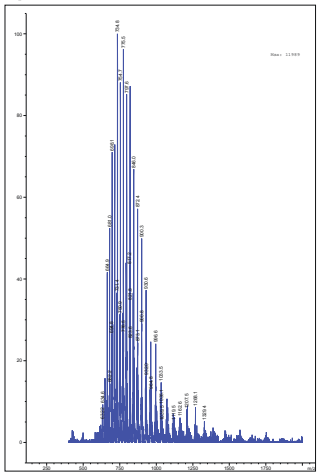
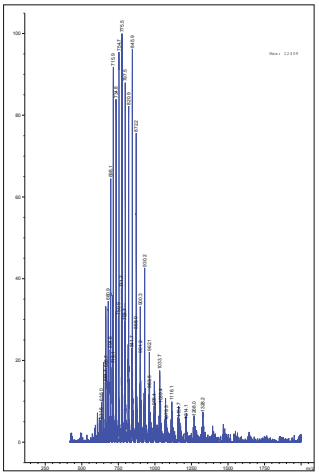


Fig.S8

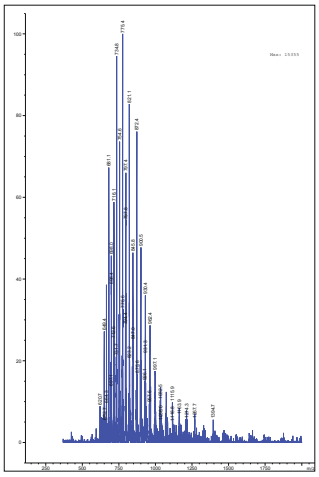
tRNA hit2



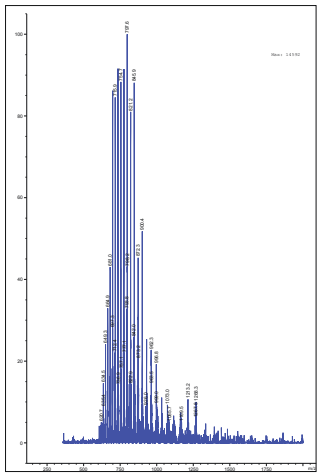
tRNA hit5



tRNA hit8



tRNA hit13



tRNA hit15

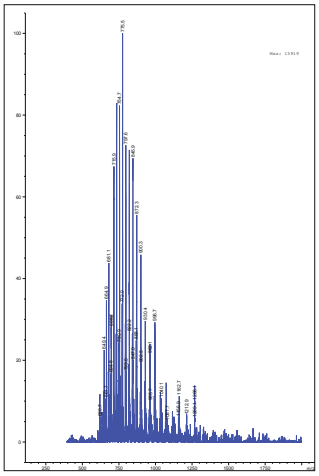
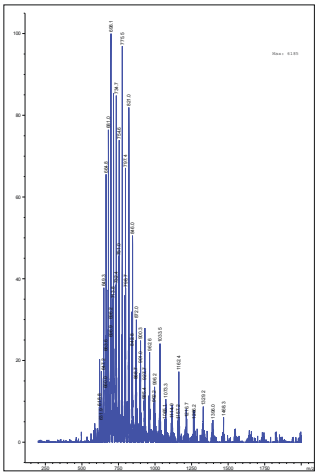


Fig.S15c

sfGFP(pSer150)-His6



sfGFP(pThr150)-His6

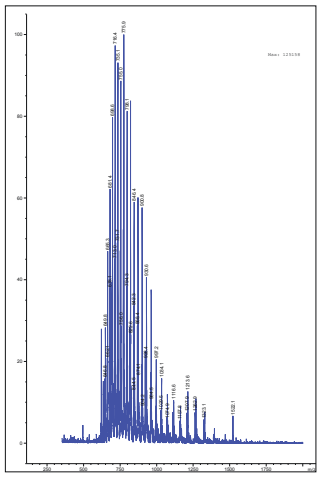


Fig.S15e

-PduX

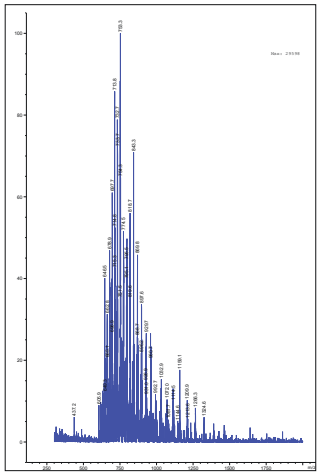
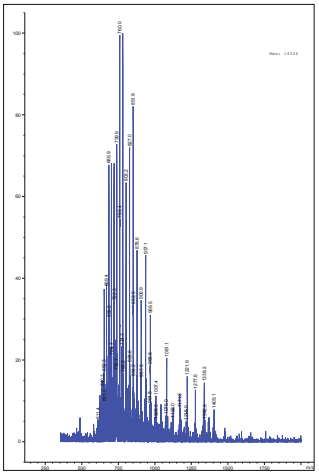
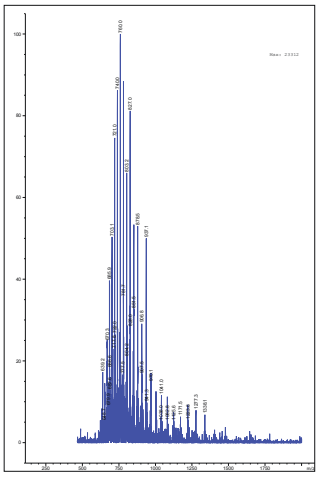


Fig.S11

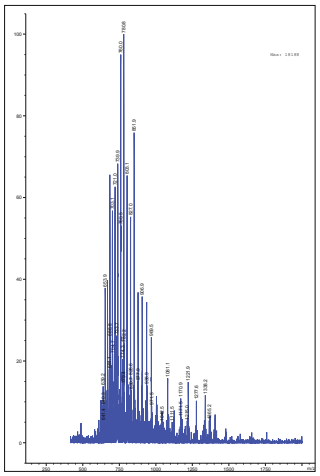
PCK-RS(evol)



PCK-RS1



PCK-RS2



PCK-RS5

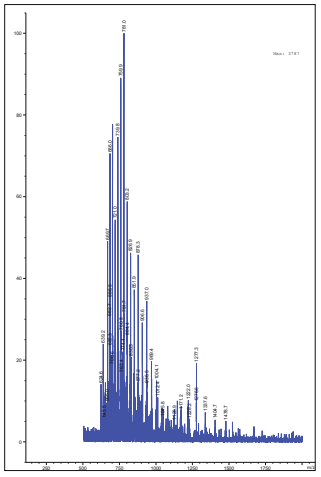


Fig.S18
sfGFP(150/158TAG)-His6

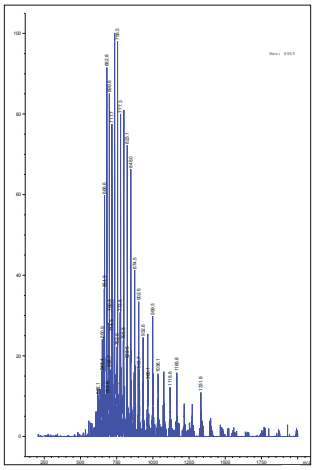
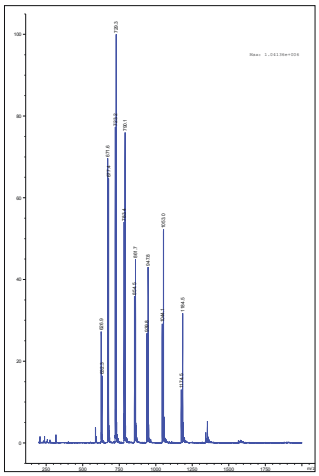
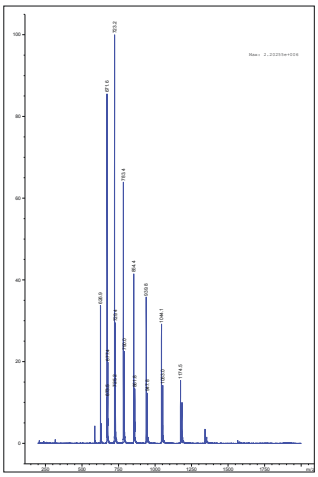


Fig.S19 (Row1)
Ub(12TAG)-His6



Ub(66TAG)-His6



Cdk2D145N(160TAG)-His6

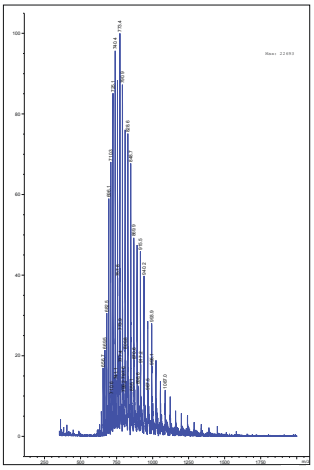
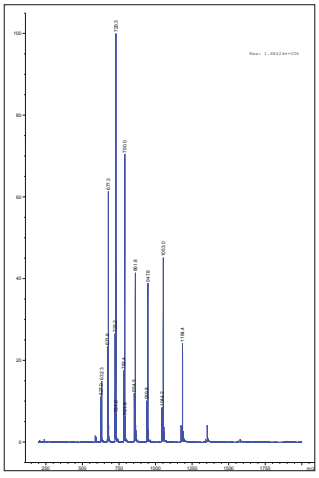
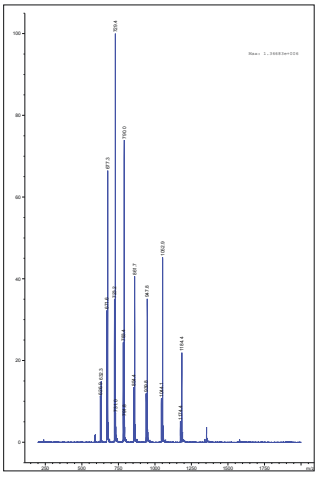


Fig.S20

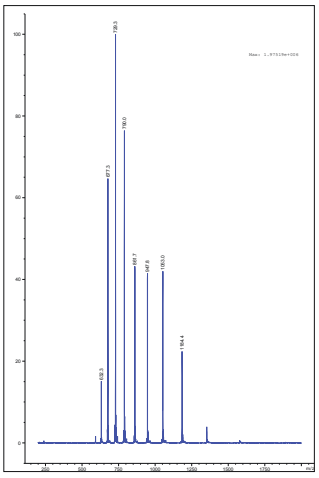
ΔaphA



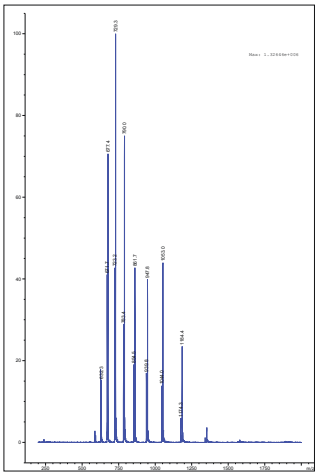
ΔphoA



ΔycdX



ΔpgpA



ΔpgpC

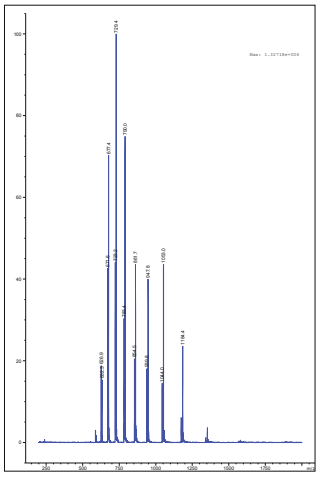
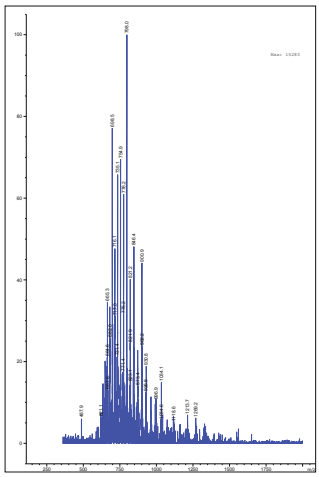
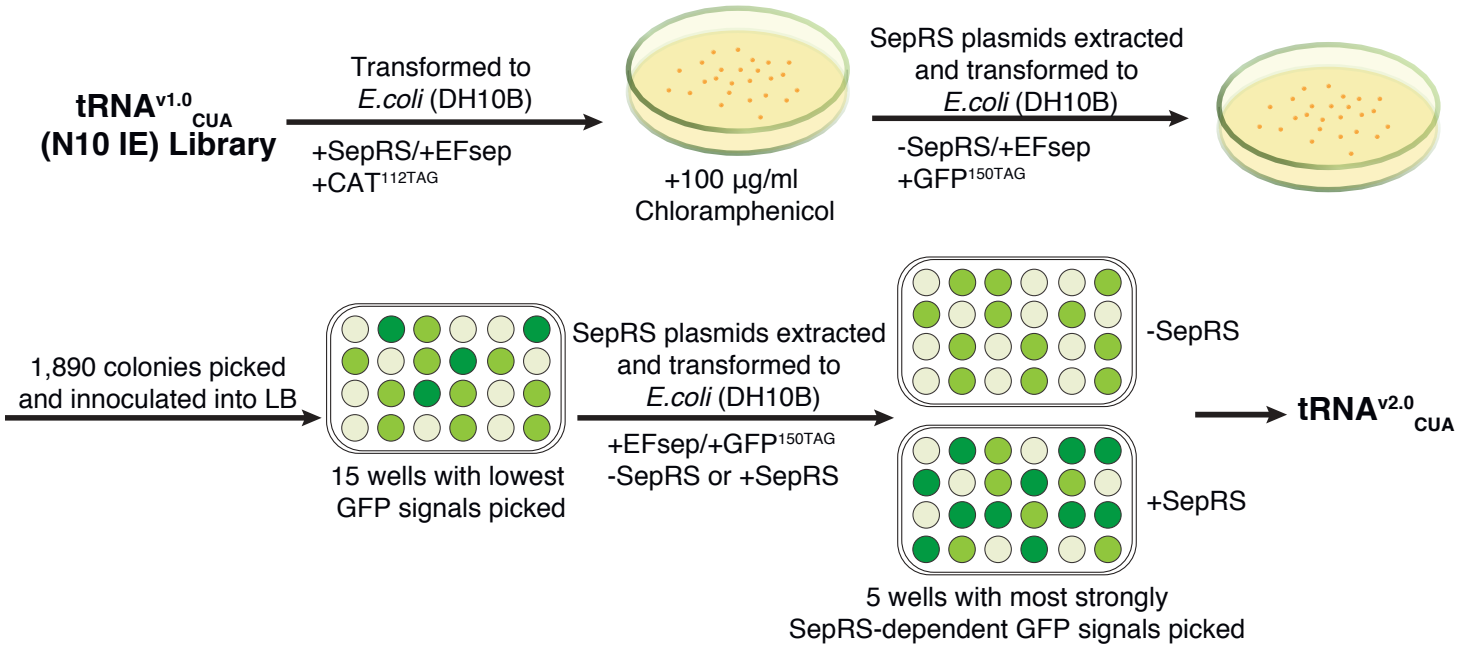


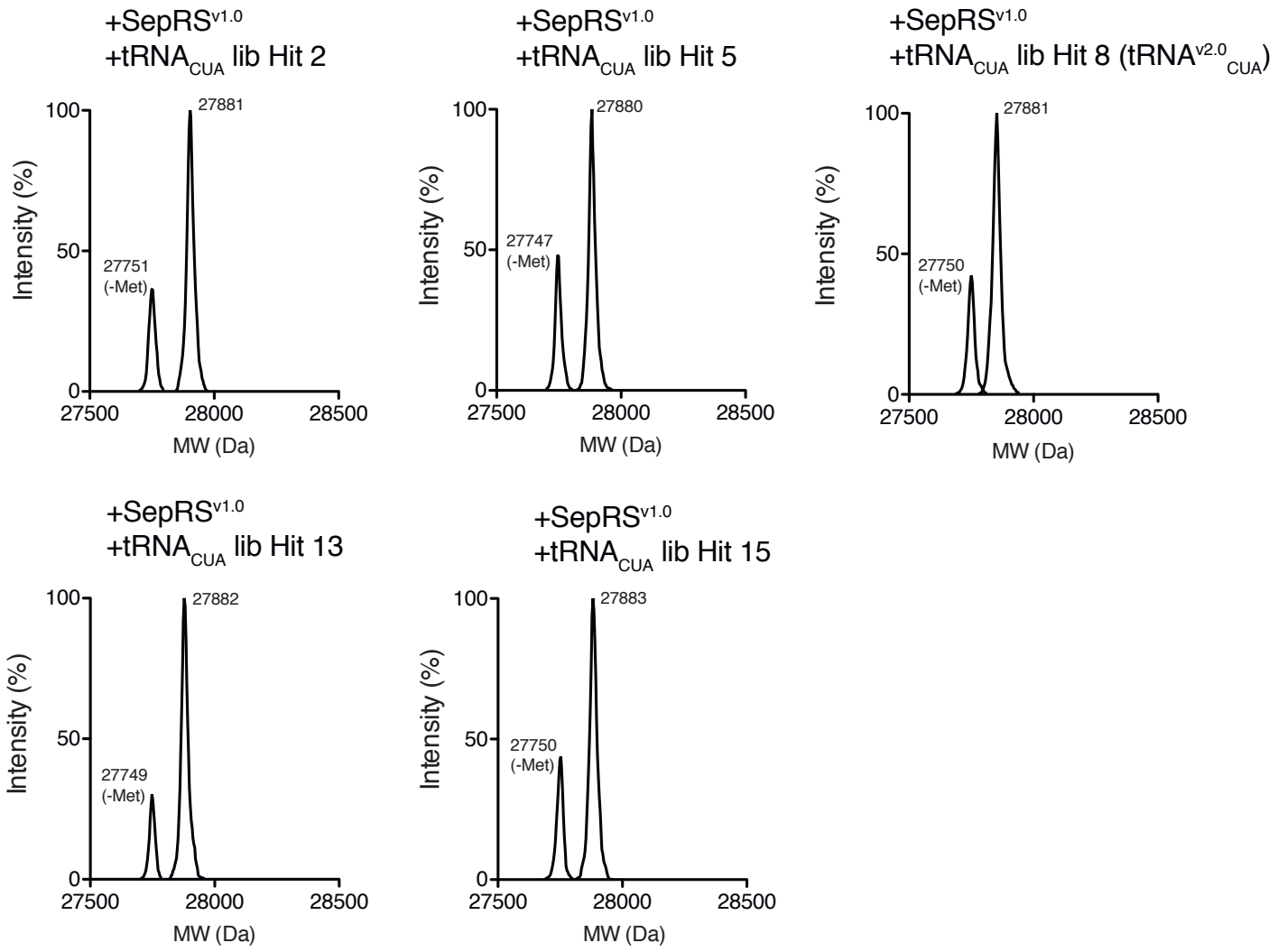
Fig.S16
-EF-Sep



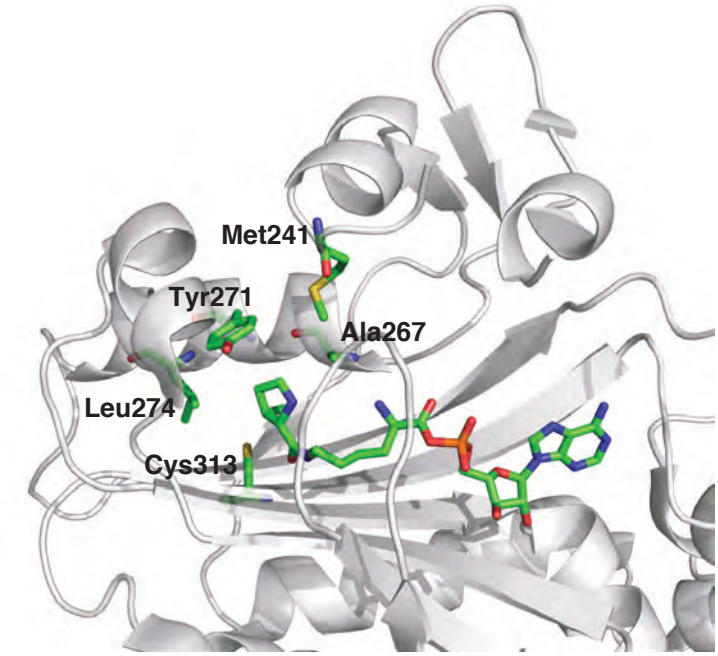
Supplementary Figure 6. Raw MS data corresponding to ESI-MS analysis results in this study.



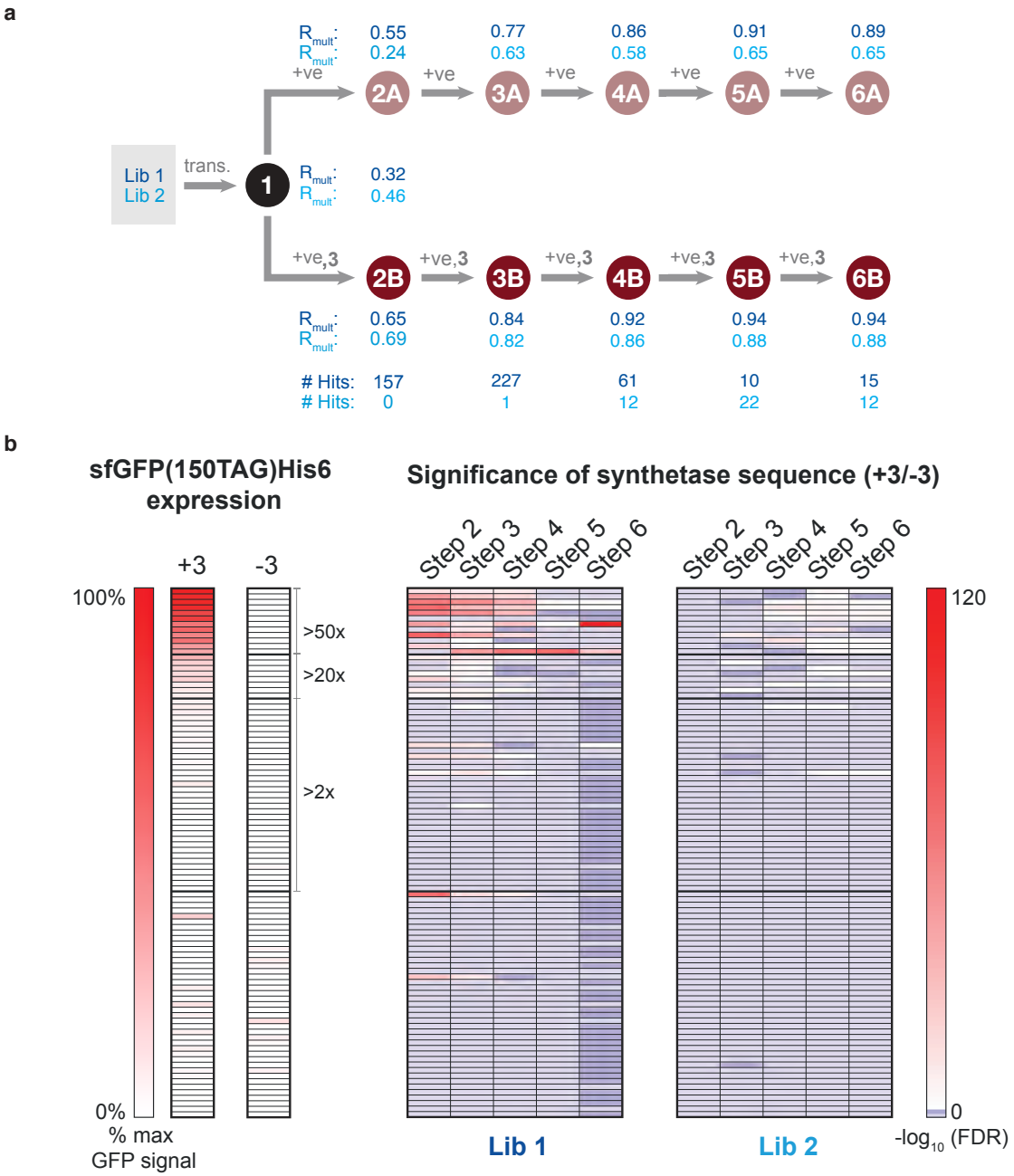
Supplementary Figure 7. Schematic process to select the tRNA$^{V1.0}_{CUA}$ library (N10 IE). tRNA$^{V1.0}_{CUA}$ library was co-transformed into DH10B cells with SepRS$^{V1.0}$ as indicated. tRNA$^{V1.0}_{CUA}$ library plasmids were extracted from cells that survived the CAT assay and were re-transformed into fresh DH10B cells without SepRS$^{V1.0}$. 1,890 colonies were picked and inoculated into LB medium for GFP assay. tRNA$^{V1.0}_{CUA}$ library plasmids from 15 wells with the lowest GFP signals were extracted and transformed again into fresh DH10B cells with or without SepRS$^{V1.0}$ for GFP assay. Five clones with the strongest SepRS$^{V1.0}$-dependent GFP signals were isolated and the tRNA plasmid from each clone was extracted and sequenced.



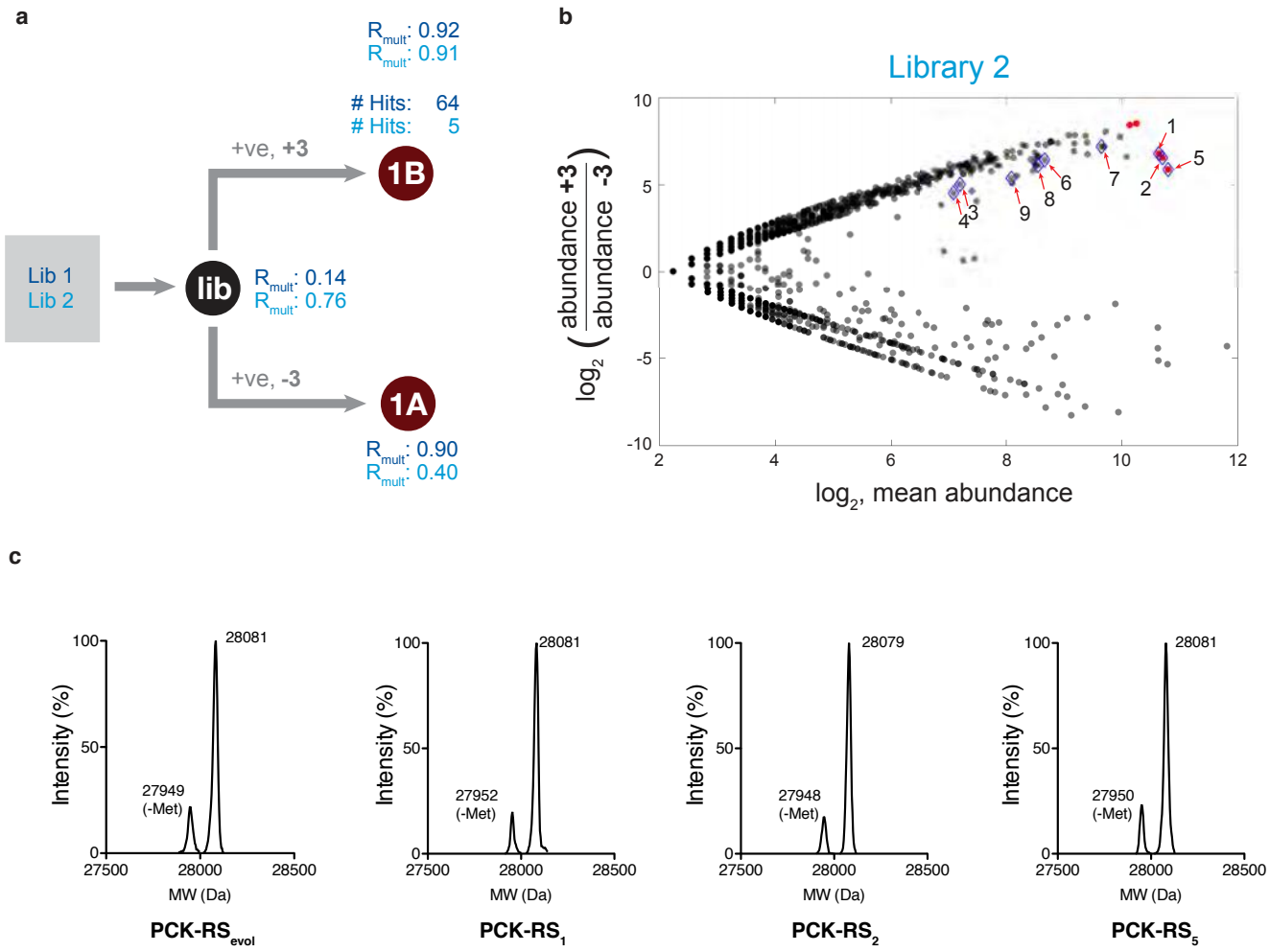
Supplementary Figure 8. Evolved SeptRNA clones incorporate phosphoserine into a.a. position 150 of sfGFP(150TAG)-His6 in the presence of SepRS$^{V1.0}$. Each of the five evolved SeptRNA clones (Hit2, 5, 8, 13&15) was transformed into DH10B cells with SepRS$^{V1.0}$ and sfGFP(150TAG)-His6. Recombinant sfGFP(150TAG)-His6 was expressed, purified by Ni-NTA resin, and analyzed by ESI-MS. The observed mass of peak is indicated and used to identify amino acid incorporated at a.a. position 150.



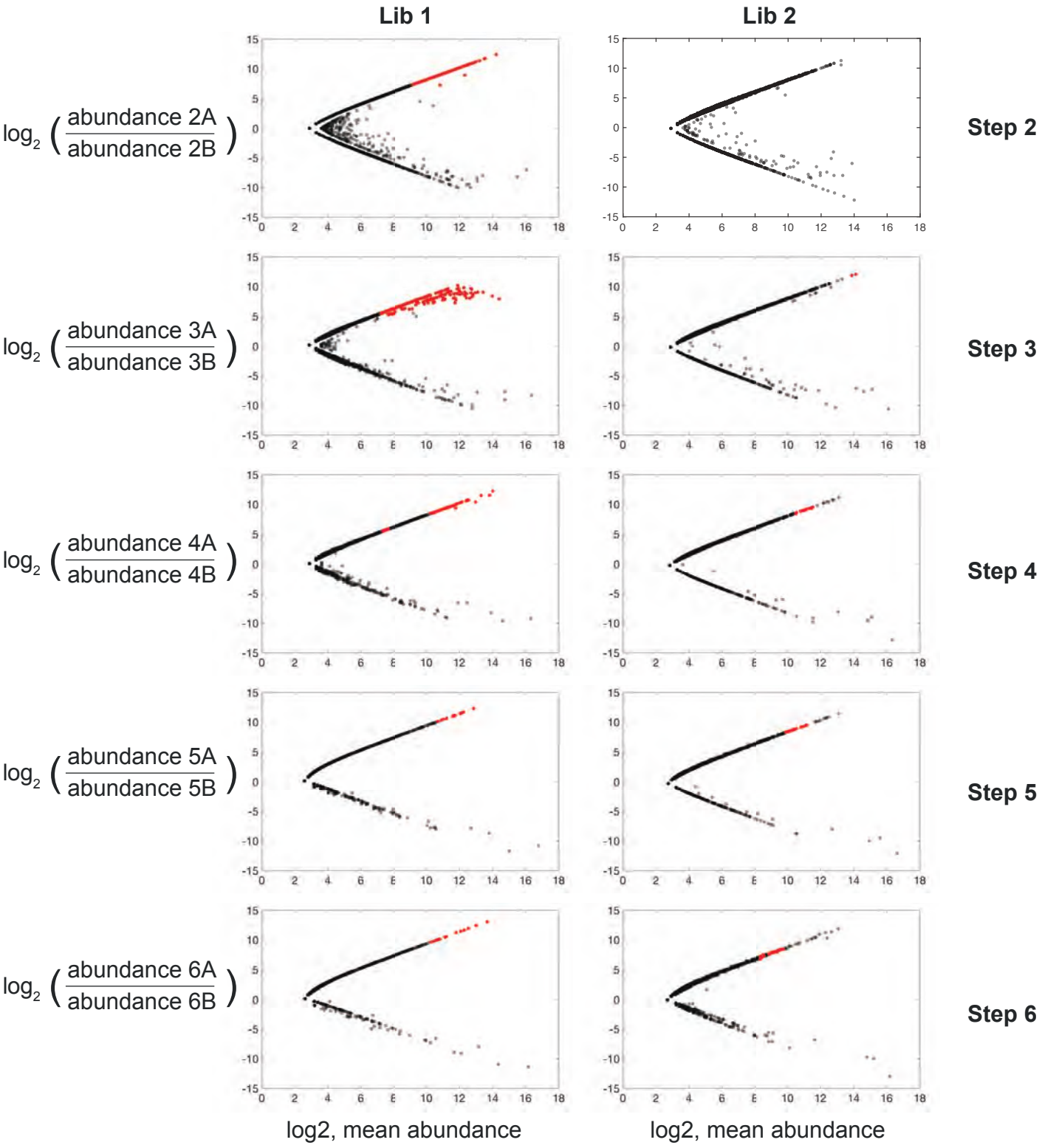
Supplementary Figure 9. Amino acid residues randomized in PylRS libraries. Structure of amino acid binding pocket of *M.mazei* PylRS (PDB ID: 2Q7H). Amino acid residues chosen for randomization in PylRS library are labeled as Met276, Ala302, Tyr306, Leu309, and Cys348 (corresponding to Met241, Ala267, Tyr271, Leu274, and Cys313 in *M. barkeri* PylRS).



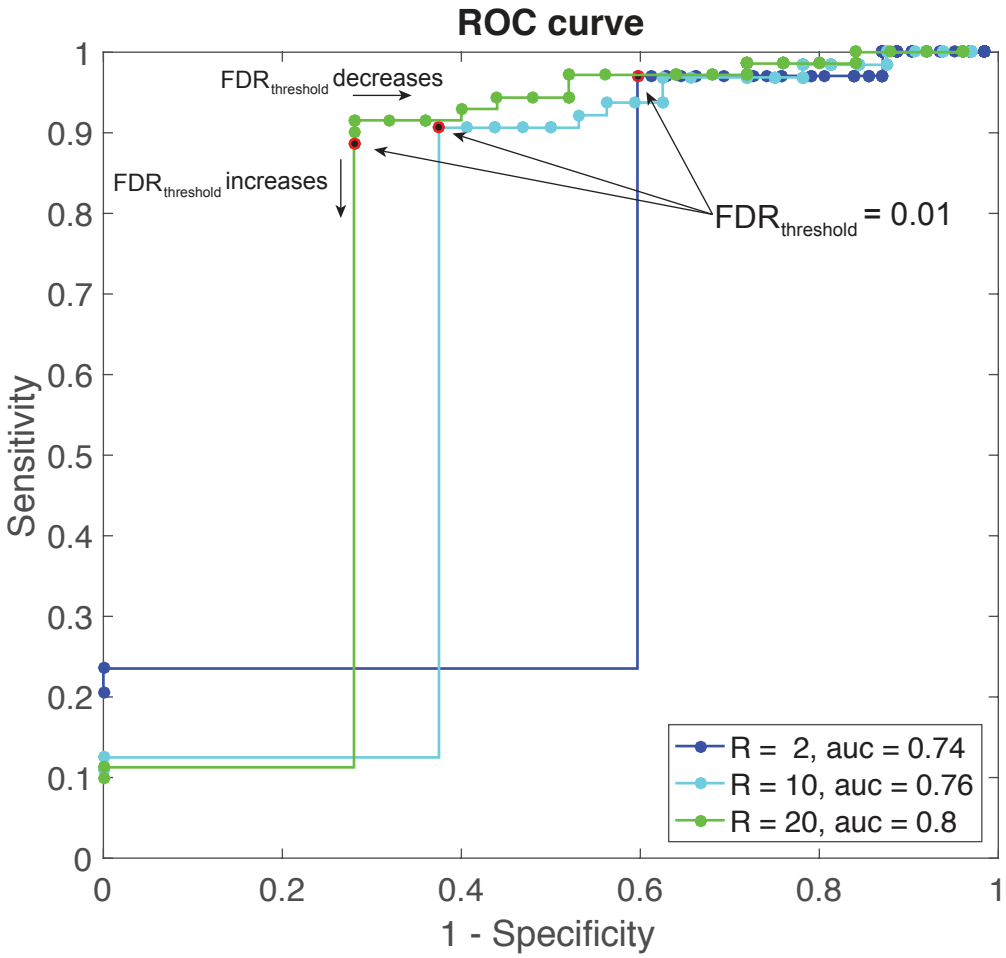
Supplementary Figure 10. Parallel positive selections coupled to deep sequencing and analysis identifies aminoacyl-tRNA synthetases/tRNA$_{CUA}$ pairs that incorporate 4. a, Schematic process of the identification of PylRS variants that direct the incorporation of **4**. *pylRS* libraries that randomize four (Lib1) or five (Lib2) positions in the active site of PylRS were subjected to indicated sequence of positive (+ve) selections, in presence or absence of **4**. Each selection was performed in three biological replicates. The resulting synthetase gene pools were deep sequenced. Correlations of sequence abundances in the three biological replicates are represented by R_{mult}. For each selection step the number of significantly (FDR < 0.01), differentially enriched mutants in presence versus absence of **4** was identified by DESeq. The number of sequences identified as hits is provided. **b,** DESeq analysis of the selection shown in **a** identifies PylRS/tRNA$_{CUA}$ pairs that are active and specific for **4**. From selection step 4B (Lib 2) 96 evolved PylRS/tRNA$_{CUA}$ pairs were picked and used to drive the expression of sfGFP(150TAG)-His6 in presence and absence of **4**. The sfGFP phenotype and statistical significance [-log$_{10}$ (FDR)] of sequencing enrichment are represented as heatmaps. Each row represents one mutant and the rows are arranged by descending ratio of GFP fluorescence in presence versus absence of **4**. The GFP signals show the mean of three biological replicates.



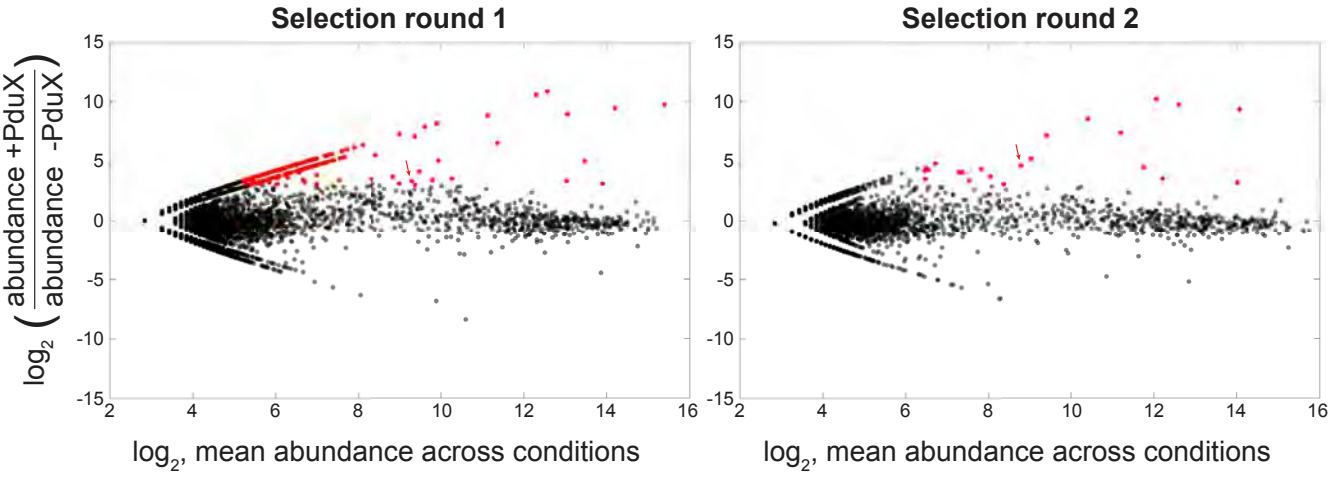
Supplementary Figure 11. Parallel positive selections coupled to deep sequencing and analysis identifies aminoacyl-tRNA synthetases/tRNA$_{CUA}$ pairs that incorporate 3. a, Schematic process of the identification of PylRS variants that direct the incorporation of **3** (same as **Fig.3b**). **b,** Deep sequencing data of five-member library (Lib 2). The X-axis represents the $\log_2$-transformed mean mutant abundance in presence and absence of **3**, the Y-axis represents the $\log_2$-transformed ratio of mutant abundance in presence versus absence of **3**. Red dots represent mutants that are significantly enriched (FDR < 0.01) in presence versus absence of **3**. Blue dimonds represent mutants that are significantly enriched (FDR < 0.01) in presence versus absence of **3** in four-member library (Lib 1) (see **Fig.3c**). Hit1, 2, and 5 are mutants that are significantly enriched (FDR < 0.01) in presence versus absence of **3** in both libraries (Lib 1 and Lib 2). **c,** Recombinant sfGFP(150TAG)-His6 expressed and purified in **Fig.3e** was analyzed by ESI-MS. The observed mass of peak is indicated and used to identify amino acid incorporated at a.a. position 150.



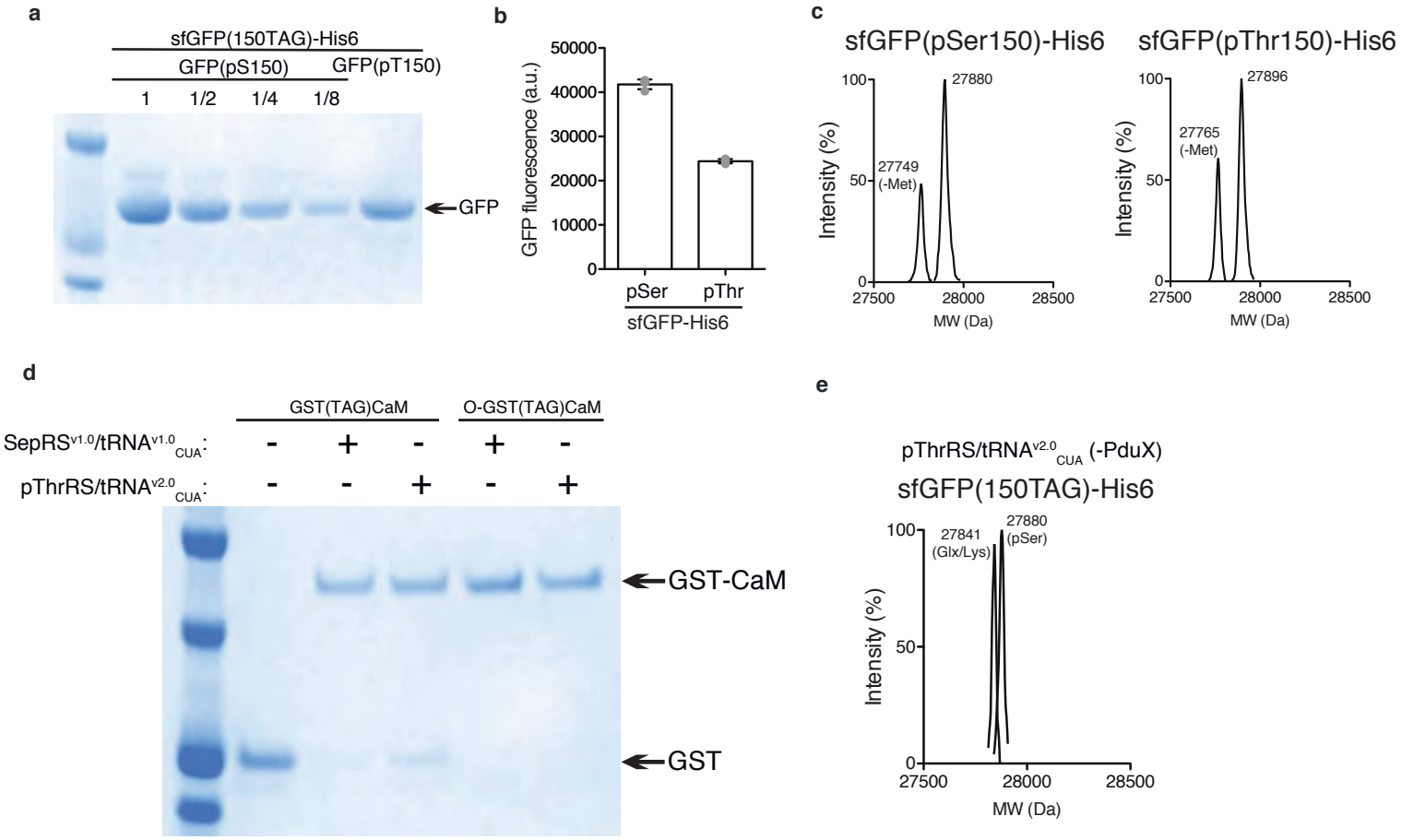
Supplementary Figure 12. Representation of the deep sequencing data from each round of *pylRS* library selection represented in Figure S10. The X-axis represents the $\log_2$-transformed mean mutant abundance in presence and absence of **4**, the Y-axis represents the $\log_2$-transformed ratio of mutant abundance in presence versus absence of **4**. Red dots represent mutants that are significantly enriched (FDR < 0.01) in presence versus absence of **4**.



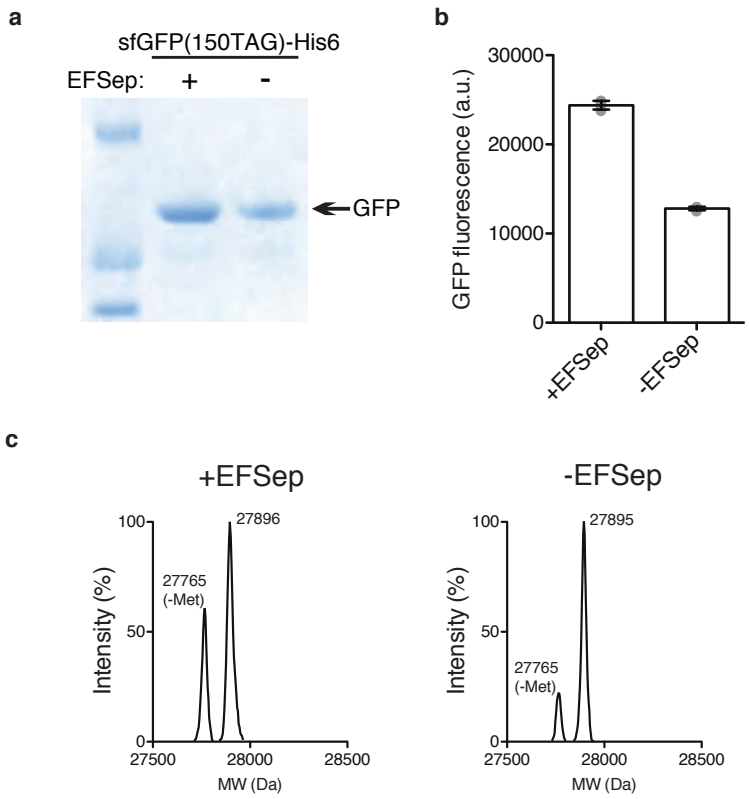
Supplementary Figure 13. Receiver Operating Characteristic (ROC) of CbzK-RS identification. Specificity and sensitivity were determined for the 96 mutants represented in Figure S10b. The ratio R of sfGFP fluorescence in presence versus absence of **4** was taken as the true activity of a mutant. Significance of differential enrichment of deep sequencing read counts after selection in presence versus absence of **4** was considered as the predicted activity. The FDR at which a mutant was predicted to be active was represented as $FDR_{threshold}$. The threshold was varied over a set of values, such that $FDR_{threshold}$ is an element of $\{10^{1}, 10^{-1}, \ldots, 10^{-100}\}$. The ROC was evaluated for three different threshold of R: 2, 10, 20. The red datapoint corresponds to $FDR_{threshold} = 10^{-2}$.



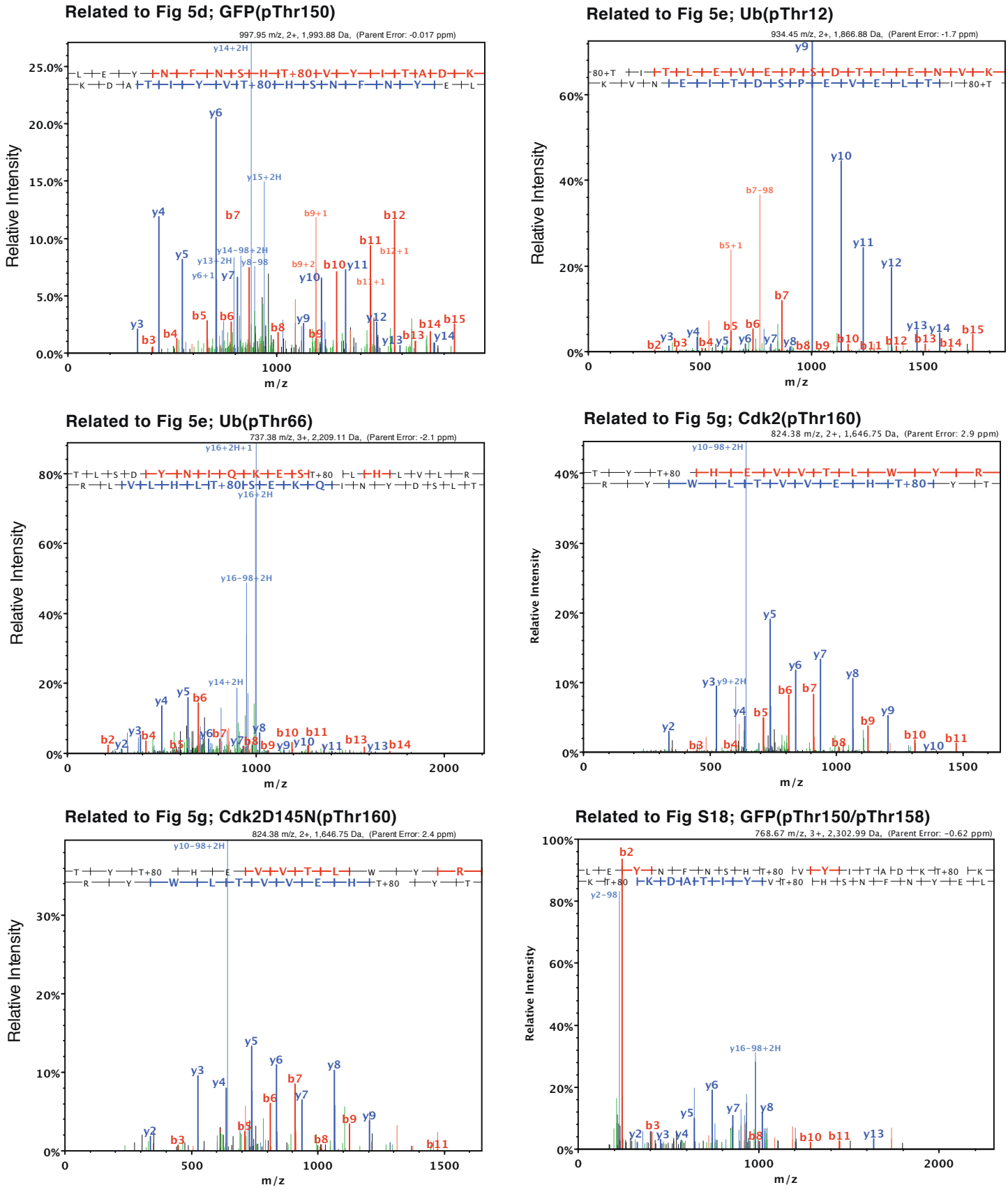
Supplementary Figure 14. Deep sequencing data from SepRS library selection represented in Figure 5a. The X-axis represents the $\log_2$-transformed mean mutant abundance in presence and absence of PduX, the Y-axis represents the $\log_2$-transformed ratio of mutant abundance in presence versus absence of PduX. Red dots represent mutants that are significantly enriched (FDR < 0.01) in presence versus absence of PduX.



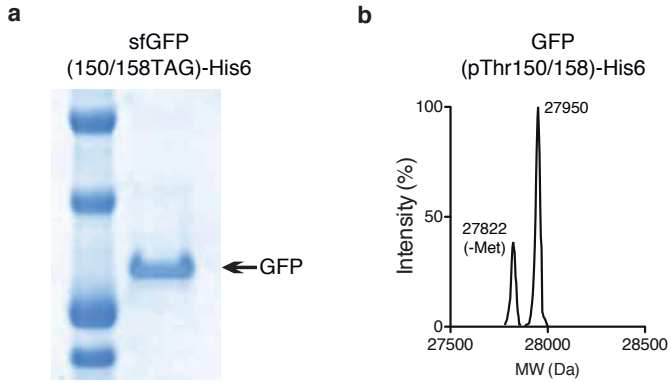
Supplementary Figure 15. The pThrRS/tRNA$^{v2.0}_{CUA}$ pair identified incorporates pThr at a.a. position 150 of sfGFP-His6 with a comparable efficiency to the SepRS$^{v1.0}$/tRNA$^{v2.0}_{CUA}$ pair's efficiency in pSer incorporation. a, sfGFP(150TAG)-His6 was co-transformed into BL21 Δ*serB* cells with a vector containing SepRS$^{v1.0}$/tRNA$_{CUA}$ pair and EF-Sep. sfGFP(150TAG)-His6 was co-transformed into BL21 Δ*serC* cells with a vector containing pThrRS/tRNA$_{CUA}$ pair, PduX, and EF-Sep. Recombinant sfGFP(150TAG)-His6 was expressed, purified by Ni-NTA resin, analyzed by SDS-PAGE, and visualized by InstantBlue staining. A serial dilution of sfGFP proteins produced with the SepRS$^{v1.0}$/tRNA$^{v2.0}_{CUA}$ pair was loaded to compare the sfGFP produced using pThrRS/tRNA$^{v2.0}_{CUA}$ pair. **b,** Fluorescent signals of recombinant sfGFP(150TAG)-His6, incorporating pThr and pSer expressed in **a,** were analyzed by Tecan plate reader. The data show the mean fluorescent intensity of GFP proteins purified from four independent expressions and the error bars represent the standard deviation. Gray dots show individual data points of the four replicates **c,** Recombinant sfGFP(150TAG)-His6 expressed and purified as described in **a** was analyzed by ESI-MS. The observed mass corresponds to incorporation of the indicated amino acid. **d,** pRSF vector containing GST(TAG)CaM or riboQ1 rRNA with O-GST(TAG)CaM was co-transformed with a vector containing SepRS/tRNA$^{v2.0}_{CUA}$ pair and EF-Sep for pSer incorporation and with a vector containing pThrRS/tRNA$_{CUA}$ pair, PduX, and EF-Sep for pThr incorporation. GST-CaM proteins were expressed, purified by glutathione resin, analyzed by SDS-PAGE, and visualized by InstantBlue staining. **e.** sfGFP(150TAG)-His6 was co-transformed into BL21 Δ*serC* cells with a vector containing pThrRS/tRNA$_{CUA}$ pair and EF-Sep in the absence of PduX. Recombinant sfGFP(150TAG)-His6 expressed and purified as described in **Fig.5d** (lane 1) was analyzed by ESI-MS. The observed mass corresponds to incorporation of the indicated amino acid.



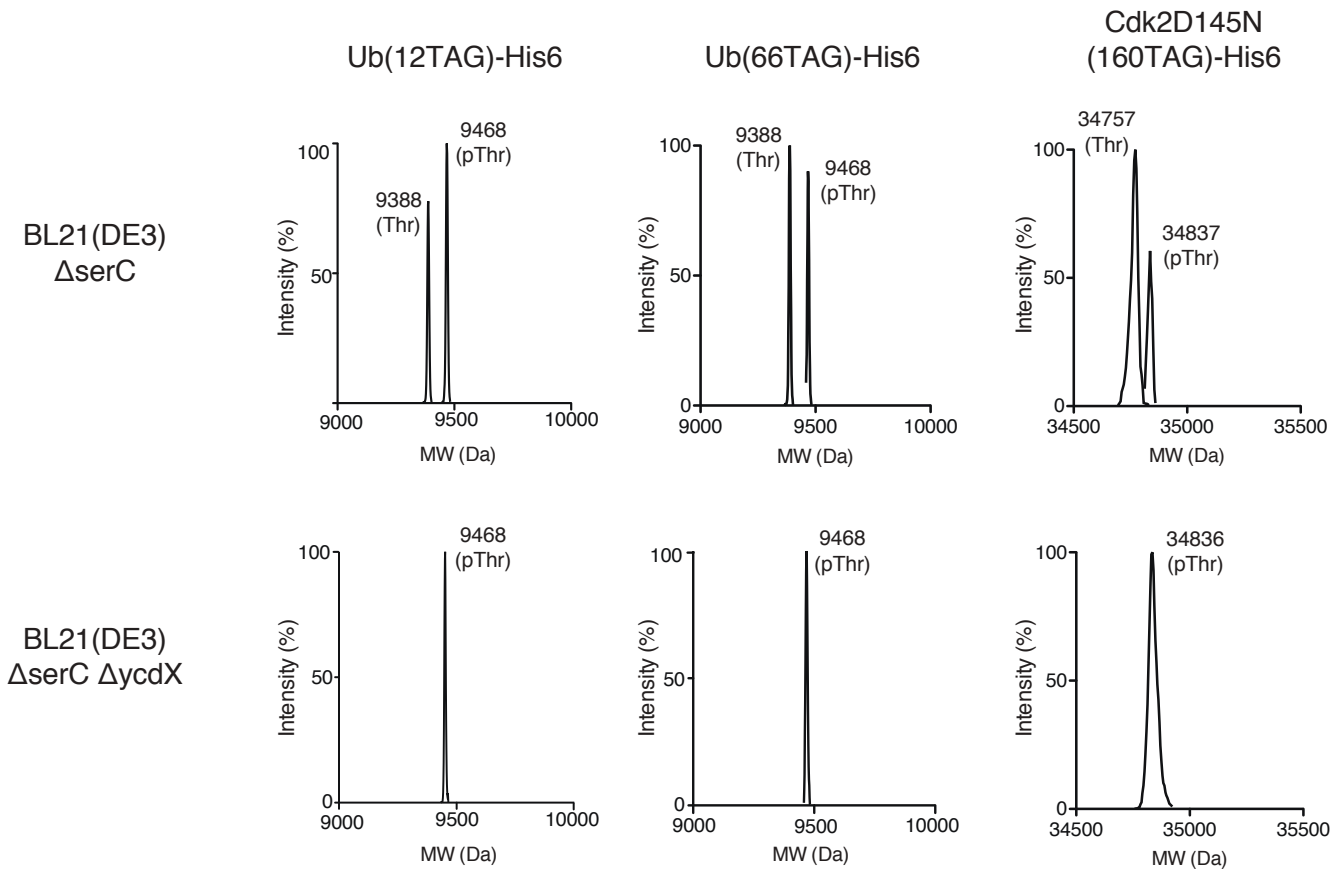
Supplementary Figure 16. The pThr incorporation efficiency of pThrRS/tRNA$_{CUA}$ pair weakly depends on the presence of EF-Sep. a, sfGFP(150TAG)-His6 was co-transformed into BL21 Δ*serC* cells with a vector containing pThrRS/tRNA$^{v2.0}_{CUA}$ pair, PduX, and EF-Sep (+EF-Sep) or a vector containing pThrRS/tRNA$^{v2.0}_{CUA}$ pair and PduX (-EF-Sep). Recombinant sfGFP(150TAG)-His6 was expressed, purified by Ni-NTA resin, analyzed by SDS-PAGE, and visualized by InstantBlue staining. **b,** Fluorescent signals of recombinant sfGFP(150TAG)-His6 expressed in **a** were analyzed by Tecan plate reader. The data show the mean fluorescent intensity of GFP proteins purified from four independent expressions and the error bars represent the standard deviation. Gray dots show individual data points of the four replicates. **c,** Recombinant sfGFP(150TAG)-His6 expressed and purified in **a** was analyzed by ESI-MS. The observed mass of peak is indicated.



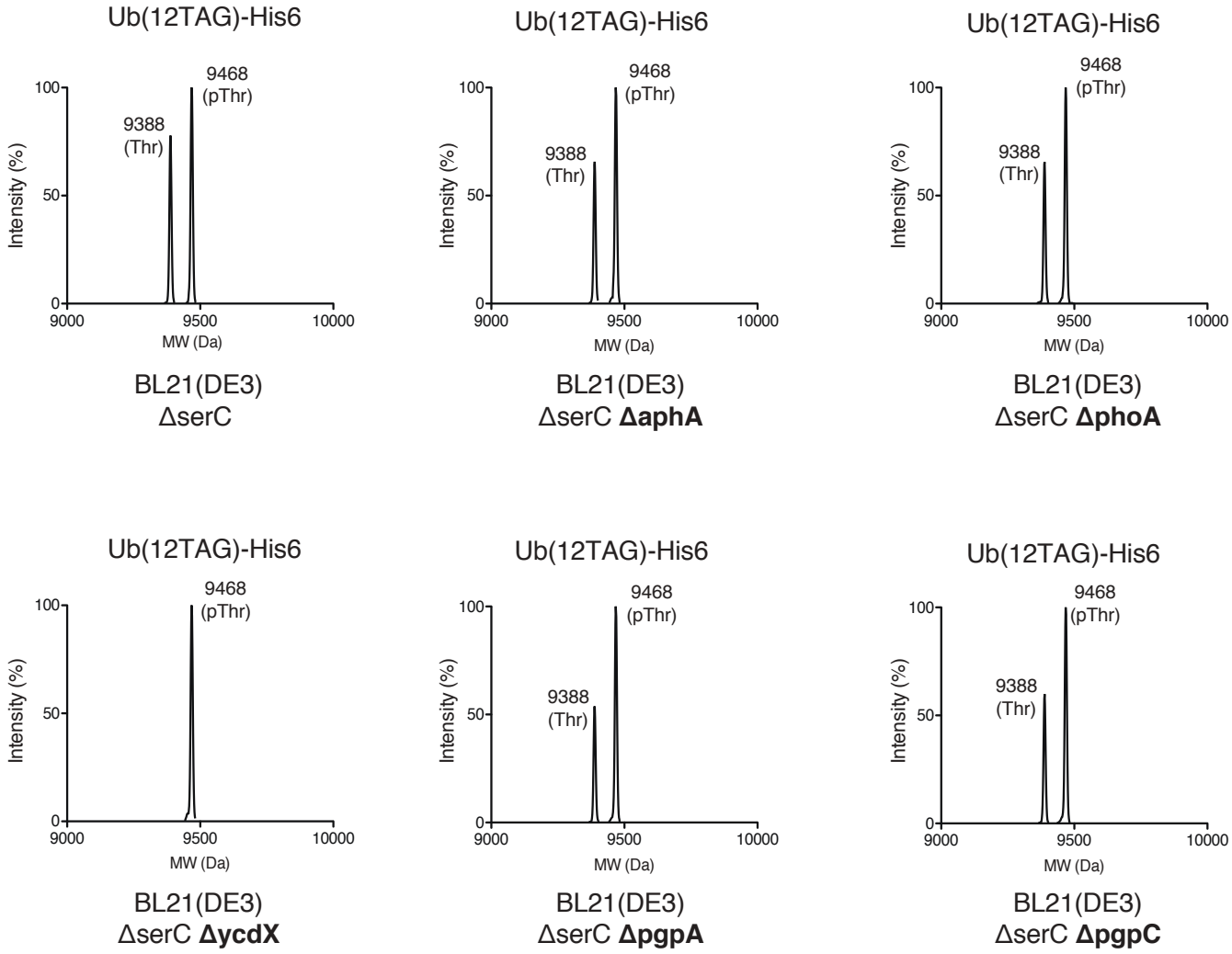
Supplementary Figure 17. ESI-MS/MS analysis data in this study. Parental peptide sequences are listed from N-terminus to C-terminus in red and from C-terminus to N-terminus in blue. Ionized fragments from parental peptide are identified and indicated (b-ions in red and y-ions in blue). Y-axis shows the intensity of each ionized fragment relative to the parental ion (not shown). The phosphorylated threonine residues are labeled as **T+80**.



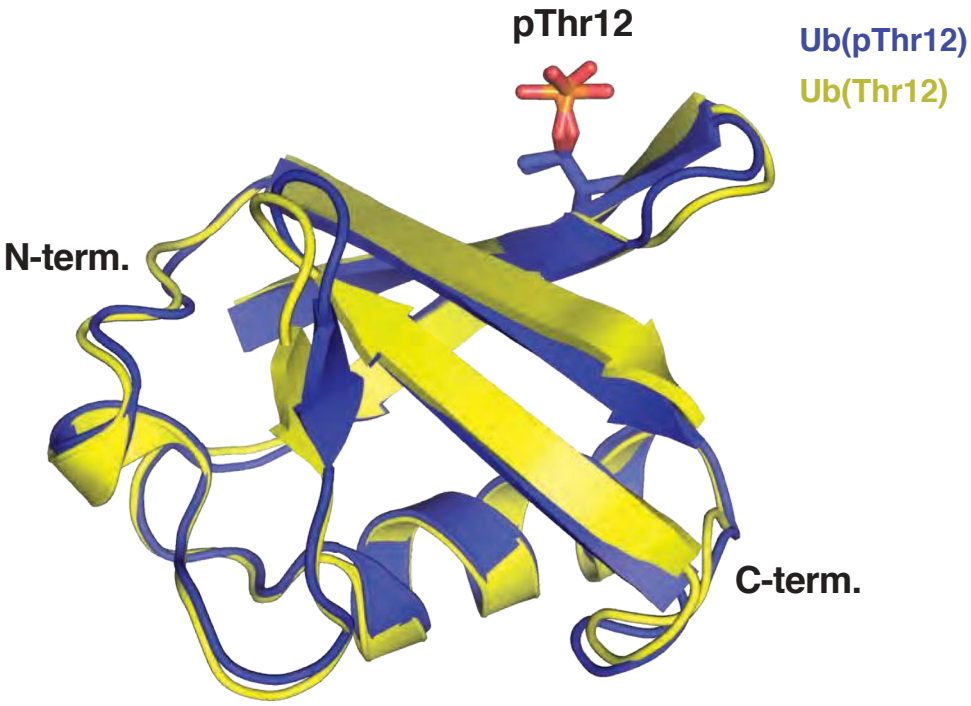
Supplementary Figure 18. The pThrRS/tRNA$_{CUA}$ can incorporate two pThr residues at a.a. positions 150 and 158 of sfGFP$^{150TAG/158TAG}$-His6. a, sfGFP(150TAG/158TAG)-His6 was co-transformed into BL21 Δ*serC* cells with a vector containing pThrRS/tRNA$^{v2.0}_{CUA}$ pair, PduX, and EF-Sep. Recombinant sfGFP(150TAG/158TAG)-His6 was expressed, purified by Ni-NTA resin, analyzed by SDS-PAGE, and visualized by InstantBlue staining. **b,** Recombinant sfGFP(150TAG/158TAG)-His6 expressed and purified in **a** was analyzed by ESI-MS. The observed mass of peak is indicated.



Supplementary Figure 19. pThr incorporation in proteins expressed in BL21 Δ*serC* cells before and after knocking out phosphatase *ycdX*. pNHD vectors containing Ub(12TAG)-His6, Ub(66TAG)-His6, or Cdk2(160TAG)-His6 were co-transformed in BL21 Δ*serC* cells or BL21 Δ*serC* Δ*ycdX* cells. Recombinant proteins were expressed, purified, and analyzed by ESI-MS. The observed mass of peak is indicated and used to identify amino acid incorporated at a.a. position 12 or 66 for Ub and a.a. position 160 for Cdk2.



Supplementary Figure 20. Screening phosphatase deletions identifies YcdX as the phosphatase responsible for pThr dephosphorylation. Each of the candidate phosphatases was individually knocked out from BL21 Δ*serC* cells. Each strain was then co-transformed with a pNHD vector containing Ub(12TAG)-His6 and a vector containing pThrRS/tRNA$^{v2.0}_{CUA}$ pair, PduX, and EF-Sep. Recombinant Ub(12TAG)-His6 was expressed, purified by Ni-NTA resin, and analyzed by ESI-MS. The observed mass of peak is indicated. The observed mass of peak is indicated and used to identify amino acid incorporated at a.a. position 12.



Supplementary Figure 21. Structure of Ub (pThr12) aligns well with the structure of a Ub reference. Alignment of the structure of Ub (pThr12) (blue) and a Ub reference structure (PDB: 1UBQ) (yellow), RMSD for residues 1–72 = 0.595 Å. The phosphorus atom is displayed in orange and oxygen atoms are in red.

	Amino acid mutations					Codon mutations				
Amino acid position	**241**	**267**	**271**	**274**	**313**	**241**	**267**	**271**	**274**	**313**
PylRS-WT	**M**	**A**	**Y**	**L**	**C**	**ATG**	**GCC**	**TAT**	**CTG**	**TGC**
Related to Figure 3 and Figure S11										
PCK-RSevol	F	S	C	M	C	TTT	AGT	TGT	ATG	TGT
PCK-RS1	M	A	A	A	A	ATG	GCT	GCG	GCG	GCG
PCK-RS2	M	S	T	M	A	ATG	AGT	ACG	ATG	GCT
PCK-RS3	M	A	A	Q	C	ATG	GCG	GCT	CAG	TGT
PCK-RS4	M	A	V	M	C	ATG	GCT	GTG	ATG	TGT
PCK-RS5	M	S	A	A	A	ATG	TCG	GCG	GCG	GCG
PCK-RS6	M	A	I	M	S	ATG	GCG	ATT	ATG	AGT
PCK-RS7	M	A	A	V	V	ATG	GCG	GCG	GTG	GTG
PCK-RS8	M	S	S	M	A	ATG	TCG	TCT	ATG	GCG
PCK-RS9	M	S	S	M	C	ATG	GCG	AGT	ATG	TGT
Related to Figure 4										
mut 1 RS	M	A	M	G	T	ATG	GCT	ATG	GGT	ACA
mut 2 RS	M	S	Y	L	C	ATG	TCT	TAT	TTG	TGT
Related to Figure S10										
CbzK-RS1	M	A	Y	A	A	ATG	GCG	TAT	GCT	GCT
CbzK-RS2	M	A	M	G	A	ATG	GCG	ATG	GGG	GCT
CbzK-RS3	M	A	A	M	V	ATG	GCG	GCT	ATG	GTT
CbzK-RS4	M	A	A	A	A	ATG	GCT	GCG	GCG	GCG
CbzK-RS5	M	A	A	L	V	ATG	GCT	GCG	CTG	GTG
CbzK-RS6	L	A	T	M	C	CTG	GCT	ACT	ATG	TGT
CbzK-RS7	M	A	G	M	C	ATG	GCT	GGG	ATG	TGT
CbzK-RS8	M	A	A	M	C	ATG	GCT	GCG	ATG	TGT
CbzK-RS9	M	D	R	G	V	ATG	GAT	AGG	GGG	GTG
CbzK-RS10	M	A	A	M	A	ATG	GCT	GCG	ATG	GCG
CbzK-RS11	M	A	A	V	V	ATG	GCG	GCG	GTG	GTG
CbzK-RS12	M	A	G	M	V	ATG	GCG	GGG	ATG	GTG
CbzK-RS13	M	A	A	L	A	ATG	GCT	GCG	CTG	GCG
CbzK-RS14	M	A	Y	M	A	ATG	GCT	TAT	ATG	GCG
CbzK-RS15	M	A	Y	V	A	ATG	GCT	TAT	GTG	GCG
CbzK-RS16	M	A	M	V	G	ATG	GCT	ATG	GTG	GGT
CbzK-RS17	M	A	G	L	A	ATG	GCG	GGT	TTG	GCG
CbzK-RS18	M	A	Y	V	G	ATG	GCT	TAT	GTG	GGG
CbzK-RS19	M	A	Y	G	C	ATG	GCG	TAT	GGT	TGT
CbzK-RS20	M	S	A	A	A	CTG	TCG	GCG	GCG	GCG
CbzK-RS21	M	A	Y	Q	C	ATG	GCT	TAT	CAG	TGT
CbzK-RS22	M	A	A	M	G	ATG	GCT	GCG	ATG	GGG
CbzK-RS23	F	A	A	V	C	TTT	GCT	GCT	GTG	TGT

CbzK-RS24	M	A	T	M	G	ATG	GCT	ACG	ATG	GGG
CbzK-RS25	M	A	A	Q	C	ATG	GCG	GCT	CAG	TGT
CbzK-RS26	M	A	I	Q	C	ATG	GCT	ATT	CAG	TGT
CbzK-RS27	M	S	C	I	A	ATG	TCG	TGT	ATT	GCT
CbzK-RS28	M	A	Y	V	C	ATG	GCT	TAT	GTT	TGT
CbzK-RS29	M	A	Y	V	V	ATG	GCT	TAT	GTT	GTG
CbzK-RS30	M	S	Y	M	G	ATG	TCT	TAT	ATG	GGT
CbzK-RS31	M	A	Y	M	G	ATG	GCC	TAT	ATG	GGT
CbzK-RS32	M	S	C	M	C	ATG	GCG	ACT	ATT	TGT
CbzK-RS33	M	A	I	M	S	ATG	GCG	ATT	ATG	AGT
CbzK-RS34	M	A	M	M	A	ATG	GCG	ATG	ATG	GCT
CbzK-RS35	M	S	T	M	A	ATG	AGT	ACG	ATG	GCT
CbzK-RS36	L	S	A	M	A	ATG	ACT	ATT	ATG	GGT
CbzK-RS37	F	S	C	M	C	TTT	AGT	TGT	ATG	TGT
CbzK-RS38	M	S	T	M	G	ATG	TCG	ACT	ATG	GGT
CbzK-RS39	L	S	A	M	C	CTT	AGT	GCG	ATG	TGT
CbzK-RS40	M	A	S	M	G	ATG	GCT	TCT	ATG	GGT
CbzK-RS41	M	A	L	I	G	ATG	GCG	TTG	ATT	GGT
CbzK-RS42	F	A	C	I	A	TTT	GCG	TGT	ATT	GCT
CbzK-RS43	M	S	Y	V	G	ATG	AGT	TAT	GTT	GGG
CbzK-RS44	F	A	A	C	C	TTT	GCG	GCT	TGT	TGT
CbzK-RS45	M	A	T	I	C	ATG	GCG	ACT	ATT	TGT
CbzK-RS46	M	A	V	M	C	ATG	GCT	GTG	ATG	TGT
CbzK-RS47	M	S	V	M	V	ATG	GCT	GCT	ATG	GGT
CbzK-RS48	M	A	I	V	C	ATG	GCT	ATT	GTG	TGT
CbzK-RS49	M	S	L	V	A	ATG	TCT	TTG	GTG	GCG
CbzK-RS50	M	A	Y	I	G	ATG	GCG	TAT	ATT	GGT
CbzK-RS51	M	H	Y	K	F	ATG	CAT	TAT	AAG	TTT
CbzK-RS52	M	S	S	M	C	ATG	TCG	AGT	ATG	TGT
CbzK-RS53	M	A	T	V	C	ATG	GCT	ACG	GTG	TGT
CbzK-RS54	M	S	S	M	A	ATG	TCG	TCT	ATG	GCG
CbzK-RS55	M	S	L	V	G	ATG	AGT	TTG	GTG	GGT
CbzK-RS56	M	A	T	A	A	ATG	GCT	ACT	GCT	GCG
CbzK-RS57	F	A	T	M	A	TTT	GCG	ACG	ATG	GCT
CbzK-RS58	F	S	T	M	C	TTT	TCG	ACG	ATG	TGT
CbzK-RS59	M	A	I	N	C	ATG	GCG	ATT	AAT	TGT
CbzK-RS60	M	S	R	L	C	ATG	TCG	CGT	CTT	TGT
CbzK-RS61	M	T	I	M	C	ATG	ACG	ATT	ATG	TGT
CbzK-RS62	M	S	R	L	H	ATG	AGT	CGT	CTT	CAT
CbzK-RS63	M	T	Y	V	G	ATG	ACG	TAT	GTT	GGT
CbzK-RS64	M	A	Y	L	C	ATG	GCC	TAT	CTG	TGC

CbzK-RS65	M	H	Y	T	V	ATG	CAT	TAT	ACG	GTT
CbzK-RS66	M	S	I	A	A	ATG	AGT	ATT	GCT	GCT
CbzK-RS67	M	H	M	L	G	ATG	CAT	ATG	TTG	GGT
CbzK-RS68	M	H	M	L	I	ATG	CAT	ATG	TTG	ATT
CbzK-RS69	F	H	Y	Q	I	TTT	CAT	TAT	CAG	ATT
CbzK-RS70	M	H	M	L	V	ATG	CAT	ATG	TTG	GTG
CbzK-RS71	M	A	G	V	V	ATG	GCT	GGG	GTG	GTT
CbzK-RS72	F	S	T	M	A	TTT	TCG	ACG	ATG	GCT
CbzK-RS73	M	H	Y	K	M	ATG	CAT	TAT	AAG	ATG
CbzK-RS74	M	S	T	M	C	ATG	TCG	ACT	ATG	TGT
CbzK-RS75	M	H	Y	N	V	ATG	CAT	TAT	AAT	GTT
CbzK-RS76	M	H	M	M	I	ATG	CAT	ATG	ATG	ATT
CbzK-RS77	G	H	M	L	V	GGG	CAT	ATG	TTG	GTG
CbzK-RS78	M	I	G	N	R	ATG	ATT	GGT	AAT	CGG
CbzK-RS79	M	H	M	L	L	ATG	CAT	ATG	TTG	CTG
CbzK-RS80	S	H	M	L	V	TCT	CAT	ATG	TTG	GTG
CbzK-RS81	R	H	M	L	V	CGG	CAT	ATG	TTG	GTG
CbzK-RS82	R	R	E	D	R	CGG	AGG	GAG	GAT	CGG
CbzK-RS83	F	H	N	L	F	TTT	CAT	AAT	CTG	TTT
CbzK-RS84	M	T	E	F	S	ATG	ACG	GAG	TTC	AGT
CbzK-RS85	M	H	M	L	C	ATG	CAT	ATG	TTG	TGT
CbzK-RS86	E	H	M	L	V	GAG	CAT	ATG	TTG	GTG
CbzK-RS87	R	R	E	D	V	CGG	AGG	GAG	GAT	GTT
CbzK-RS88	M	H	M	L	S	ATG	CAT	ATG	TTG	AGT
CbzK-RS89	T	H	M	L	V	ACG	CAT	ATG	TTG	GTG
CbzK-RS90	M	H	C	K	V	ATG	CAT	TGT	AAG	GTT
CbzK-RS91	M	H	M	L	R	ATG	CAT	ATG	TTG	CGG
CbzK-RS92	R	R	E	D	G	CGG	AGG	GAG	GAT	GGT
CbzK-RS93	I	H	M	L	V	ATT	CAT	ATG	TTG	GTG
CbzK-RS94	M	H	M	L	W	ATG	CAT	ATG	TTG	TGG
CbzK-RS95	M	H	Y	K	V	ATG	CAT	TAT	AAG	GTT
CbzK-RS96	M	W	G	V	S	ATG	TGG	GGG	GTG	AGT

	Amino acid mutations					Codon mutations				
Amino acid position	**317**	**318**	**319**	**320**	**321**	**317**	**318**	**319**	**320**	**321**
SepRS(v1.0)	**M**	**N**	**L**	**G**	**L**	**ATG**	**AAC**	**CTG**	**GGT**	**CTG**
pThrRS	M	N	L	A	Y	ATG	AAT	TTG	GCT	TAT

Supplementary Table 1. Sequences of aminoacyl-tRNA synthetase mutants described in this study.

Ub12TPO	
Wavelength [Å]	0.97199
Resolution range [Å]	43.81 - 1.07 (1.09 - 1.07)
Space group	C 1 2 1
Unit cell (a, b, c) (α, ß, γ) [°]	(50.46 24.74 47.65) (90 113.17 120)
Total reflections	66501
Unique reflections	22955
Multiplicity	2.9 (3.0)
Completeness (%)	95.6 (94.7)
Mean I/σ(I)	7.5 (2.5)
Rmerge (all I+ and I-)	0.087 (0.652)
R-meas	0.105 (0.786)
Reflections used in refinement	21851
R-work	0.13
R-free	0.164
Number of non-hydrogen atoms	834
Protein	757
Protein residues	76
RMS(bonds)	0.035
RMS(angles)	3.227
Ramachandran favored (%)	100
Ramachandran allowed (%)	0
Ramachandran outliers (%)	0
Rotamer outliers (%)	5.88
Clashscore	9.53
Average B-factor	19.54
Macromolecules	15.41
Solvent	66.72

Supplementary Table 2. Data collection and refinement statistics for Ub12pThr.

Barcode ID	Sequence
P7-1	ACCTCA
P7-2	CCTTCT
P7-3	CTACCT
P7-4	CTCACA
P7-5	CTCGAA
P7-6	GGAGAA
P7-7	GGATCT
P7-8	TCGACT
P7-9	AACACC
P7-10	GAGTCA
P7-11	GTAACC
P5-1	ATGTCC
P5-2	CAACTC
P5-3	CACAGT
P5-4	CGTAGA

Supplementary Table 3. Hexanucleotide barcode sequences used for deep sequencing sample multiplexing.

		Deep sequencing-based result	
		Substrate-selective	Promiscuous
Phenotype result	Substrate-selective	True positive	False negative
	Promiscuous	False positive	True negative

Supplementary Table 4. Example of a contingency table.