Supporting Information

Efficient Site-Specific Prokaryotic and Eukaryotic Incorporation of Halotyrosine Amino Acids into Proteins

Hyo Sang Jang[1], Xiaodong Gu[2], Richard B. Cooley[1], Joseph J. Porter[1], Rachel L. Henson[1], Taylor Willi[1], Joseph A. DiDonato[2,3], Stanley L. Hazen[2,3,4,*], Ryan A. Mehl[1,*]

[1] Department of Biochemistry and Biophysics, Oregon State University, Corvallis, Oregon 97331
[2] Department of Cardiovascular & Metabolic Sciences, Lerner Research Institute, Cleveland Clinic, Cleveland, Ohio 44195
[3] Center for Microbiome & Human Health, Lerner Research Institute, Cleveland Clinic, Cleveland, Ohio 44195
[4] Department of Cardiovascular Medicine, Heart and Vascular Institute, Cleveland Clinic, Cleveland, Ohio 44195

Running title: Genetic incorporation of halotyrosine derivatives into proteins

<u>Materials included</u>:
Tables S1-S2
Figures S1-S14
Halotyrosine machinery system sequences

Table S1. Sequences of *M. barkeri* halotyrosine synthetase hits.

Mb Pyl RS hit	Leu270	Tyr271	Leu274	Asn311	Cys313	Off Site
D1	Leu	Tyr	Leu	Ser	Cys	I215V / I285V
B2	Leu	Tyr	Leu	Ser	Cys	
C2	Leu	Tyr	Leu	Ser	Ala	
B3	Phe	Tyr	Leu	Ser	Cys	E330G
E3	Leu	Tyr	Leu	Ser	Cys	N401T
H3	Leu	Tyr	Leu	Cys	Ala	
A4	Leu	Tyr	Leu	Ser	Ser	
D4	Leu	Tyr	Leu	Cys	Cys	
E4	Phe	Tyr	Leu	Cys	Cys	T269A
F4	Leu	Tyr	Leu	Ser	Ser	
G4	Leu	Tyr	Leu	Cys	Ser	
C5	Leu	Tyr	Leu	Gly	Ser	
B6	Gln	Tyr	Leu	Gly	Thr	
C6	**Ser**	**Leu**	**Leu**	**Gly**	**Thr**	
E6	Leu	Tyr	Leu	Gly	Thr	

Table S2. DNA oligomers used to construct eukaryotic expression vectors.

Primer name	DNA oligo sequence
JP1	5’-TAAACTTgctagcgccGCCACCATGGCGTGTCCGGTTCCTTTGCAGTTGCCTCC-3’
JP2	5’-ccgcaggtagttGAGTGAggtgggggccagGGTaggccgcaggcac-3’
JP3	5’-gtgcctgcggcctACCctggcccccaccTCACTCaactacctgcgg-3’
JP4	5’-cagccgctgcccatctgGGTaaaACCcaccattgtaaac-3’
JP5	5’-gtttacaatggtgGGTtttACCcagatgggcagcggctg-3’
JP6	5’-CAGCGGGTTTAAACGGGCCCTCTAGTCATCACAGGTTGGTGCTGATGCCGTTGTAGTAGC-3’
JP7	5’-CTGTGTGCTAGCgccgccaccATGGTTTCTAAAGGTGAAGAACTTTTTACTGG-3’
JP8	5’-ctgcaaGAATTCTTAGTGGTGATGGTGGTGATGAGTAGAATCCAGTCCC-3’
JP9	5’-TAAACTTgctagcGCCACCATGGCGTGTCCGGTTCCTTTGCAGTTGCCTCC-3’
JP10	5’-AGTggagaattcTCATCACAGGTTGGTGCTGATGCCGTTGTAGTAGCTCTCGC-3’

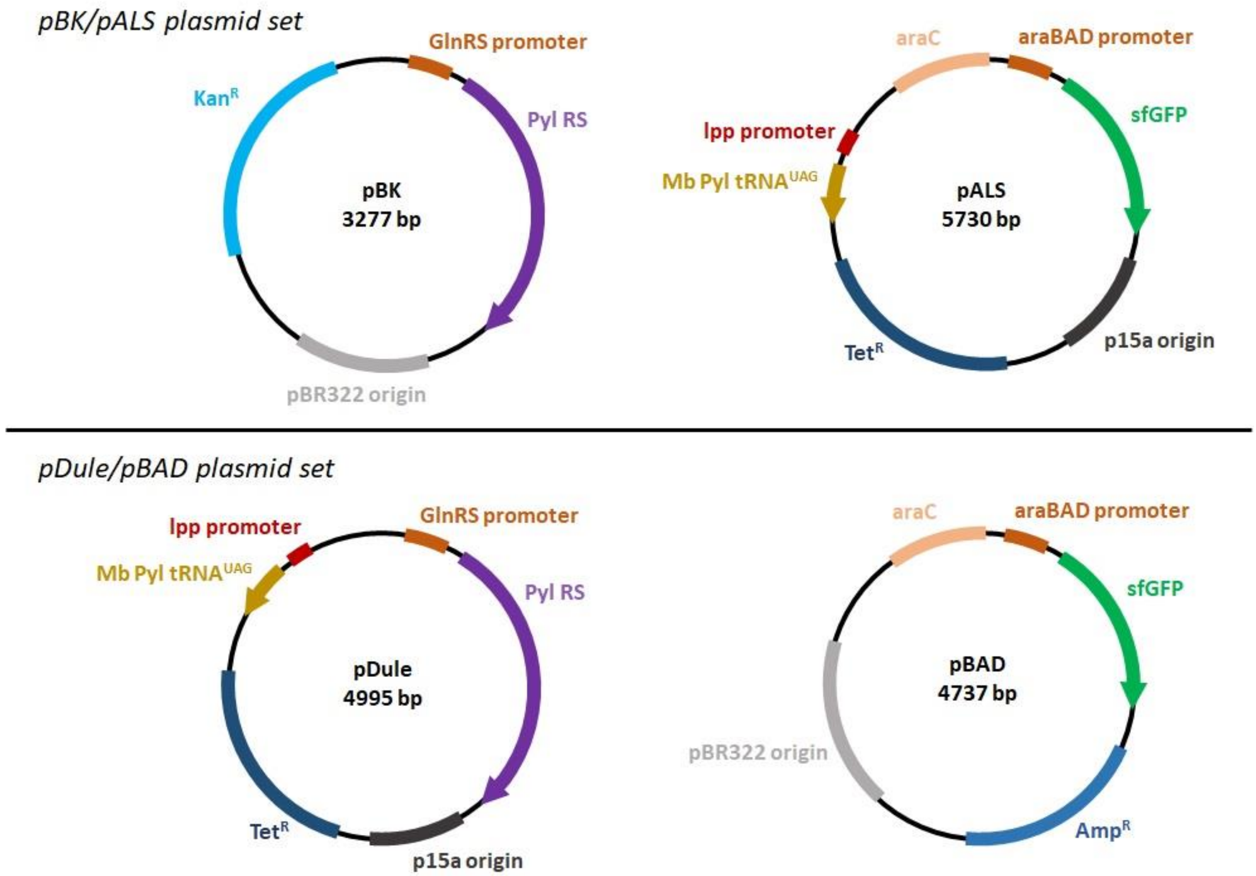

Figure S1. Plasmid maps for the two plasmid sets: The pBK/pALS pairing is used in the initial screening of library members. The pDule/pBAD pairing was used for standard halotyrosine protein expression.

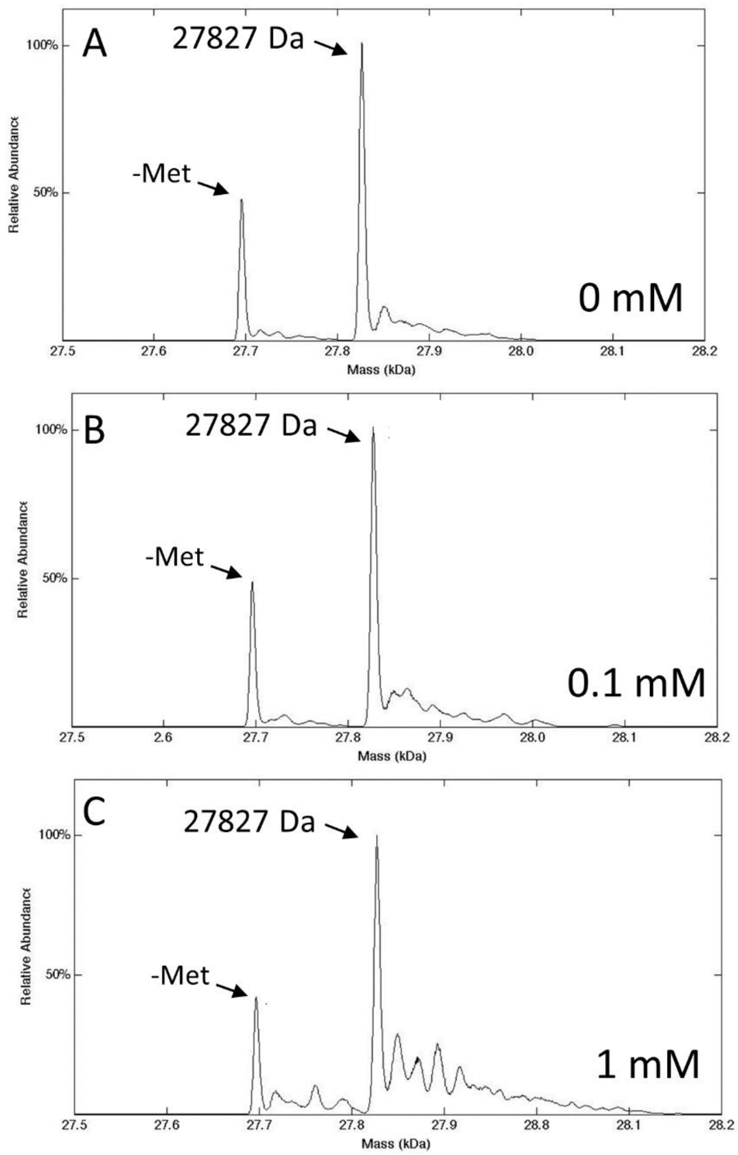

Figure S2. ESI-MS spectra of purified wild-type sfGFP expressed in (A) 0 mM, (B) 0.1 mM and (C) 1.0 mM 3-chloroTyr using DH10b cells and 50 mL media in a 250 mL baffled flask.

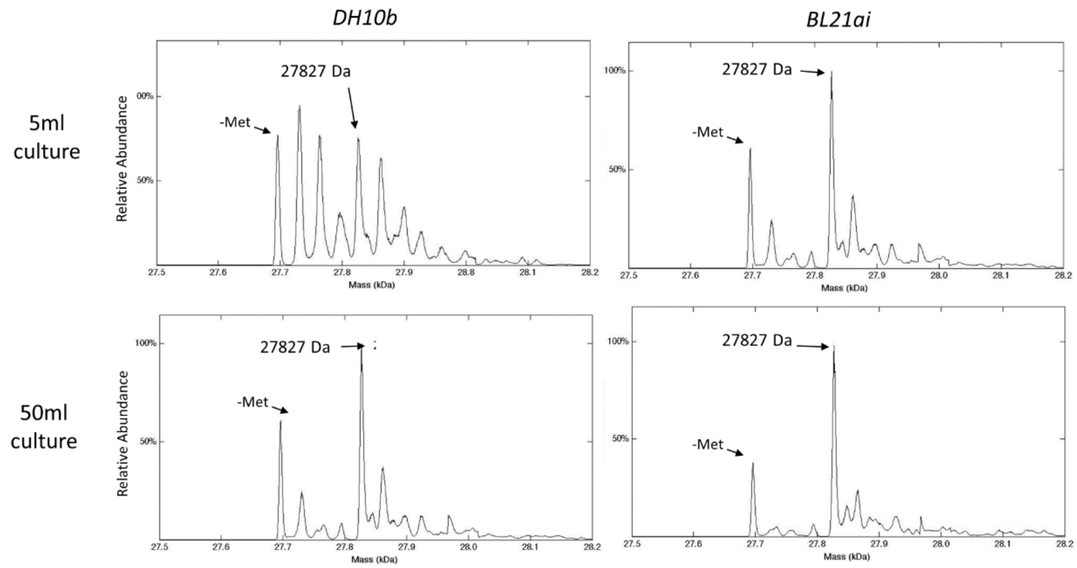

Figure S3. Dependence of tyrosine residue specific incorporation of 3-chloroTyr into wild-type sfGFP on culture volume and *E. coli* strain. Left column: DH10b cells, Right column: BL21ai cells. Top row: 5 mL culture volume in a 15 mL culture tube (low aeration), Bottom row: 50 mL cultures in a 250 mL baffled flask (high aeration).

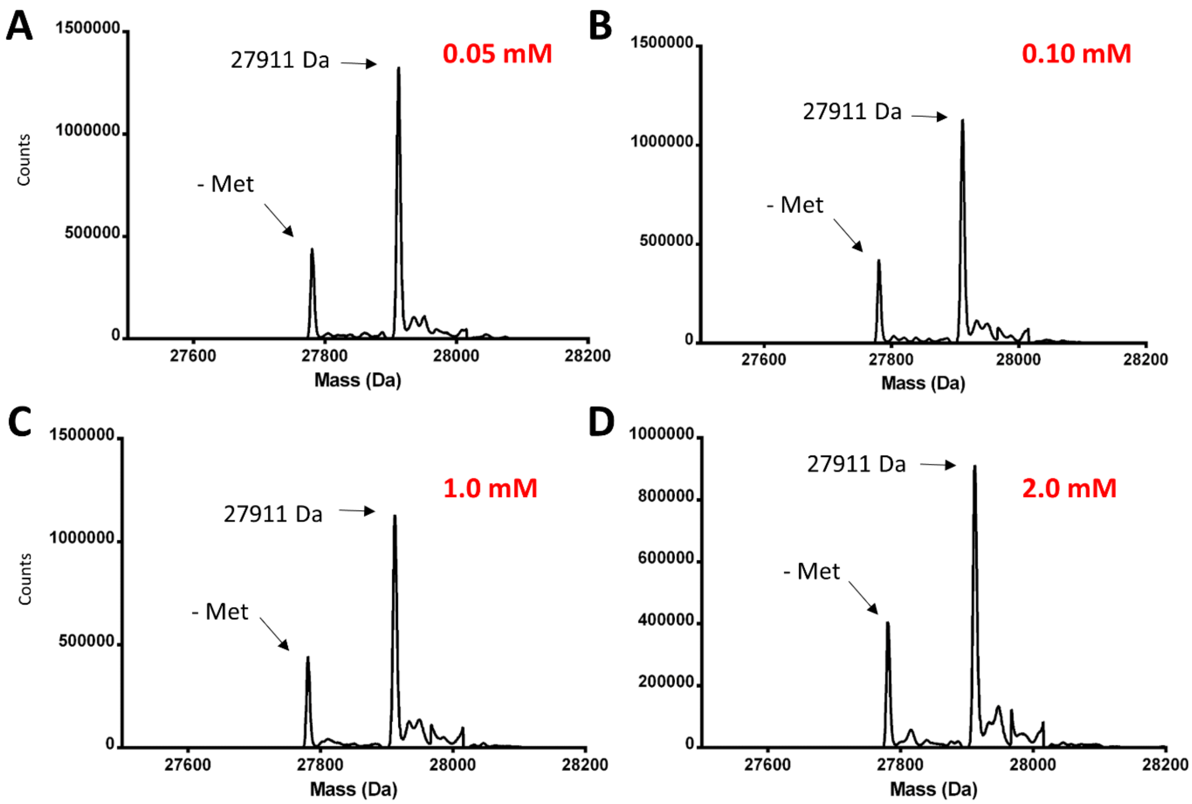

Figure S4. Mass spectrometry of sfGFP-150-3-chloroTyr purified from BL21ai cell expression cultures grown at different concentrations of 3-chloroTyr. (A) 0.05 mM 3-chloroTyr, (B) 0.1 mM 3-chloroTyr, (C) 1.0 mM 3-chloroTyr and (D) 2.0 mM 3-chloroTyr in the media.

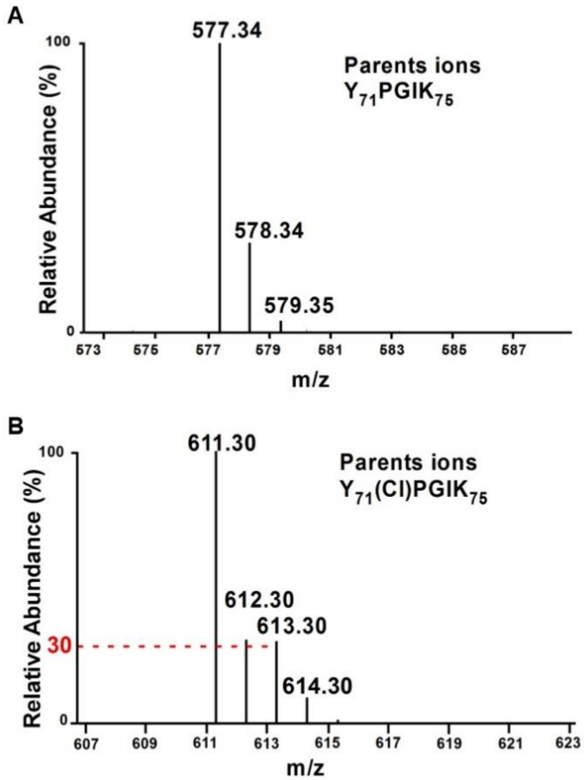

Figure S5. MS spectra of purified recombinant PON1 expressed in *E. coli*. Parents ions of peptides YPGIK (A) and Y71(Cl)PGIK75 (B) are shown. m/z 611.30 is the parent ion of Y71(^{35}Cl)PGIK75 and m/z 613.30 is the parent ion of Y71(^{37}Cl)PGIK75. The ratio of intensities of m/z 611.30 and m/z 613.30 is about 3:1, which is consistent with the natural abundance of ^{35}Cl and ^{37}Cl.

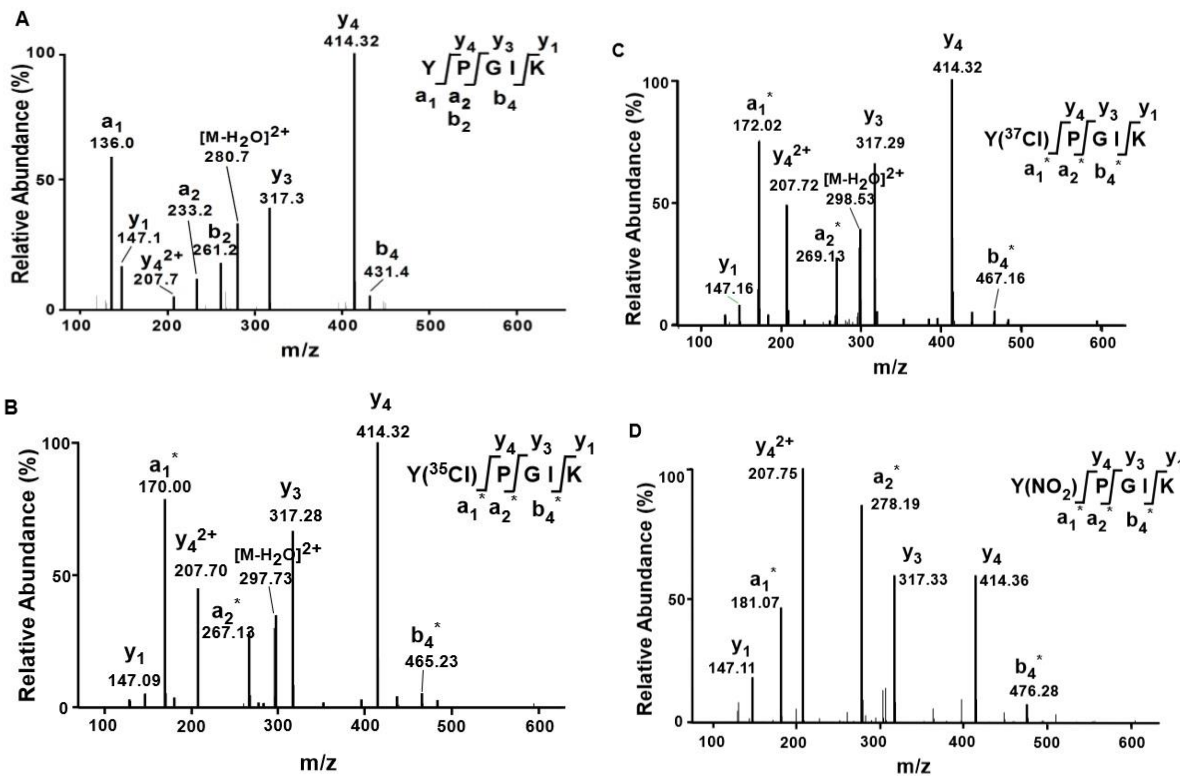

Figure S6. MS/MS spectra of purified recombinant PON1 in *E. coli*. (A) MS/MS of Y71PGIK75, (B) Y71(^{35}Cl)PGIK75, (C) Y71(^{37}Cl)PGIK75 and (D) Y71(NO_2)PGIK75

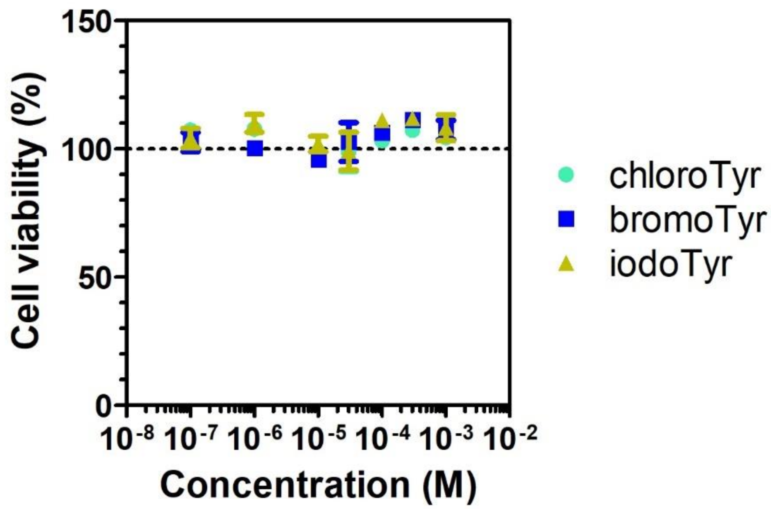

Figure S7. The effect of halotyrosine on the viability of HEK293T cells. The viability of HEK293T cells exposed for 48 h to 3-chloroTyr, 3-bromoTyr, or 3-iodoTyr was measured using the CellTiter Glo assay kit. n =3 ± S.E.M.

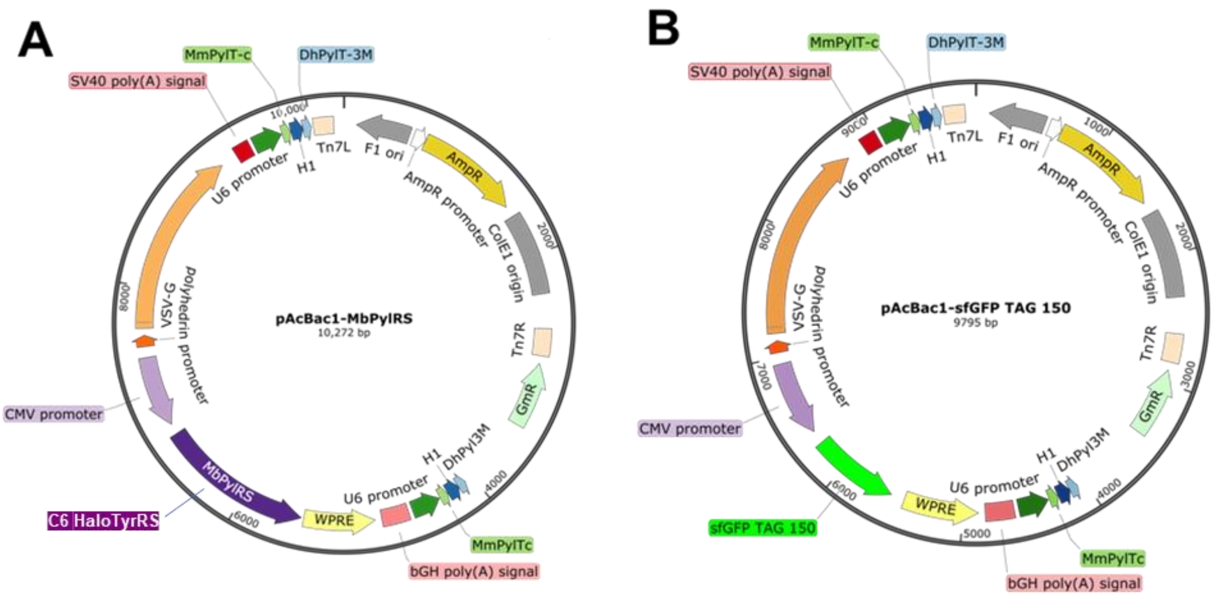

Figure S8. pAcBac1 expression vector maps. (A) pAcBac1-HaloTyrRS. (B) pAcBac1-sfGFP-150TAG.

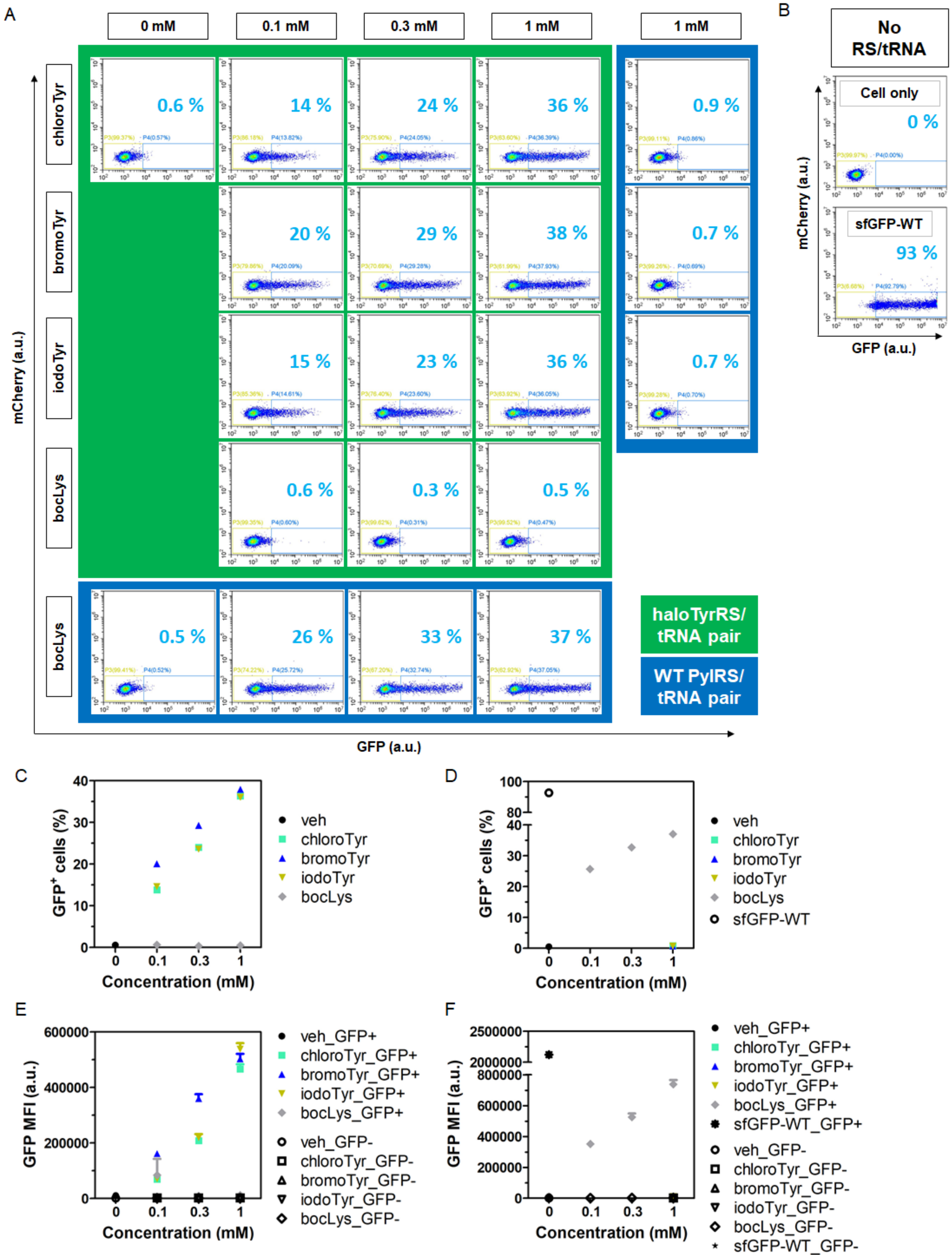
A
0 mM
0.1 mM
0.3 mM
1 mM
1 mM
chloroTyr
bromoTyr
iodoTyr
bocLys
bocLys
mCherry (a.u.)
GFP (a.u.)
0.6 %
14 %
24 %
36 %
0.9 %
20 %
29 %
38 %
0.7 %
15 %
23 %
36 %
0.7 %
0.6 %
0.3 %
0.5 %
0.5 %
26 %
33 %
37 %
haloTyrRS/ tRNA pair
WT PylRS/ tRNA pair
B
No RS/tRNA
Cell only
0 %
sfGFP-WT
93 %
mCherry (a.u.)
GFP (a.u.)
C
GFP+ cells (%)
Concentration (mM)
veh
chloroTyr
bromoTyr
iodoTyr
bocLys
D
GFP+ cells (%)
Concentration (mM)
veh
chloroTyr
bromoTyr
iodoTyr
bocLys
sfGFP-WT
E
GFP MFI (a.u.)
Concentration (mM)
veh_GFP+
chloroTyr_GFP+
bromoTyr_GFP+
iodoTyr_GFP+
bocLys_GFP+
veh_GFP-
chloroTyr_GFP-
bromoTyr_GFP-
iodoTyr_GFP-
bocLys_GFP-
F
GFP MFI (a.u.)
Concentration (mM)
veh_GFP+
chloroTyr_GFP+
bromoTyr_GFP+
iodoTyr_GFP+
bocLys_GFP+
sfGFP-WT_GFP+
veh_GFP-
chloroTyr_GFP-
bromoTyr_GFP-
iodoTyr_GFP-
bocLys_GFP-
sfGFP-WT_GFP-

Figure S9. Suppression of the TAG mutation using the haloTyrRS in mammalian cells. HEK293T cells were transfected for 24 h with pAcBac1-haloTyrRS and pAcBac1-sfGFP-150TAG (A, green panel), or pAcBac1-WT pylRS and pAcBac1-sfGFP-150TAG (A, blue panel) followed by flow cytometry analysis. The cells without any transfection (cell only) and sfGFP-WT were used as a negative and positive control, respectively (B). The percentage numbers in blue in panels A and B indicate the percentage ratio of the GFP expressing cells. The percentage of the GFP expressing cells were plotted for the haloTyrRS/tRNA pair (C) and the WT pylRS/tRNA pair (D). The MFI of the GFP expressing (GFP+) and the GFP non-expressing cells (GFP-) is shown for the haloTyrRS/tRNA pair (E) and the WT pylRS/tRNA pair (F).

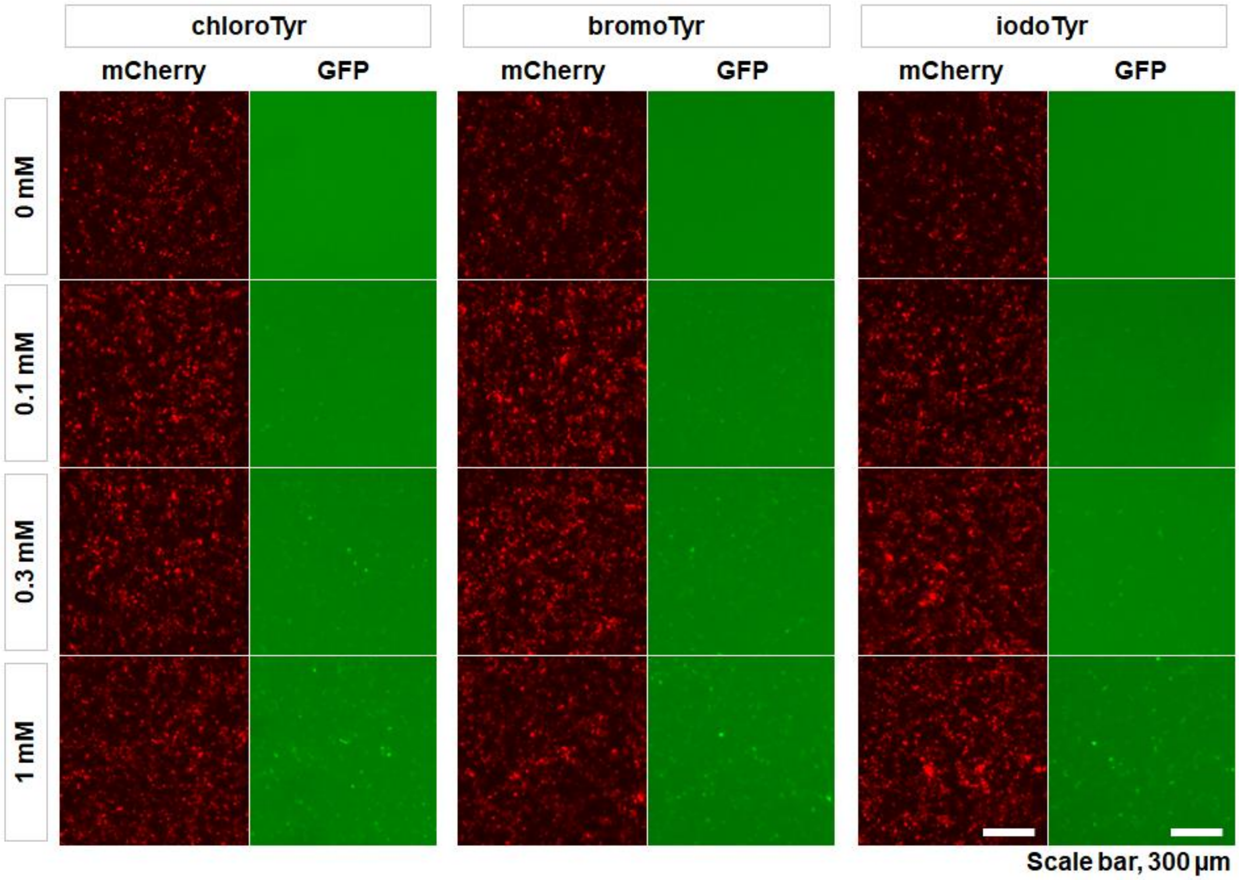

Figure S10. Suppression of the TAG mutation on the fluorescent reporter. HEK293T cells expressing *mCherry-TAG-EGFP-HA* and the haloTyrRS/tRNA pair in the presence or absence of haloTyr were subjected to epifluorescence microscopy at 24 h after transfection. Scale bar, 300 µm

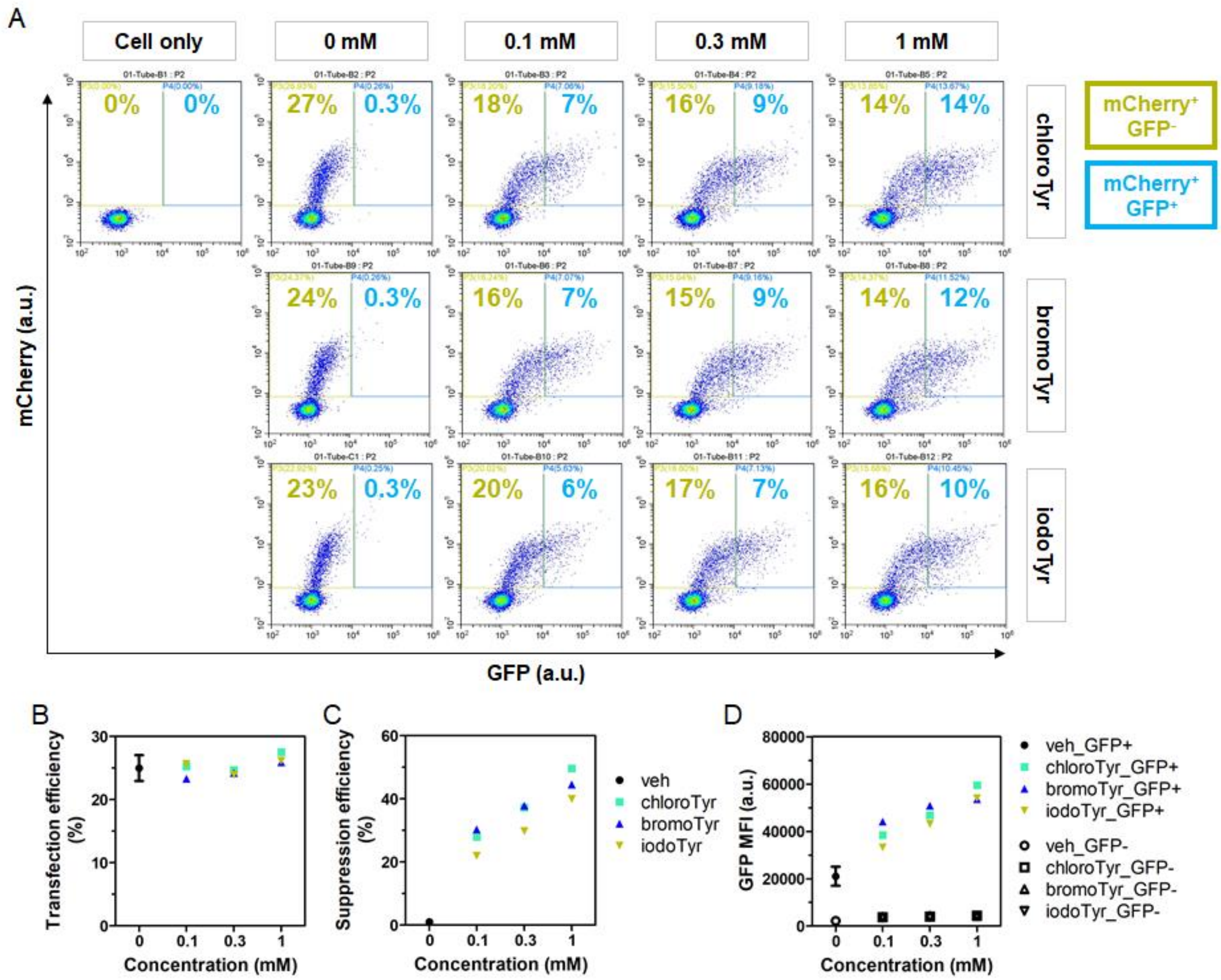

Figure S11. Flow cytometry analysis of TAG suppression in eukaryotic cells. HEK293T cells transfected for 24 h with haloTyrRS/tRNA and mCherry-TAG-EGFP-HA expression vectors were analyzed by flow cytometry (A). The transfection efficiency was defined as the percentage of the mCherry expressing cells (mCherry$^+$) in the total number of cells and calculated from the flow cytometry data (B). The suppression efficiency was defined as the percentage ratio of the mCherry$^+$GFP$^+$ cells and mCherry$^+$ cells (C). The average GFP expression level in the mCherry$^+$GFP$^+$ cells (GFP+) and the mCherry$^+$GFP$^-$ cells (GFP-) is shown as the GFP mean fluorescence intensity (MFI) (D).

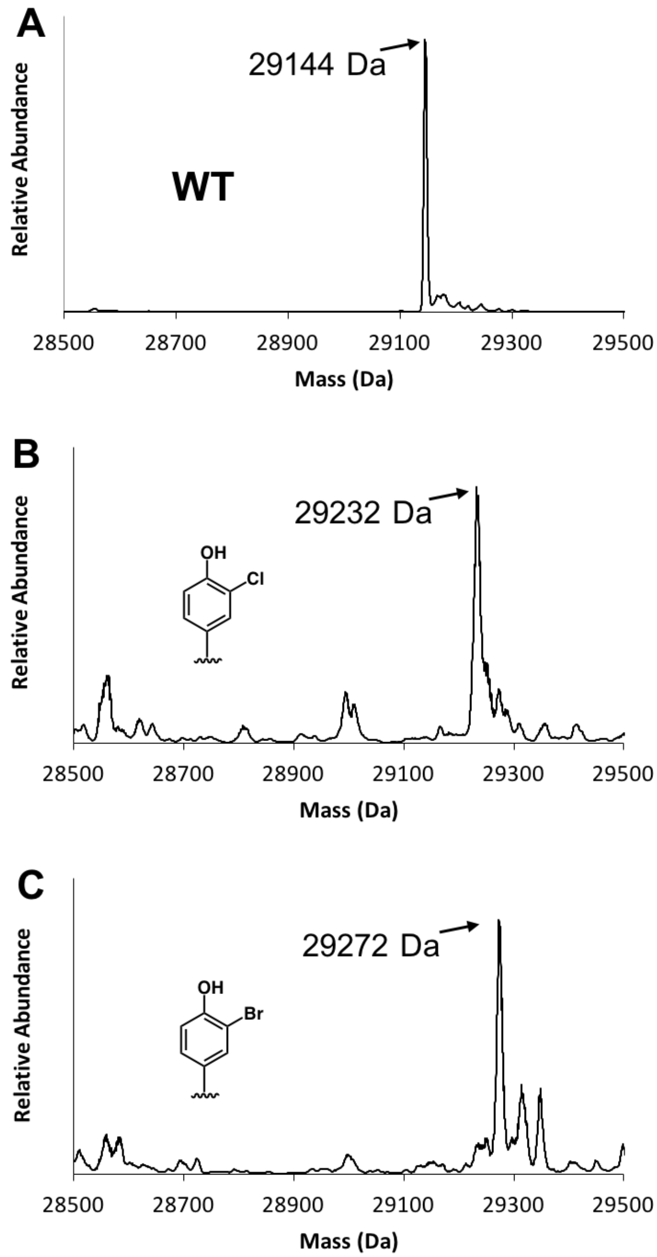

Figure S12. ESI-MS spectra of purified sfGFP expressed in HEK293T cells (A) ESI-MS spectra of wild-type GFP, (B) ESI-MS spectra of sfGFP-150-3-chloroTyr, and (C) ESI-MS spectra of sfGFP-150-3-bromoTyr.

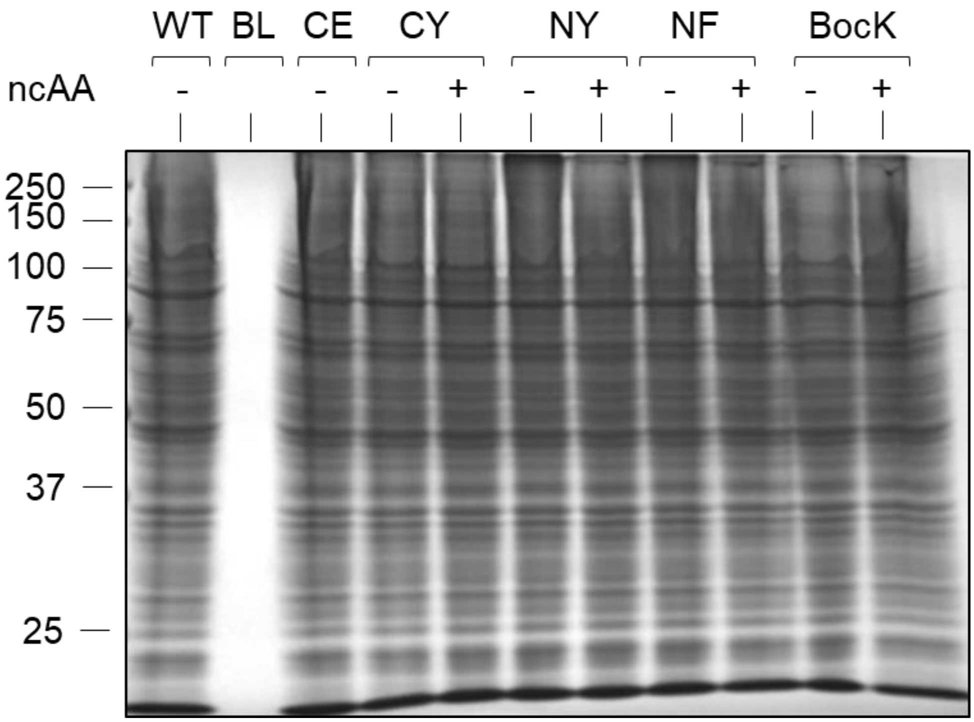

Figure S13. Coomassie Blue stained SDS-PAGE gel for Figure 8. To show even loading of the samples, the gel was stained with Coomassie Brilliant Blue R-250 and destained before imaging.

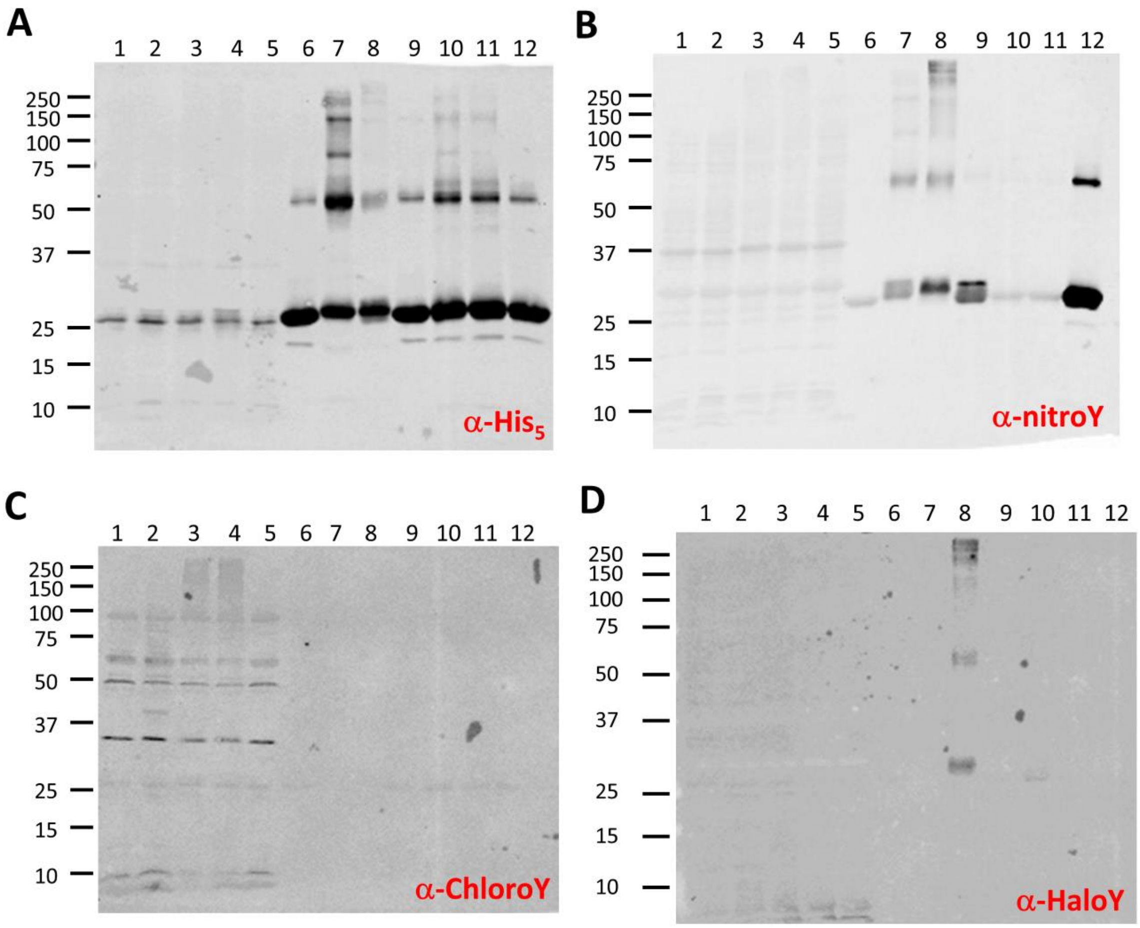

Figure S14. Evaluation of halotyrosine antibodies for detecting halotyrosine containing proteins in Western blot analysis. Immunoblots were probed with (A) anti-His_5 antibody, (B) anti-nitroTyr antibody, (C) anti-chloroTyr antibody, and (D) anti-haloTyr antibody. Lane 1: *E. coli* cell free extract (CFE) from cells expressing His_6-tagged wild-type ApoA1 protein. Lane 2: same as 1 except expression performed in the presence of 1 mM 3-chloroTyr. Lanes 3-5: Same as lane 1 except CFE was treated with 1 mM HOCl, HOBr, and $ONOO^-$, respectively, prior to electrophoresis. Lane 6: wild-type His_6-tagged sfGFP. Lanes 7-9: same as 6 except purified sfGFP was treated with 0.2 mM HOCl, HOBr, and $ONOO^-$, respectively, prior to electrophoresis. Lanes 10-12: His_6-tagged sfGFP with 3-chloroTyr, 3-bromoTyr, and 3-nitroTyr incorporated at position 150.

Halotyrosine GCE machinery system sequences

DNA sequences of components used for expression of halotyrosine-containing protein in HEK cells.

(A) sfGFP-150TAG

ATGGTTTCTAAAGGTGAAGAACTtTTTACtGGCGTTGTGCCGATTCTGGTGGAACTGG
ATGGTGATGTGAATGGCCATAAATTTAGCGTTCGTGGCGAAGGCGAAGGTGATGCGA
CCAACGGTAAACTGACCCTGAAATTTATTTGCACCACCGGTAAACTGCCGGTTCCGT
GGCCGACCCTGGTGACCACCCTGACCTATGGCGTTCAGTGCTTTAGCCGCTATCCGG
ATCATATGAAACGCCATGATTTCTTTAAAAGCGCGATGCCGGAAGGCTATGTGCAGG
AACGTACCATTAGCTTCAAAGATGATGGCACCTATAAAACCCGTGCGGAAGTTAAAT
TTGAAGGCGATACCCTGGTGAACCGCATTGAACTGAAAGGTATTGATTTTAAAGAAG
ATGGCAACATTCTGGGTCATAAACTGGAATATAATTTCAACAGCCAT***tag***GTGTATATT
ACCGCCGATAAACAGAAAAATGGCATCAAAGCGAACTTTAAAATCCGTCACAACGT
GGAAGATGGTAGCGTGCAGCTGGCGGATCATTATCAGCAGAATACCCCGATTGGTGA
TGGCCCGGTGCTGCTGCCGGATAATCATTATCTGAGCACCCAGAGCGTTCTGAGCAA
AGATCCGAATGAAAAACGTGATCATATGGTGCTGCTGGAATTTGTTACCGCCGCGGG
CATTACCCACGGTATGGATGAACTTTATAAAGGTTCTGGAAAGCCGATTCCAAATCC
CCTGTTGGGACTGGATTCTACTCATCACCACCATCACCACTAA

sfGFP-150TAG (the TAG codon is indicated as underlined bold italic letters)
V5 tag
6xHis tag

(B) HaloTyrRS

ATGGCGTGTCCGGTTCCTTTGCAGTTGCCTCCACTGGAGCGCCTCACACTCGAC
AAGGACAAGAAACCCCTGGACGTGCTGATCAGCGCCACCGGCCTGTGGATGAGCCG
GACCGGCACCCTGCACAAGATCAAGCACCACGAGGTGTCAAGAAGCAAAATCTACA
TCGAGATGGCCTGCGGCGACCACCTGGTGGTGAACAACAGCAGAAGCTGCCGGACC
GCCAGAGCCTTCCGGCACCACAAGTACAGAAAGACCTGCAAGCGGTGCCGGGTGTC
CGACGAGGACATCAACAACTTTCTGACCAGAAGCACCGAGAGCAAGAACAGCGTGA
AAGTGCGGGTGGTGTCCGCCCCCAAAGTGAAGAAAGCCATGCCCAAGAGCGTGTCC
AGAGCCCCCAAGCCCCTGGAAAACAGCGTGTCCGCCAAGGCCAGCACCAACACCAG
CCGCAGCGTGCCCAGCCCCGCCAAGAGCACCCCCAACAGCTCCGTGCCCGCCTCTGC
TCCTGCTCCCAGCCTGACACGGTCCCAGCTGGACAGAGTGGAGGCCCTGCTGTCCCC
CGAGGACAAGATCAGCCTGAACATGGCCAAGCCCTTCCGGGAGCTGGAACCCGAGC
TGGTGACCCGGCGGAAGAACGACTTCCAGCGGCTGTACACCAACGACCGGGAGGAC
TACCTGGGCAAGCTGGAACGGGACATCACCAAGTTCTTCGTGGACCGGGGCTTCCTG
GAAATCAAGAGCCCCATCCTGATCCCCGCCGAGTACGTGGAGCGGATGGGCATCAA
CAACGACACCGAGCTGTCCAAGCAGATTTTCCGGGTGGACAAGAAcctgtgcctgcggcct*AC*
*C*ctggcccccacc*TCACTC*aactacctgcggaaactggacagaatcctgcctggccccatcaagattttcgaagtgggaccctgctac
cggaaagagagcgacggcaaagagcacctggaagagtttacaatggtg*GGT*ttt*ACC*cagatgggcagcggctgcACCCGGG
AGAACCTGGAAGCCCTGATCAAAGAGTTCCTGGATTACCTGGAAATCGACTTCGAGA
TCGTGGGCGACAGCTGCATGGTGTACGGCGACACCCTGGACATCATGCACGGCGACC
TGGAACTGAGCAGCGCCGTGGTGGGACCCGTGTCCCTGGACCGGGAGTGGGGCATC
GACAAGCCCTGGATCGGAGCCGGCTTCGGCCTGGAACGGCTGCTGAAAGTGATGCA

CGGCTTCAAGAACATCAAGCGGGCCAGCAGAAGCGAGAGCTACTACAACGGCATCA
GCACCAACCTGTGATGA

N-terminal nuclear export signal (bold and underlined)
Active site mutations as compared to WT Mb PylRS (italics and underlined)