Supporting Information

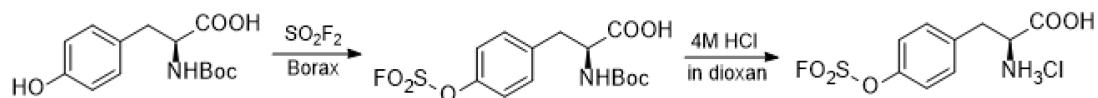
Genetically encoding fluorosulfate-L-tyrosine to react with lysine, histidine and tyrosine via SuFEx in proteins in vivo

Nanxi Wang[1‡], Bing Yang[1‡], Caiyun Fu[1,4‡], He Zhu[3], Feng Zheng[2], Tomonori Kobayashi[1], Jun Liu[1], Shanshan Li[1], Cheng Ma[3], Peng G. Wang[3], Qian Wang[2], Lei Wang[1]*

1. Department of Pharmaceutical Chemistry and the Cardiovascular Research Institute, University of California San Francisco, 555 Mission Bay Blvd. South, San Francisco, California 94158, USA.
2. Hangzhou Research Institute of Technical Institute of Physics and Chemistry, Chinese Academy of Sciences, 310018, China.
3. Department of Chemistry and Center for Therapeutics and Diagnostics, Georgia State University, Atlanta, Georgia 30302, USA.
4. College of Life Sciences, Zhejiang Sci-Tech University, Hangzhou, 310018, China.

Chemical synthesis of FSY

The fluorosulfate-L-tyrosine HCl salt was synthesized based on the classic SO_2F_2/borax method.[1]

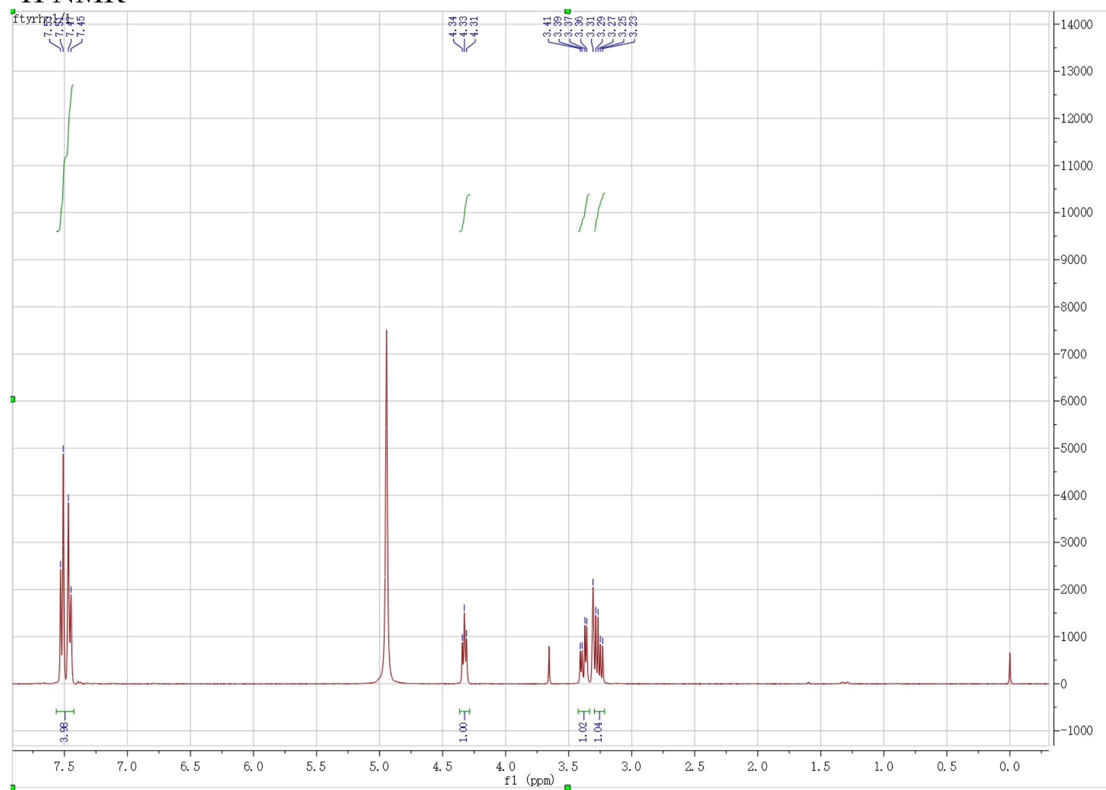


To a 2 L two-neck round-bottom flask containing a magnetic stir bar was added Boc-Tyr-OH (5.00 g, 17.8 mmol), 210 mL of CH_2Cl_2 and 860 mL of a saturated Borax solution. The mixture was stirred vigorously for 20 minutes. The reaction system was vacuumed until the biphasic solution started to degas and refilled with SO_2F_2 for three times. The reaction mixture was stirred vigorously at 25°C overnight. CH_2Cl_2 was carefully removed using a rotary evaporator. Then 1 M aqueous HCl (210 mL) was slowly added to the reaction mixture while stirring and white solid precipitated. The mixture was filtered and the solid was washed with water (80 mL x 3). The white solid was dried under vacuum (1 mm Hg) at 40°C for 4 h affording 6.07 g (16.7 mmol) of the Boc-Tyr-OSO_2F, which was directly used in the next step without any further purification.

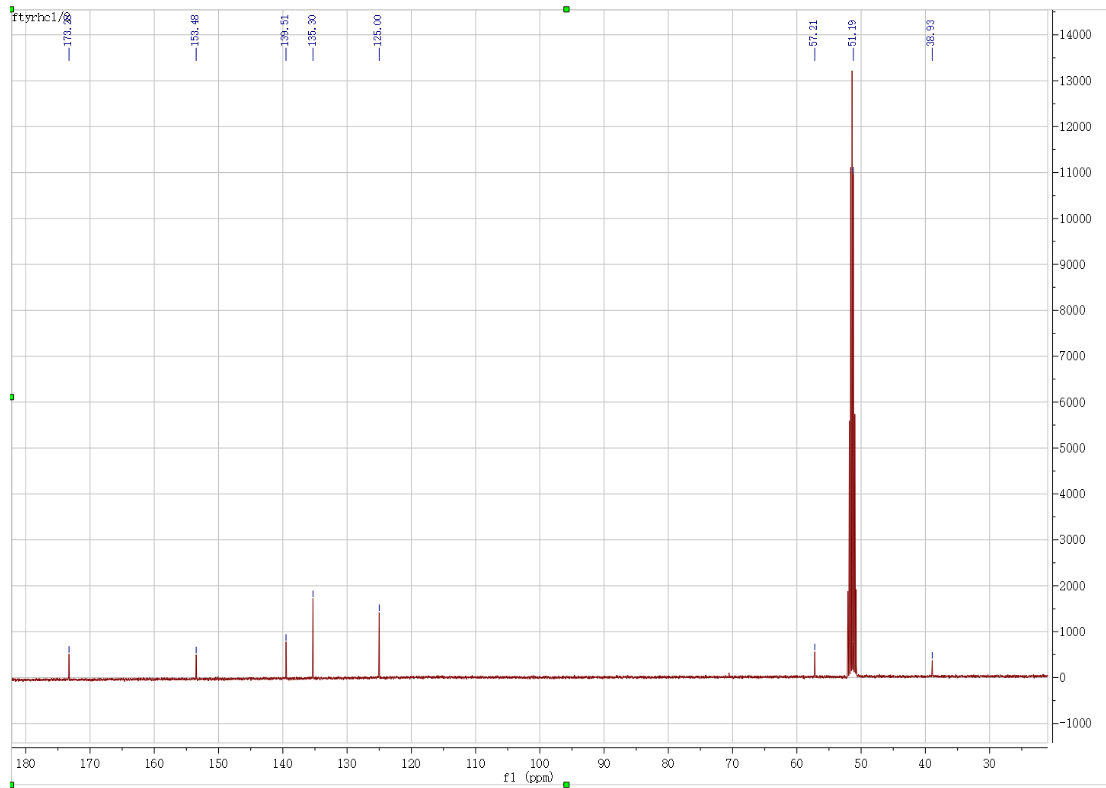
Boc-Tyr-OSO_2F (2.0 g, 5.5 mmol) was treated with 4 M HCl in dioxane (11 mL) and the reaction mixture was stirred overnight, during which white solid precipitated. The solid was filtered and washed by cool ether (5 mL x 2), affording the targeted fluorosulfate-L-tyrosine HCl salt as a white solid (1.46 g, 88% yield).

^1H NMR (400 MHz, CD_3OD): δ (ppm) 3.23-3.41 (m, 2H), 4.32-4.34 (m, 1H), 7.45-7.53 (m, 4H); ^{13}C NMR (400 MHz, CD_3OD): δ (ppm) 38.9, 57.2, 125.0, 135.3, 139.5, 153.5, 173.3; MS: 264.0 $[NH_3\text{-Tyr-}OSO_2F]^+$, 286.0 $[NH_2\text{-Tyr-}OSO_2F + Na]^+$

^{1}H NMR



^{13}C NMR



Synthetase library construction and selection

The pBK-TK3 mutant library of MmPylRS was constructed using the new small-intelligent mutagenesis approach, which uses a single codon for each amino acid and thus allows a greater number of residues to be mutated simultaneously. The following residues of MmPylRS were mutated using the procedures previously described:[2] 302NYT, 305WTG, 306WTG/TAC, 309KYA, 322AYA, 346NDT/VMA/ATG/TGG, 348NDT/VMA/ATG/TGG, 384TTM/TAT, 401VTT, 417NDT/VMA/ATG/TGG.

DH10B cells (100 uL) harboring the pREP positive selection reporter was transformed with 100 ng of pBK-TK3 library via electroporation. The electroporated cells were immediately recovered with 1 mL of pre-warmed SOC media and agitated vigorously at 37 °C for 1 h. The recovered cells were directly plated on a LB-agar selection plate supplemented with 1 mM FSY, 12.5 μg mL^{-1} of tetracycline (Tet), 25 μg mL^{-1} of kanamycin (Kan), and 68 μg mL^{-1} of chloramphenicol (Cm). The selection plate was incubated at 37 °C for 48 h and then stored at room temperature. Colonies showing green fluorescence were diluted in 100 uL of LB and replicated on LB-agar screening plates containing 1) Tet12.5Kan25; 2) Tet12.5Kan25Cm100; 3) Tet12.5Kan25Cm100 supplemented with 1 mM FSY. After 48 h of incubation at 37 °C, 6 clones present FSY-dependent fluorescence and growth were considered as hits and further characterized. The pBK plasmids encoding PylRS mutants were extracted by miniprep and then separated from reporter plasmids by DNA gel electrophoresis. The purified pBK plasmids were analyzed by Sanger-sequencing.

Plasmid construction

pEvol-FSY

pEvol-FSY plasmid was generated by introducing the FSYRS encoding gene into pEvol vector via ligation independent cloning.[3] Briefly, the *FSYRS* gene was amplified with following primers, purified, and ligated into pEvol vectors (linearized with *Bgl* II and *Sal* I) with T4 DNA polymerase.

FSRYS-BglII-F	CTAACAGGAGGAATTAGATCTATGGATAAAAAGCCT
FSYRS-SalI-R	GATGATGATGATGATGGTCGACTTACAGGTTAGTAGAA

pMP-3×$tRNA^{Pyl}_{CUA}$-FSYRS

The pMP-3×$tRNA^{Pyl}_{CUA}$-FSYRS plasmid was constructed by introducing the *FSYRS* gene into pMP vector via standard cloning. The *FSYRS* gene was amplified with following primers, digested with *Nco* I and *Nhe* I, and ligated into the pMP vector pre-treated with the same restriction enzymes.

FSYRS-NcoI-F	TATGCCATGGATAAAAAGCCTTTG
FSYRS-NheI-R	CTATGCTAGCTTACAGGTTAGTAGA

pET-Duet-Afb_{4A}-7X-MBP-Z24TAG

To evaluate the *in vivo* crosslinking ability of FSY, pET-Duet-Afb_{4A}-7X-MBP-Z24TAG plasmids were generated by introducing mutations at residue 7 of Afb_{4A}-7X (X= Lys, Tyr, Cys, Ser, Thr, His, or Ala) gene within the pET-Duet-MBP-Z24TAG expression vector[4] via site-directed mutagenesis. The following primers were used.

Afb-4A7A-F	AACGCGGAACTATCAGTCGCCGGC
Afb-4A7K-F	AACAAAGAACTATCAGTCGCCGGC
Afb-4A7C-F	AACTGCGAACTATCAGTCGCCGGC
Afb-4A7S-F	AACAGCGAACTATCAGTCGCCGGC
Afb-4A7T-F	AACACCGAACTATCAGTCGCCGGC
Afb-4A7H-F	AACCATGAACTATCAGTCGCCGGC
Afb-4A-R	GAACGCGTTGTCTACCATGGTATATCTCC
Afb-4A7Y-F	CCATGGTAGACAACGCGTTCAACTATGAACTATCAGTCGCC
Afb-4A7Y-R	TATATCTCCTTCTTAAAGTTAAACAAAATTATTTCTAGAGGGG

pTak-CaM-76TAG-80Tyr

To investigate the intramolecular crosslinking ability of FSY, residue 76 and 80 of calmodulin encoding gene *CaM* were mutated to an amber stop codon TAG and Tyr respectively. Meanwhile, residue 75, 77, 79, 81 of *CaM* were mutated to Ala via overlapping PCR to assist the crosslinking reaction. The *CaM* gene was amplified with following primers, digested with *Spe* I and *Blp* I, and ligated into the pTak-CaM vector pre-treated with the same restriction enzymes.

CaM-SpeI-F	AACTATGACTAGTCATGACCAACTGAC
80Tyr-R	CGCATACGCGTCCGCCTACGCTCTAGCCATCATAGT
80Tyr-F	TGGCTAGAGCGTAGGCGGACGCGTATGCGGAAGAGGAAATCCG
CaM-BlpI-R	CCAAGCTCAGCTTATTAGTGATGGTGATG

pBad-CysH

To generate pBad-CysH plasmid, the PAPS reductase encoding gene *CysH* was amplified by colony PCR, digested with *Nde* I and *Hind* III, and ligated into the pBad vector pre-treated with the same restriction enzymes.

CysH-NdeI-F	TATACATATGTCCAAACTCGATCTAAACG
CysH-Hind3-R	AGCCAAGCTTTTAATGATGATGATGATGATGCCCTTCGTGTAACCCACATTCC

pBad-Trx35A62TAG

To generate pBad-Trx35A62TAG plasmid, residue 62 of Trx35A gene was mutated into an amber stop codon TAG using site-directed mutagenesis with following primers.

Trx-62TAG-F	GAACATCGATTAGAACCCTGGCAC
Trx-62TAG-R	AGTTTTGCAACGGTCAGTTTG

Protein expression

Afb36FSY

pTak-Afb36TAG-His and pBK-FSYRS were co-transformed into DH10B *E. coli* chemical competent cells. The transformants were plated on an LB-Kan50Cm34 agar plate and incubated overnight at 37 °C. A single colony was inoculated into 5 mL of 2xYT-Kan50Cm34 and cultured overnight at 37 °C. On the following day, 2 mL of overnight cell culture was diluted into 100 mL 2xYT-Kan50Cm34 and agitated vigorously at 37 °C. When OD_{600} reached 0.4~0.6, half of the cell culture (50 mL) was supplemented with 1 mM FSY and 0.5 mM IPTG, then induced at 30 °C for 6 h. As a negative control, the rest 50 mL cell culture was induced with 0.5 mM IPTG at 30 °C for 6 h. Cell pellets were collected by centrifugation at 4200 g for 30 min at 4 °C and stored at -80 °C.

Afb_{4A}-7X and MBP-Z24FSY

The pEvol-FSYRS and pET-Duet-Afb_{4A}-7X-MBP-Z24TAG were co-transformed into BL21(DE3) *E. coli* chemical competent cells. The transformants were plated on an LB-Amp100Cm34 agar plate and incubated overnight at 37 °C. A single colony was inoculated into 5 mL of 2xYT- Amp100Cm34 and cultured overnight at 37 °C. On the following day, 1 mL of overnight cell culture was diluted into 50 mL 2xYT- Amp100Cm34 and agitated vigorously at 37 °C. When OD_{600} reached 0.4~0.6, the cell culture was induced with 0.5 mM IPTG and 0.2% arabinose, then incubated at 37 °C for 6 h. Cell pellets were collected by centrifugation at 4200 g for 30 min at 4 °C and stored at -80 °C.

CaM-76FSY-80Tyr

pBad-CaM76TAG80Tyr and pEvol-FSYRS were co-transformed into BL21(DE3) *E. coli* chemical competent cells. The transformants were plated on an LB-Amp100Cm34 agar plate and incubated overnight at 37 °C. A single colony was inoculated into 5 mL of 2xYT- Amp100Cm34 and cultured overnight at 37 °C. On the following day, 1 mL of overnight cell culture was diluted into 50 mL 2xYT- Amp100Cm34 and agitated vigorously at 37 °C. When OD_{600} reached 0.4~0.6, the cell culture was induced with 0.2% arabinose, then incubated at 37 °C for 6 h. Cell pellets were collected by centrifugation at 4200 g for 30 min at 4 °C and stored at -80 °C.

Trx35A62FSY

pBad-Trx35A62TAG and pEvol-FSYRS were co-transformed into BL21(DE3) *E. coli* chemical competent cells. The transformants were plated on an LB-Amp100Cm34 agar plate and incubated overnight at 37 °C. A single colony was inoculated into 5 mL of 2xYT- Amp100Cm34 and cultured overnight at 37 °C. On the following day, 1 mL of overnight cell culture was diluted into 50 mL 2xYT- Amp100Cm34 and agitated vigorously at 37 °C. When OD_{600} reached 0.4~0.6, the cell culture was induced with 0.2% arabinose, then incubated at 30 °C for 6 h. Cell pellets were collected by centrifugation at 4200 g for 30 min at 4 °C and stored at -80 °C.

PAPS reductase

pBad-CysH was transformed into DH10B *E. coli* chemical competent cells. The transformants were plated on an LB-Amp100 agar plate and incubated overnight at 37 °C. A single colony was

inoculated into 10 mL of 2xYT- Amp100 and cultured overnight at 37 °C. On the following day, 10 mL of overnight cell culture was diluted into 1 L 2xYT- Amp100 and agitated vigorously at 37 °C. When OD_{600} reached 0.4~0.6, the cell culture was induced with 0.2% arabinose, then incubated at 30 °C for 6 h. Cell pellets were collected by centrifugation at 4200 g for 30 min at 4 °C and stored at -80 °C.

His-tag protein purification

Above cell pellets were resuspended in 14 mL lysis buffer (50 mM Tris-HCl pH 8.0, 500 mM NaCl, 20 mM imidazole, 1% v/v Tween 20, 10%v/v glycerol, lysozyme 1 mg/mL, DNase 0.1 mg/mL, and protease inhibitors). The cell suspension was lysed at 4 °C for 30 min. Cell lysate was sonicated with Sonic Dismembrator (Fisher Scientific, 30% output, 3 min, 1 sec off, 1 sec on) in an ice-water bath, followed by centrifugation (20,000 g, 30 min, 4 °C). The soluble fractions were collected and incubated with pre-equilibrated Protino®Ni-NTA Agarose resin (400 µL) at 4 °C for 1 h with constant mechanical rotation. The slurry was loaded onto a Poly-Prep® Chromatography Column, washed with 5 mL of wash buffer (50 mM Tris-HCl pH 8.0, 500 mM NaCl, 20 mM imidazole, and 10%v/v glycerol) for 3 times, and eluted with 200 µL of elution buffer (50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 250 mM imidazole, and 10%v/v glycerol) for 5 times. The eluates were concentrated and buffer exchanged into 100 µL of protein storage buffer (50 mM Tris-HCl, pH 7.4 or 8.0, and 150 mM NaCl) using Amicon Ultra columns, and stored at -80 °C for future analysis.

FACS analysis of Uaa incorporation into HeLa-GFP-182TAG reporter cells

One day before transfection, 4.5×10^4 HeLa-EGFP-182TAG reporter cells[5] were seeded in a Greiner bio-one 24 well-cell culture dish containing 500 µL of DMEM media with 10% FBS, and incubated at 37 °C in a CO_2 incubator. Plasmid pMP-3xtRNA-FSYRS (500 ng, encoding FSYRS and 3 copies of $tRNA^{Pyl}_{CUA}$) was transfected into target cells using 2.5 µL of lipofectamine 2000 following manufacturer's instructions. Six hours post transfection, the media containing transfection complex were replaced with fresh DMEM media with 10% FBS in the presence or absence of 1 mM FSY. For AzF incorporation, plasmid pIre-Azi3[6] was similarly transfected and the DMEM media containing 10% FBS with or without 1 mM AzF were used. After incubation at

37 °C for 24-48 h, transfected cells were trypsinized and collected by centrifugation (1500 rpm, 5 min, r.t.). The cells were resuspended in 300 μL of FACS buffer (1×PBS, 2% FBS, 1 mM EDTA, 0.1% sodium azide, 0.28 μM DAPI) and analyzed by BD LSRFortessa™ cell analyzer.

Fluorescence confocal microscopy of HeLa-EGFP-182TAG reporter cells

One day before transfection, 4.5×10^4 HeLa-EGFP-182TAG cells were seeded in a Greiner bio-one CELLview glass bottom dish containing 500 μL of DMEM media with 10% FBS, and incubated at 37 °C in a CO_2 incubator. Plasmid pMP-3xtRNA-FSYRS (500 ng) was transfected into target cells using 2.5 μL of lipofectamine 2000 following manufacturer's instructions. Six hours post transfection, the media were replaced with complete DMEM media with or without 1 mM FSY. The cells were incubated at 37 °C for additional 24-48 h and imaged with Nikon Eclipse Ti confocal microscope.

Western blot

One day before transfection, 1.8×10^5 HEK293T cells were seeded in an Olympus 6-well tissue culture plate. Plasmid pcDNA-PCNA-165TAG (1.5 μg) and pMP-3xtRNA-FSYRS (1.5 μg) were co-transfected into target cells using 5 uL of lipofectamine 2000 following manufacturer's instructions. Six hours post transfection, the media containing transfection complex were replaced with fresh DMEM media with 10% FBS in the presence or absence of 1 mM FSY. For AzF incorporation, plasmid pcDNA-PCNA-165TAG and pIre-Azi3[6] were similarly co-transfected and DMEM media containing 10% FBS with or without 1 mM AzF were used.

After incubation at 37 °C for 48 h, transfected cells were rinsed with 1× PBS twice, and lysed with western blot solution buffer (containing 0.5 mM EDTA, 20 mM HEPES, and 2% SDS) on ice for 20 min. Cells were scraped off, boiled, and centrifuged at 13000 rpm for 1 min. The total protein concentration was analyzed with Pierce BCA assay following manufacturer's instructions. The samples were separated on SDS-PAGE and immunoblotted with 1:10000 anti-his monoclonal antibody (Proteintech #HRP-66005) to detect full-length PCNA. The internal standard β-actin was probed with 1:500 β-Actin Antibody (C4) (Santa-cruz #sc-47778 HRP). After normalization of

the loaded samples with internal standard β–actin, the incorporation efficiency of FSY was calculated to reach 41% of that of AzF.

Mass spectrometric analysis

Intact FSY-containing Afb were analyzed by ESI-TOF MS using an Agilent 6210 mass spectrometer coupled to an Agilent 1100 HPLC system. Two micrograms of protein samples were injected by an auto-sampler and separated on an Agilent Zorbax SB-C8 column (2.1 mm ID × 10 cm length) by a reverse-phase gradient of 0–80% acetonitrile for 15 min. Mass calibration was performed right before the analysis. Protein spectra were averaged and the charge states were deconvoluted using Agilent MassHunter software.

Protein digestion and tandem mass spectrometry measurement were performed as previously described.[4] The Afb/MBP-Z samples were digested with Glu-C. The CaM and Trx1/PAPS reductase samples were digested by trypsin. Digested peptides were analyzed with an in-line EASY-spray source and nano-LC UltiMate 3000 high-performance liquid chromatography system (Thermo Fisher) interfaced with Elite mass spectrometer (Thermo Fisher). Peptides were eluted over gradient of 2% - 40% buffer B (80% acetonitrile, 20% H_2O, 0.1% formic acid) at flow rate 300 nL/min from EASY-Spray PepMap C18 Columns (50 cm; particle size, 2 μm; pore size, 100 Å; Thermo Fisher). For different samples, slight modifications were made to the separation method. The Elite mass spectrometer was operated in data-dependent mode with one full MS scan at R = 60,000 (m/z = 200) mass range from 375 to 1800 (AGC target 1×10^6), followed by ten CID MS/MS scans. A dynamic exclusion time of 30 s was used, and singly charged ions were excluded. Mass spectrometry raw data was searched by Maxquant.

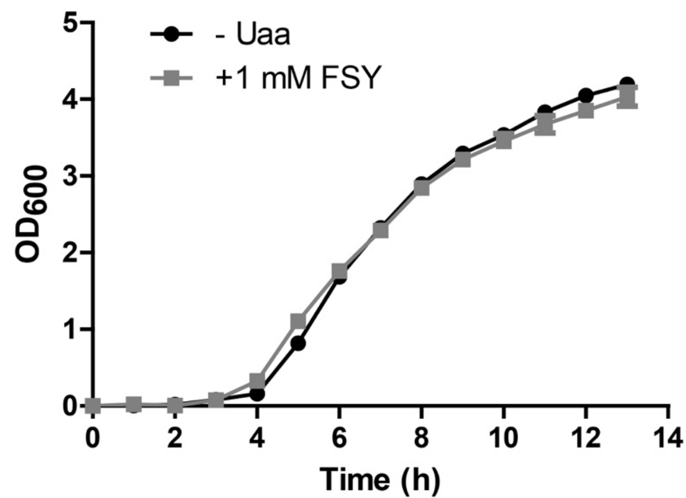


Figure S1. Growth curve of *E. coli* DH10B cells at 37 °C in the presence or absence of 1 mM FSY. The experiments were repeated for three times.

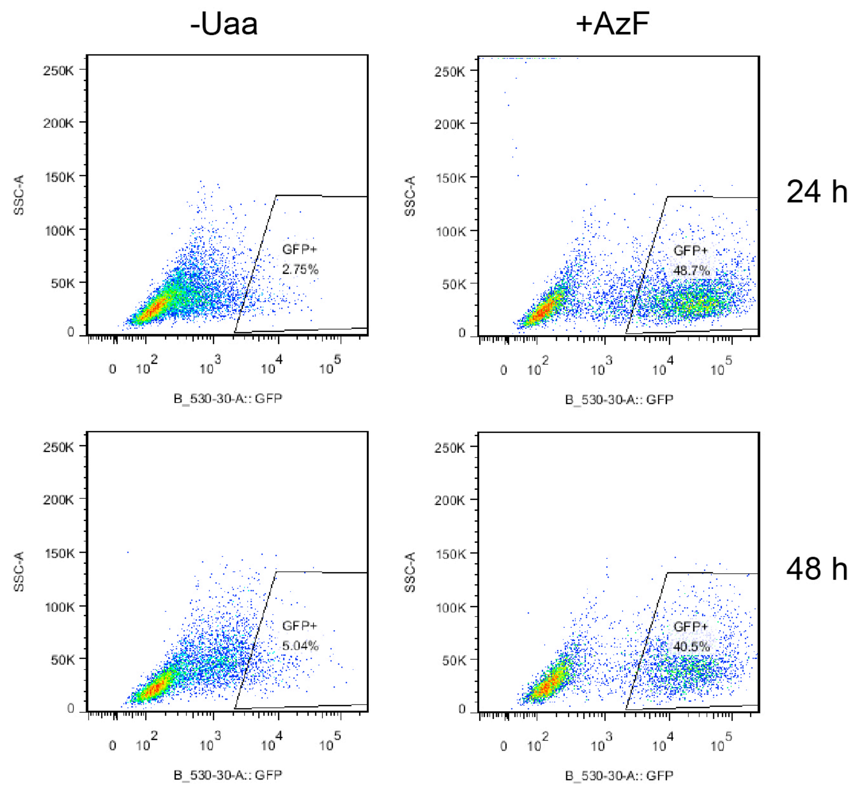


Figure S2. FACS analysis of AzF incorporation into EGFP-182TAG HeLa reporter cells.

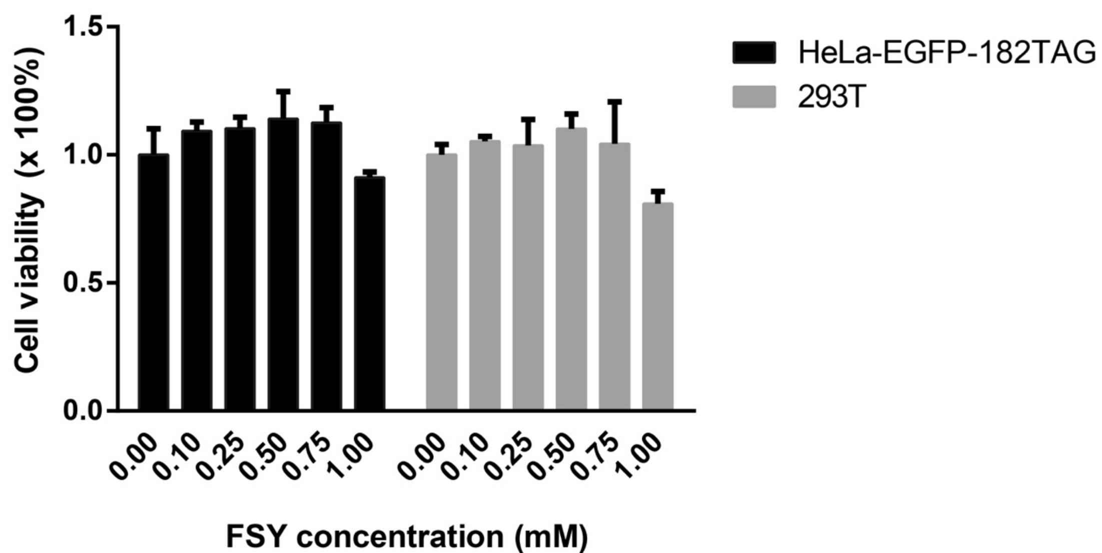


Figure S3. Cell viability assay for HeLa-EGFP-182TAG reporter cells and 293T cells incubated with various concentrations of FSY. Error bars represent s.e.m.; n = 3.

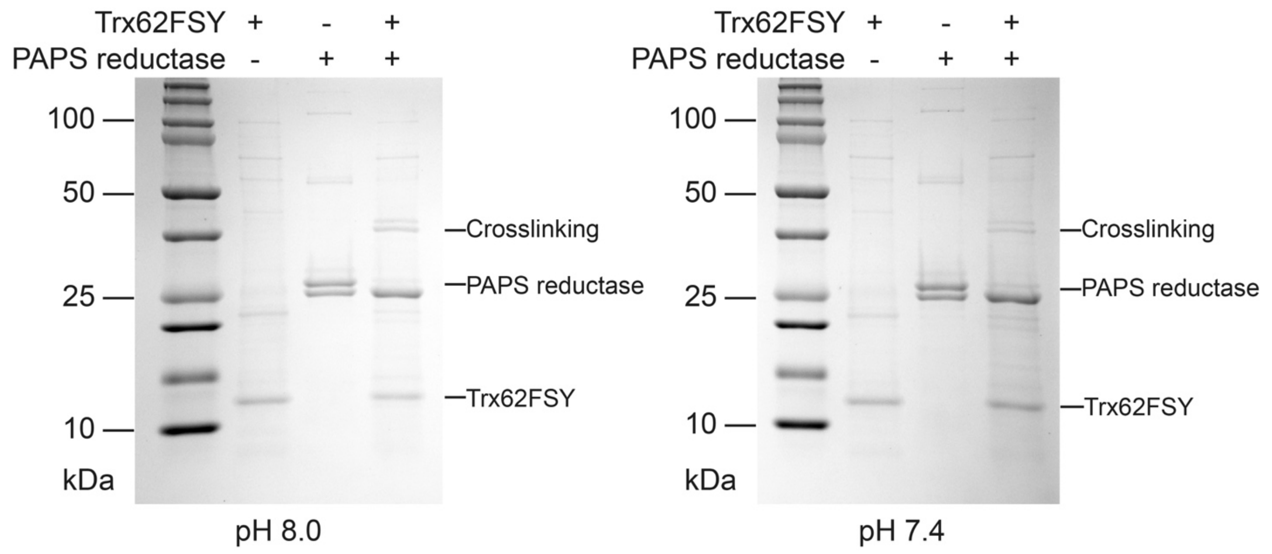


Figure S4. SDS-PAGE analysis of Trx62FSY crosslinking with PAPS reductase at pH 7.4 and 8.0.

Supplementary references

(1) (a) Chen, W.; Dong, J.; Li, S.; Liu, Y.; Wang, Y.; Yoon, L.; Wu, P.; Sharpless, K. B.; Kelly, J. W. *Angew. Chem. Int. Ed. Engl.* **2016,** 55, 1835-1838. (b) Dong, J.; Krasnova, L.; Finn, M. G.; Sharpless, K. B. *Angew. Chem. Int. Ed. Engl.* **2014,** 53, 9430-9448.

(2) Lacey, V. K.; Louie, G. V.; Noel, J. P.; Wang, L. *ChemBioChem* **2013,** 14, 2100-2105.

(3) Li, M. Z.; Elledge, S. J. *Nat. Methods* **2007,** 4, 251-256.

(4) Yang, B.; Tang, S.; Ma, C.; Li, S. T.; Shao, G. C.; Dang, B.; DeGrado, W. F.; Dong, M. Q.; Wang, P. G.; Ding, S.; Wang, L. *Nat. Communi.* **2017,** 8, 2240.

(5) Wang, W.; Takimoto, J. K.; Louie, G. V.; Baiga, T. J.; Noel, J. P.; Lee, K. F.; Slesinger, P. A.; Wang, L. *Nat. Neurosci.* **2007,** 10, 1063-1072.

(6) Coin, I.; Katritch, V.; Sun, T.; Xiang, Z.; Siu, F. Y.; Beyermann, M.; Stevens, R. C.; Wang, L. *Cell* **2013,** 155, 1258-1269.