Encoding Latent SuFEx Reactive meta-Fluorosulfate Tyrosine to Expand Covalent Bonding of Proteins

Paul C. Klauser,[1] Viktoriya Y. Berdan,[1] Li Cao,[1] Lei Wang[1]*

1. University of California San Francisco, Department of Pharmaceutical Chemistry and the Cardiovascular Research Institute, 555 Mission Bay Blvd. South, San Francisco, California 94158, United States

* E-mail: Lei.wang2@ucsf.edu

Table of Contents

Experimental Procedures:

Supplementary table 1: Primers used for cloning

Primer	Oligonucleotide Sequence (5'→3')
mFSRYS-BglII-F	CTAACAGGAGGAATTAGATCTATGGATAAAAAGCCT
mFSYRS-SalI-R	GATGATGATGATGATGGTCGACTTACAGGTTAGTAGAA
mFSYRS-NcoI-F	TATGCCATGGATAAAAAGCCTTTG
mFSYRS-NheI-R	CTATGCTAGCTTACAGGTTAGTAGA
2Rs15d-37TAG-For	TCTTGTGGTATGGGTTGGTAGCGTCAGAGCCCGGGT
2Rs15d-37TAG-Rev	CTACCAACCCATACCACAAGAGTTGAAGATATAGC
TrasFab-50TAG-For	ACTAGGCATCCTTTCTCTACTCTGGAGTCCCT
TrasFab-50TAG-Rev	GAGAAAGGATGCctaGTAAATCAGAAGCTTCGGAGCTTT
TrasFab-92TAG-For	TCAGCAACATTAGACCACACCAC
TrasFab-92TAG-Rev	CAGTAATAAGTTGCGAAGTCTTC
7D12-116TAG-For	ATGAATACGACTACTGGGGTTAGGGTACGCAGG
7D12-116TAG-Rev	AACCCCAGTAGTCGTATTCATACAGAGTGCCAT

Reagents and molecular biology

Primers were synthesized and purified by Integrated DNA Technologies (IDT), and plasmids were sequenced by GENEWIZ. All molecular biology reagents were either obtained from New England Biolabs or Vazyme. His-HRP antibody were obtained from ProteinTech Group. pBAD-EGFP, pBAD-dZ$_{HER2}$, and pBAD-7D12 were used as previously described. [1,2]

mFSYRS amino acid sequence

DKKPLNTLISATGLWMSRTGTIHKIKHHEVSRSKIYIEMACGDHLVVNNSRSSRTARALRHHKY
RKTCKRCRVSDEDLNKFLTKANEDQTSVKVKVVSAPTRTKKAMPKSVARAPKPLENTEAAQAQ
PSGSKFSPAIPVSTQESVSVPASVSTSISSISTGATASALVKGNTNPITSMSAPVQASAPALTKQ
TDRLEVLLNPKDEISLNSGKPFRELESELLSRRKKDLQQIYAEERENYLGKLEREITRFFVDRGF
LEIKSPILIPLEYIERMGIDNDTELSKQIFRVDKNFCLRPMLAPNMYNYLRKLDRALPDPIKTFEIG
PCYRKESDGKEHLEEFTMLGFCQMGSGCTRENLESIITDFLNHLGIDFKIVGDSCMVYGDTLDV
MHGDLELSSAVVGPIPLDREWGIDKPWIGAGFGLERLLKVKHDFKNIKRAARSESYYNGISTNL

Red: mutated residues.

pBAD-dZ$_{HER2}$-D36/D37TAG

MAVDNKFNKEMRNAYWEIALLPNLNNQQKRAFIRSLYDDPSQSANLLAEAKKLNDAQAPKVEV
DNKFNKEMRNAYWEIALLPNLNNQQKRAFIRSLYDDPSQSANLLAEAKKLNDAQAPKHHHHHH

Red: amber codon TAG at 36th/37th position.

pBAD-2Rs15d-Y37TAG

MKYLLPTAAAGLLLLAAQPAMAMGQVQLQESGGGSVQAGGSLKLTCAASGYIFNSCGMGWY
RQSPGRERELVSRISGDGDTWHKESVKGRFTISQDNVKKTLYLQMNSLKPEDTAVYFCAVCYN
LETYWGQGTQVTVSSHHHHHH

Red: amber codon TAG at 37th position.

pBR322-TrasFab-S50/Y92TAG

Light Chain
MKSLLPTAAAGLLLLAAQPAMASDIQMTQSPSSLSASVGDRVTITCRASQDVNTAVAWYQQKP
GKAPKLLIYSASFLYSGVPSRFSGSRSGTDFTLTISSLQPEDFATYYCQQHYTTPPTFGQGTKV
EIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQD
SKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC

Heavy Chain
MKKNIAFLLASMFVFSIATNAYAEISEVQLVESGGGLVQPGGSLRLSCAASGFNIKDTYIHWVRQ
APGKGLEWVARIYPTNGYTRYADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCSRWGG
DGFYALDYWGQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNS
GALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTH
TGGSGSAGGLNDIFEAQKIEWHE

Red: amber codon TAG at 50th/92nd position of the Light Chain.

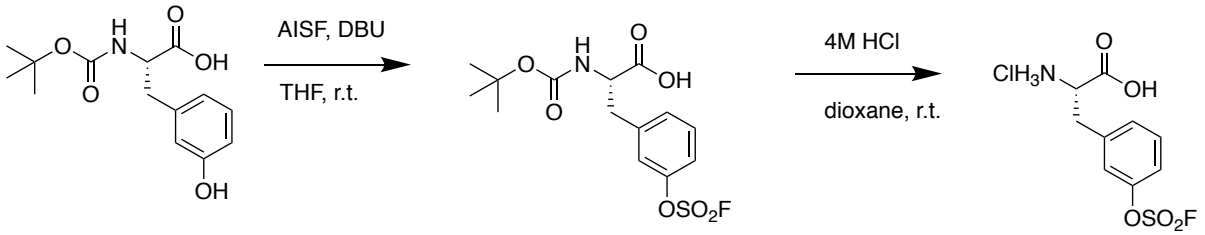
pBAD-7D12-Q116TAG

MKYLLPTAAAGLLLLAAQPAMAMGQVKLEESGGGSVQTGGSLRLTCAASGRTSRSYGMGW
FRQAPGKEREFVSGISWRGDSTGYADSVKGRFTISRDNAKNTVDLQMNSLKPEDTAIYYC
AAAAGSAWYGTLYEYDYWGQGTQVTVSSHHHHHH

Red: amber codon TAG at 116th position.

Chemical Synthesis of mFSY

Synthesis of aryl fluorosulfates was based on recent methods to synthesize sulfur (IV) fluorides using [4-(acetylamino)phenyl]imidodisulfuryl difluoride (AISF) reagent.[3]



1 **2** **3**

*Synthesis of (S)-2-amino-3-(3-((fluorosulfonyl)oxy)phenyl)propanoic acid (**3,** mFSY).* To a 100 mL round-bottom flask were added Boc-(S)-2-Amino-3-(3-hydroxyphenyl)propionic acid (1, 1.15 g, 4.09 mmol) and [4-(acetylamino)phenyl]imidodisulfuryl difluoride (AISF) reagent (1.54 g, 4.90 mmol, 1.2 equiv.). The mixture was dissolved in 25 mL anhydrous tetrahydrofuran and 1,8-diazabicyclo[5.4.0]undec-7-ene (1.37 mL, 9 mmol, 2.2 equiv.) was added dropwise while stirring. The solution was then stirred at r.t. for 30 minutes. The reaction was then diluted with 50 mL ethyl acetate and washed with 1 M HCl (100 mL x 2) and brine (100 mL x 1). The organic fraction was dried with anhydrous sodium sulfate and concentrated under vacuum. The crude product was then purified by column chromatography using MeOH:CH_2Cl_2 (1:200). The product, (*S*)-2-((*tert*-butoxycarbonyl)amino)-3-(3-((fluorosulfonyl)oxy)phenyl)propanoic acid, was isolated as a white solid (2, 0.774 g, 2.13 mmol, 52%).

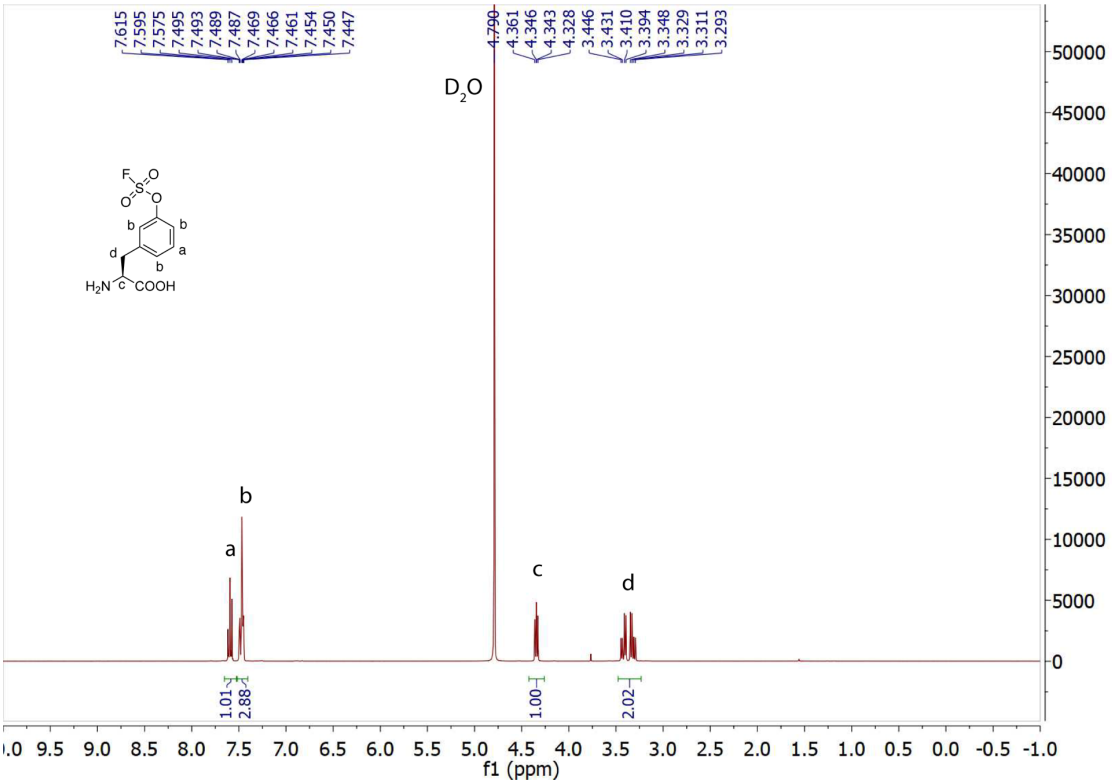
S)-2-((*tert*-butoxycarbonyl)amino)-3-(3-((fluorosulfonyl)oxy)phenyl)propanoic acid (2, 0.774 g , 2.13 mmol) was added to a scintillation vial and dissolved in 4 M HCl in dioxane (10 mL). The reaction was stirred overnight. The resultant solid was filtered off and washed with cool ether (10 mL x 2) affording the product mFSY-HCl as a white solid (3, 554 mg, 1.85 mmol, 87%).

^{1}H NMR (400 MHz, D_2O): δ (ppm) 7.62-7.58 (t, J= 16.0 Hz, 1H), 7.50-7.45 (m, J= 19.2 Hz, 3H), 4.36-4.33 (t, J= 13.2, 1H), 3.45-3.29 (m, J= 61.2 Hz, 2H)

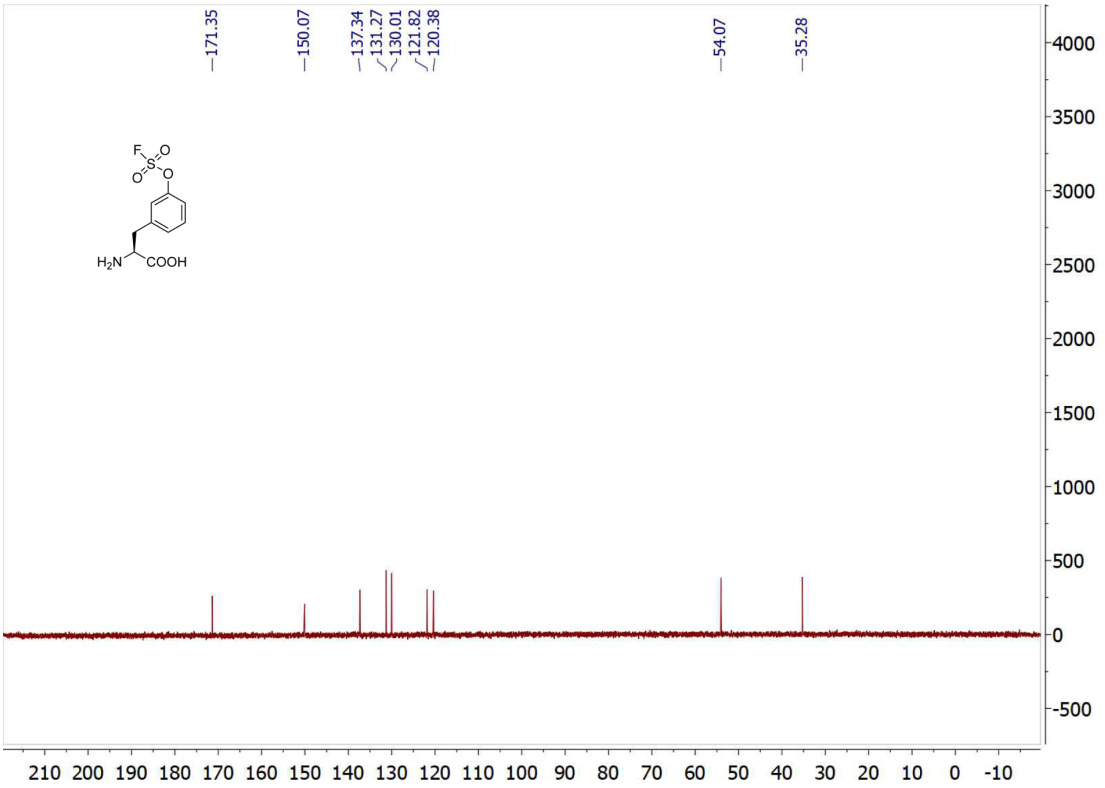
^{13}C NMR (400 MHz, D_2O): δ (ppm) 171.4, 150.1, 137.3, 131.3, 130.0, 121.8, 120.4, 54.1, 35.3

HR-ESI (+)*m/z*: calculated for $C_9H_{10}FNO_5S$ $[M+H]^+$, 264.0264; found 264.0351.

^{1}H NMR



^{13}C NMR



Library construction and mFSYRS mutant 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: 302NYT, 305WTG, 306WTG/TAC, 309KYA, 322AYA, 346NDT/VMA/ATG/TGG, 348NDT/VMA/ATG/TGG, 384TTM/TAT, 401VTT, 417NDT/VMA/ATG/TGG.[4] The selection was performed as previously described.[5]

Briefly, pBK-TK3 library was transformed into DH10b-pRep positive selection reporter cells via electroporation. The cells were then plated onto an LB-agar selection plate containing 1 mM mFSY, 12.5 μg/mL of tetracycline (Tet), 25 μg/mL of kanamycin (Kan), and 75 μg/mL of chloramphenicol (Cm). The selection plate was incubated at 37 °C for 72 h and then stored at 4 °C. Colonies showing green fluorescence was picked and streaked on a fresh LB-agar plate containing either Tet12.5Kan25Cm100 or Tet12.5Kan25Cm100 +1mM mFSY. After 24 h of incubation at 37 °C, 2 clones present mFSY-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.

Incorporation of mFSY into EGFP(182TAG)

pBAD-EGFP(182TAG) was co-transformed with pEVOL-mFSYRS into DH10b and plated on LB agar plate supplemented with 100 μg/mL ampicillin and 34 μg/mL chloramphenicol. A single colony was picked and inoculated into 1 mL 2xYT (5 g/L NaCl, 16 g/L Tryptone, 10 g/L Yeast extract). The cells were left grown at 37 °C, 220 rpm, for 16 h. The cells were then diluted to an OD_{600} of 0.6 in fresh 2XYT supplemented with relevant antibiotics, with or without 1 mM mFSY. The cells were then induced with 0.2% arabinose at either 30 °C for 6 h. The fluorescence intensity was measured with a plate reader (excitation at 485 nm, emission at 528 nm) and normalized to OD at 600 nm.

General incorporation of mFSY into proteins for expression and purification

For the incorporation of mFSY into dZ_{HER2}, 2Rs15d, TrasFab, and 7D12 the procedure of transformation was the same as described above. After transformation, a single colony was picked and left grown at 37 °C, 220 rpm for 16 h. Next morning, the cell culture was diluted 100 times and then regrown to an OD 0.6-0.8 in 100 mL scale, with good aeration and the relevant antibiotic selection. Then the medium was added with 0.2% arabinose (and 1 mM IPTG for TrasFab) with or without 1 mM mFSY, and the expression were carried out at 18 °C, 220 rpm for 18 hr, 18 °C, or 25 °C. The IMAC chromatography was used for protein purification and the procedure was done as described elsewhere.[1] TrasFab was purified on the Äkta Pure FPLC protein purification system using an HiTrap® Protein A column. Procedure was described previously.[6]

***In vitro* cross-linking of dZ_{HER2}, 2Rs15d, TrasFab, and 7D12 with HER2 and EGFR**

Recombinant extra-cellular domain (ECD) of HER2 receptor was purchased from Abcam (Cat# ab168896); EGFR-ECD receptor was purchased from Abcam (Cat# ab155726). The *in vitro* protein cross-linking was performed as previously described.[2, 7] Briefly, purified 1 μM dZ_{HER2} or TrasFab was incubated with 1 μM HER2 ECD in 20 uL buffer at 37 °C for 16 h. Purified 1 μM 7D12 was incubated with 1 μM EGFR in 20 uL buffer at 37 °C for 16 h. Purified 5 μM 2Rs15d was incubated with 5 μM HER2 ECD in 20 μL buffer at 37 °C for 24 h. For the competition assay, 1 μM dZ_{HER2} (37mFSY) and 1 μM HER2 ECD were incubated with different concentrations of dZ_{HER2} (WT) at 37 °C for 4 h. After incubation, 4x Laemmli Sample Buffer (Bio Rad, Cat# 161-0747) was added into the incubation and heated at 95 °C for 10 min. The samples were separated on SDS-

PAGE and either analyzed by Coomassie blue staining or immunoblotted with 1:10000 anti-his monoclonal antibody (Proteintech #HRP66005).

Mass spectrometry

Mass spectrometric measurements were performed as previously described.[8] Briefly for electrospray ionization mass spectrometry, mass spectra of intact proteins were obtained using a QTOF Ultima (Waters) mass spectrometer, operating under positive electrospray ionization (+ESI) mode, connected to an LC-20AD (Shimadzu) liquid chromatography unit. Protein samples were separated from small molecules by reverse phase chromatography on a Waters Xbridge BEH C4 column (300 Å, 3.5 µm, 2.1 mm x 50 mm), using an acetonitrile gradient from 30-71.4%, with 0.1% formic acid. Each analysis was 25 min under constant flow rate of 0.2 mL/min at RT. Data were acquired from m/z 350 to 2500, at a rate of 1 sec/scan. Alternatively, spectra were acquired by Xevo G2-S QTOF on a Waters ACQUITY UPLC Protein BEH C4 reverse-phase column (300 Å, 1.7 µm, 2.1 mm x 150 mm). An acetonitrile gradient from 5%-95% was used with 0.1% formic acid, over a run time of 5 min and constant flow rate of 0.5 mL/min at RT. Spectrum were acquired from m/z 350 to 2000, at a rate of 1 sec/scan. The spectra were deconvoluted using maximum entropy in MassLynx.

For tandem mass spectrometry, analysis and sequencing of peptides were carried out using a Q Exactive Orbitrap interfaced with Ultimate 3000 LC system. Data acquisition by Q Exactive Orbitrap was as follows: 10 µL of trypsin-digested protein was loaded on an Ace UltraCore super C18 reverse-phase column (300 Å, 2.5 µm, 75 mm × 2.1 mm) via an autosampler. An acetonitrile gradient from 5%-95% was used with 0.1% formic acid, over a run time of 45 min and constant flow rate of 0.2 mL/min at RT. MS data were acquired using a data-dependent top10 method dynamically choosing the most abundant precursor ions from the survey scan for HCD fragmentation using a stepped normalized collision energy of 28, 30 35 eV. Survey scans were acquired at a resolution of 70,000 at m/z 200 on the Q Exactive. Tandem MS data was analyze on MaxQuant.

Flow cytometric analysis of mFSY incorporation into HeLa-GFP(182TAG)

One day before transfection, 4.5×10^4 HeLa-GFP(182TAG) reporter cells were seeded in 9 wells of a Greiner bio-one 12 well-cell culture dish containing 1 mL of DMEM media with 10% FBS, and incubated at 37 °C in a CO_2 incubator. Plasmid pMP-mFSYRS (1 µg) was transfected into target cells using 9 µL polyethylenimine (PEI) transfection agent. pMP-mFSYRS plasmid was not added to three of the wells (negative control). Six hours post transfection, 1 mM mFSY was added to three wells. The remaining three wells were transfected with pMP-mFSYRS plasmid but did not have mFSY Uaa added. After incubation at 37 °C for 48 h, cells were non-enzymatically detached from the plates using Gibco Cell Dissociation Buffer and collected by centrifugation (500 g, 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-GFP(182mFSY)

One day before transfection, 4.5×10^4 HeLa-GFP(182TAG) cells were seeded in a Greiner bioone CELLview glass bottom dish containing 2 mL of DMEM media with 10% FBS, and incubated at 37 °C in a CO2 incubator. Plasmid pMP-mFSYRS (2 µg) was transfected into the HeLa-GFP(182TAG) cells using 9 µL polyethylenimine (PEI) transfection agent. Six hours post transfection, 1 mM mFSY was added to the media. A Hela-GFP(182TAG) cell group that was not transfected with any plasmid was used as a negative control. The cells were incubated at 37 °C for an additional 48 h post transfection and imaged with a Nikon CSU-X1 Spinning Disk microscope.

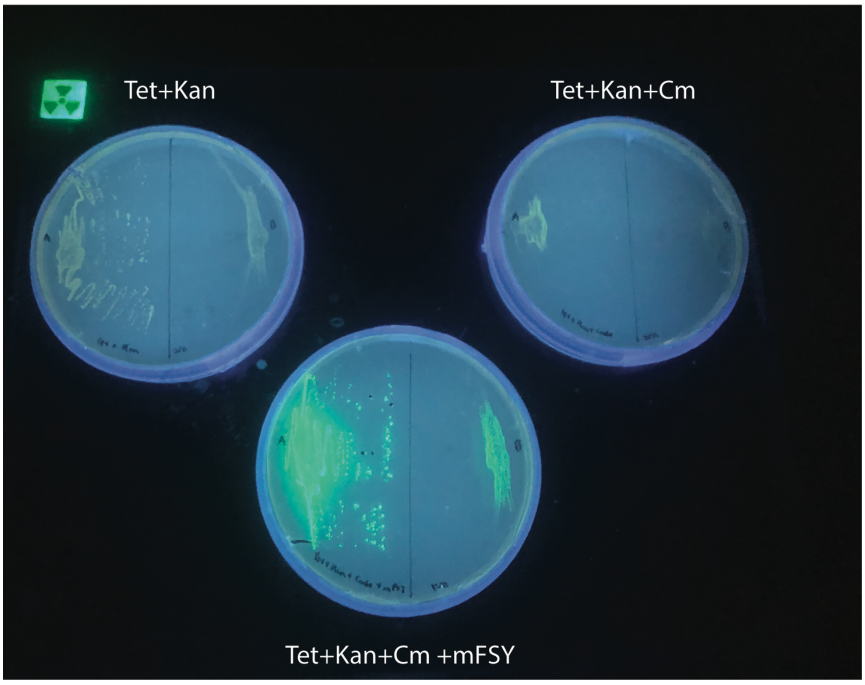


Figure S1. Selection plates for mFSY-specific synthetase. Addition of mFSY to plates showed robust incorporation of the Uaa into EGFP via the orthogonal $tRNA^{Pyl}$/mFSYRS pair, rendering cells green fluorescent. When mFSY was not added to plates, there was insignificant EGFP fluorescence detected, suggesting negligible misincorporation of native amino acids.

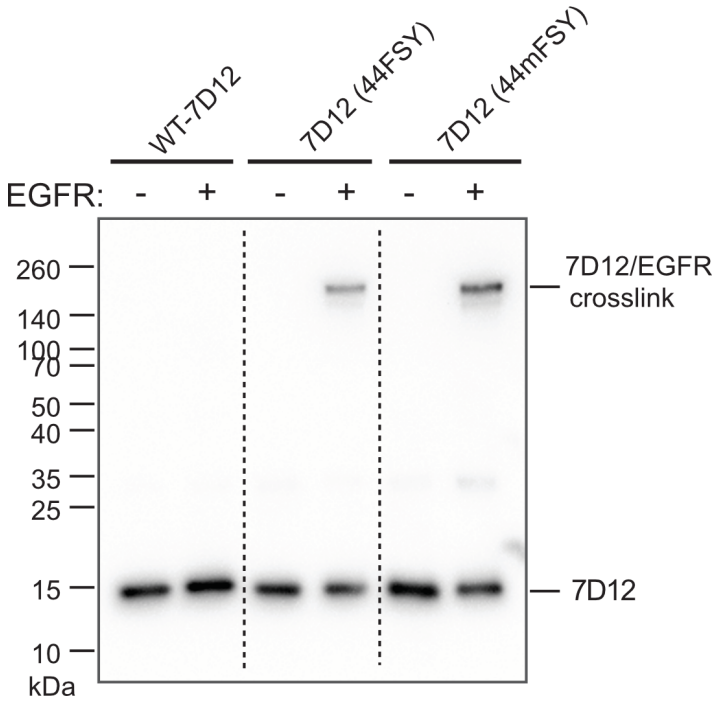


Figure S2. Western blot analysis of WT-7D12, 7D12(44FSY), and 7D12(44mFSY) incubation with the EGFR receptor. WT-7D12 did not crosslink EGFR, while 7D12(44mFSY) crosslinked EGFR in a higher efficiency than 7D12(44FSY) as also observed in Figure 3b.

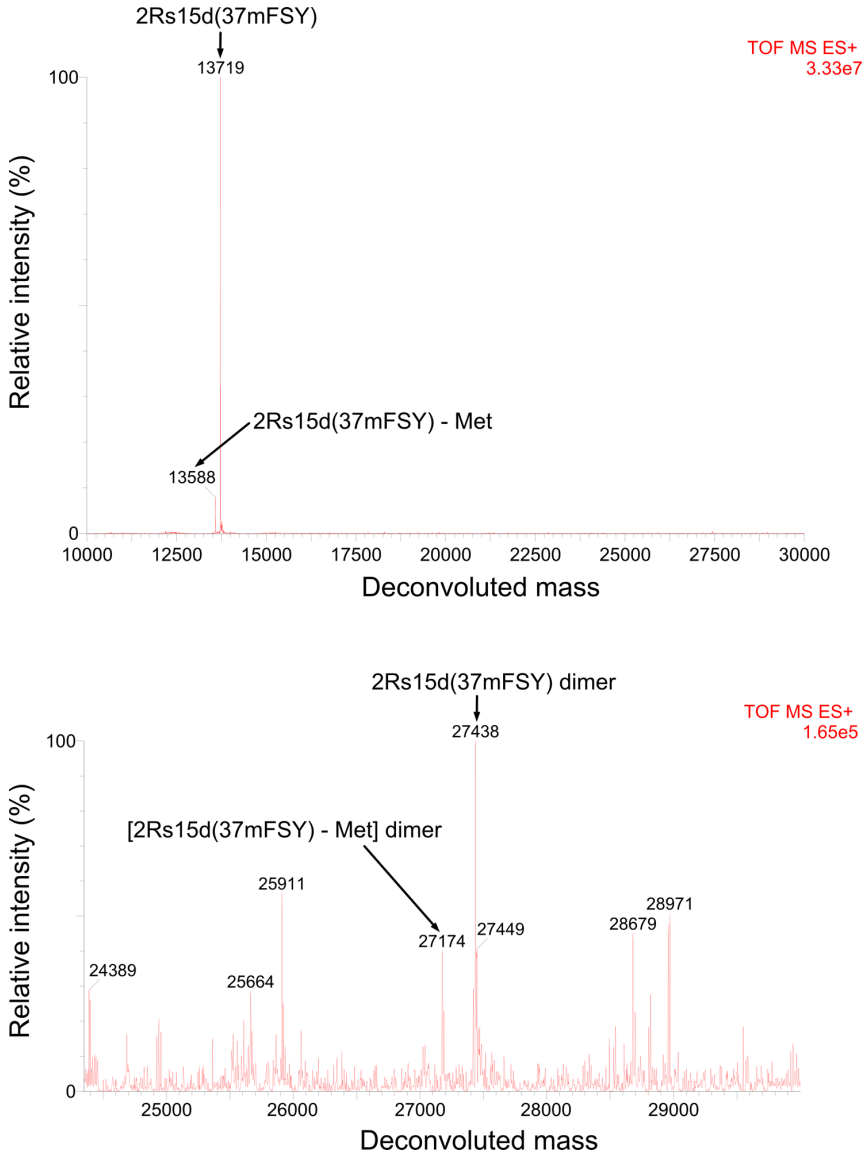


Figure S3. Mass spectra for intact 2Rs15d(37mFSY). Both monomeric and non-cross-linked dimeric 2Rs15d(37mFSY) were detected, together with their –Met (loss of N-terminal Met) species. Monomeric 2Rs15d(37mFSY): expected 13719 Da, observed 13719 Da. Monomeric [2Rs15d(37mFSY) – Met]: expected 13588 Da, observed 13588. Non-cross-linked 2Rs15d(37mFSY) dimer: expected 27438 Da, observed 27438 Da. Non-cross-linked [2Rs15d(37mFSY) – Met] dimer: expected 27176 Da, observed 27174 Da.

Supplementary Reference

(1) Li, Q.; Chen, Q.; Klauser, P. C.; Li, M.; Zheng, F.; Wang, N.; Li, X.; Zhang, Q.; Fu, X.; Wang, Q.; Xu, Y.; Wang, L. Developing Covalent Protein Drugs via Proximity-Enabled Reactive Therapeutics. *Cell* 2020, *182* (1), 85-97.e16. https://doi.org/10.1016/j.cell.2020.05.028.

(2) Liu, J.; Cao, L.; Klauser, P. C.; Cheng, R.; Berdan, V. Y.; Sun, W.; Wang, N.; Ghelichkhani, F.; Yu, B.; Rozovsky, S.; Wang, L. A Genetically Encoded Fluorosulfonyloxybenzoyl-l-lysine for Expansive Covalent Bonding of Proteins via SuFEx Chemistry. *J Am Chem Soc* 2021, *143* (27), 10341–10351. https://doi.org/10.1021/jacs.1c04259.

(3) Zhou, H.; Mukherjee, P.; Liu, R.; Evrard, E.; Wang, D.; Humphrey, J. M.; Butler, T. W.; Hoth, L. R.; Sperry, J. B.; Sakata, S. K.; Helal, C. J.; Ende, C. W. am. Introduction of a Crystalline, Shelf-Stable Reagent for the Synthesis of Sulfur(VI) Fluorides. *Org Lett* 2018, *20* (3), 812–815. https://doi.org/10.1021/acs.orglett.7b03950.

(4) Lacey, V. K.; Louie, G. V.; Noel, J. P.; Wang, L. Expanding the Library and Substrate Diversity of the Pyrrolysyl‐tRNA Synthetase to Incorporate Unnatural Amino Acids Containing Conjugated Rings. *Chembiochem* 2013, *14* (16), 2100–2105. https://doi.org/10.1002/cbic.201300400.

(5) Wang, N.; Yang, B.; Fu, C.; Zhu, H.; Zheng, F.; Kobayashi, T.; Liu, J.; Li, S.; Ma, C.; Wang, P. G.; Wang, Q.; Wang, L. Genetically Encoding Fluorosulfate- l -Tyrosine To React with Lysine, Histidine, and Tyrosine via SuFEx in Proteins in Vivo. *J Am Chem Soc* 2018, *140* (15), 4995–4999. https://doi.org/10.1021/jacs.8b01087.

(6) Hornsby, M.; Paduch, M.; Miersch, S.; Sääf, A.; Matsuguchi, T.; Lee, B.; Wypisniak, K.; Doak, A.; King, D.; Usatyuk, S.; Perry, K.; Lu, V.; Thomas, W.; Luke, J.; Goodman, J.; Hoey, R. J.; Lai, D.; Griffin, C.; Li, Z.; Vizeacoumar, F. J.; Dong, D.; Campbell, E.; Anderson, S.; Zhong, N.; Gräslund, S.; Koide, S.; Moffat, J.; Sidhu, S.; Kossiakoff, A.; Wells, J. A High Through-Put Platform for Recombinant Antibodies to Folded Proteins*. *Mol Cell Proteom Mcp* 2015, *14* (10), 2833–2847. https://doi.org/10.1074/mcp.o115.052209.

(7) Kobayashi, T.; Hoppmann, C.; Yang, B.; Wang, L. Using Protein-Confined Proximity to Determine Chemical Reactivity. *J Am Chem Soc*. 2016 Oct 31; 138(45):14832-14835. https://doi.org/10.1021/jacs.6b08656.

(8) Liu, J.; Chen, Q.; Rozovsky, S. Utilizing Selenocysteine for Expressed Protein Ligation and Bioconjugations. *J Am Chem Soc* 2017, *139* (9), 3430–3437. https://doi.org/10.1021/jacs.6b10991.