Supplementary Information for

Genetic Encoding of 2-Aryl-5-carboxytetrazole-Based Protein Photo-cross-linkers

Yulin Tian and Qing Lin*

Department of Chemistry, State University of New York at Buffalo, Buffalo, New York 14260-3000, United States

E-mail: qinglin@buffalo.edu

Table of Contents

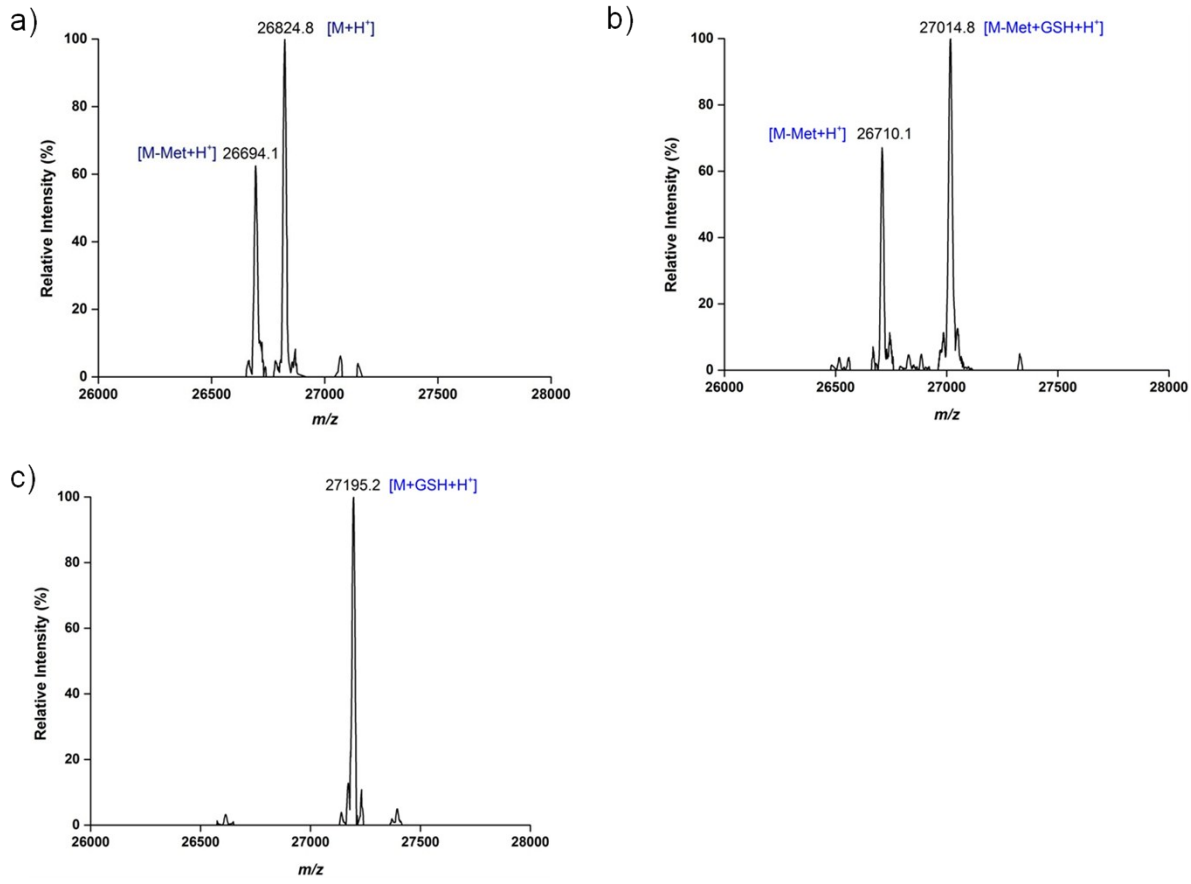


Figure S1. (a) Deconvoluted mass spectrum of GST-E52mPyTOK: calcd, 26693.2 Da [M − Met + H^+], found, 26694.1 ± 4.7 Da; calcd, 26824.2 Da [M − Met + H^+], found, 26824.8 ± 3.5 Da. (b) Deconvoluted mass spectrum of GST-E52mPyTSK: calcd, 26709.2 Da [M − Met + H^+], found, 26710.1 ± 3.0 Da; calcd, 27016.2 Da [M − Met + GSH + H^+], found, 27014.8 ± 2.9 Da. (c) Deconvoluted mass spectrum of GST-E52mPyTSeK: calcd, 27194.2 Da [M + GSH + H^+], found, 29195.2 ± 5.3 Da.

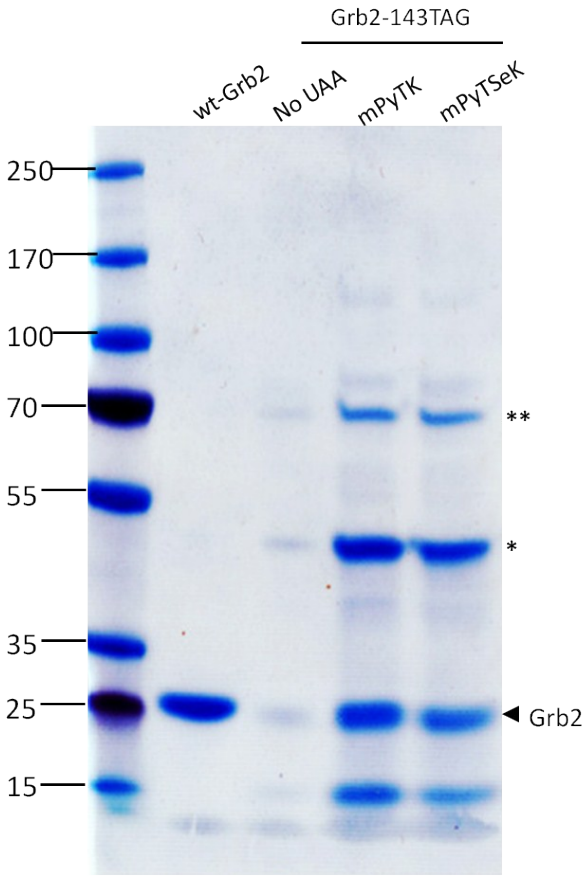


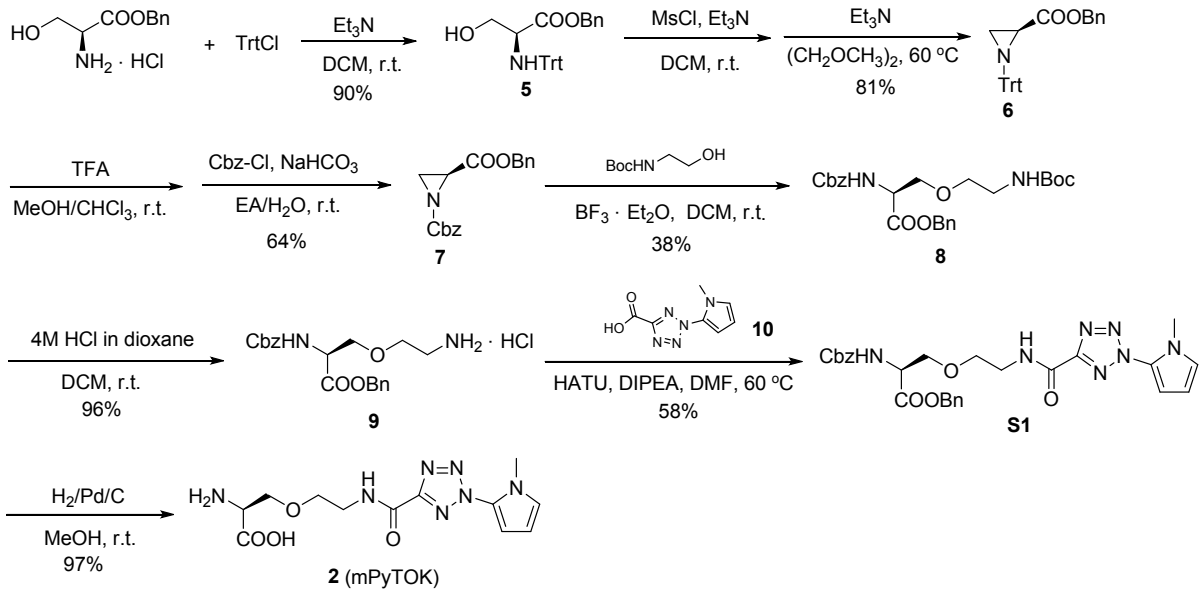
Figure S2. Coomassie blue stained SDS-PAGE gel of wt-Grb2 and Grb2-N143 mutants encoding mPyTK and mPyTSeK. Asterisk indicates impurities derived from Ni-NTA affinity purification.

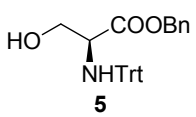
General Information

Solvents and chemicals were purchased from commercial sources and used directly without further purification. Flash chromatography was performed with SiliCycle P60 silica gel (40-63 μm, 60 Å). ^{1}H and ^{13}C NMR spectra were recorded with Varian Mercury-300, Inova-400 or -500 MHz spectrometer. Chemical shifts were reported in ppm using either TMS or deuterated solvents as internal standards (TMS, 0.00; $CDCl_3$, 7.26; CD_3OD, 3.31; DMSO-d_6, 2.50). Multiplicity was reported as follows: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, brs = broad. ^{13}C NMR spectra were recorded at 126 MHz, and chemical shifts were reported in ppm using deuterated solvents as internal standards ($CDCl_3$, 77.0; DMSO-d_6, 39.5; CD_3OD, 49.05). Electrospray LC-MS analysis was performed using a Finnigan LCQ Advantage IonTrap mass spectrometry coupled with a Surveyor HPLC system. Protein liquid chromatography was performed using a Phenomenex Aeris C4 column (3.6 μm, 200 Å, 2.10 × 100 mm) with a flow rate of 250 μL/min and a linear gradient of 5-50% ACN/H_2O containing 0.1% HCOOH for 10 min or 5-95% ACN/H_2O containing 0.1% HCOOH for 15 min. The intact protein masses were derived through deconvoluting charge ladders using ProMass software (Thermo Scientific). High resolution mass spectrometry was performed on a Bruker solariX XR Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR-MS).

Experimental Procedures and Characterization Data

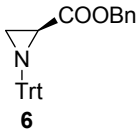
Scheme S1. Synthesis of mPyTOK[S1]





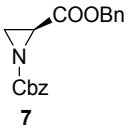
Benzyl 3-hydroxy-2-(tritylamino)propanoate (5): To a solution of benzyl 2-amino-3-hydroxypropanoate hydrochloride (4.1 g, 14.2 mmol) and Et_3N (5.7 g, 56.8 mmol) in dichloromethane (20 mL) was added trityl chloride (4 g, 14.2 mmol) at 0 °C. The mixture was stirred at 0 °C for 24 h, then washed with 10% citric acid, brine, dried over anhydrous Na_2SO_4, filtered and concentrated to afford the titled compound as colorless syrup (5.6 g, 90% yield): ^{1}H NMR

(500 MHz, $CDCl_3$) δ 7.50 – 7.17 (m, 20H), 4.77 (d, J = 12.3 Hz, 1H), 4.63 (d, J = 12.3 Hz, 1H), 3.70 (dd, J = 10.5, 4.0 Hz, 1H), 3.61 – 3.50 (m, 2H); ESI-MS calcd for $C_{29}H_{27}NO_3Na$ 460.2 [M+Na^+], found 460.1.



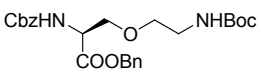
6

Benzyl 1-tritylaziridine-2-carboxylate (6): To a solution of **5** (3.7 g, 8.5 mmol) and MsCl (1.1 g, 9.3 mmol) in dichloromethane (20 mL) was added Et_3N (1.3 g, 12.8 mmol) dropwise at 0 °C. The mixture was stirred at 0 °C for 30 min, then washed with 10% citric acid, brine, dried over anhydrous Na_2SO_4, filtered and concentrated. The residue was suspended in $(CH_2OCH_3)_2$ (20 mL), and Et_3N (3.4 g, 34 mmol) was added. The mixture was stirred at 60 °C for 24 h, and then concentrated. The residue was purified by silica gel flash column chromatography (hexanes/EtOAc = 3:1) to afford the titled compound as a white solid (2.9 g, 81% yield): ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.51 – 7.19 (m, 20H), 5.27 (d, J = 12.3 Hz, 1H), 5.22 (d, J = 12.3 Hz, 1H), 2.33 (dd, J = 2.7, 1.3 Hz, 1H), 1.97 (dd, J = 6.2, 2.8 Hz, 1H), 1.45 (dd, J = 6.2, 1.2 Hz, 1H); ESI-MS calcd for $C_{29}H_{25}NO_2Na$ 442.2 [M+Na^+], found 442.1.



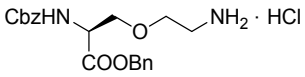
7

Dibenzyl aziridine-1,2-dicarboxylate (7): To a solution of **6** (2.8 g, 6.7 mmol) in MeOH/$CHCl_3$ (20 mL, 1:1) was added TFA (4 mL) at 0 °C. The mixture was stirred at room temperature for 3 h, and then concentrated. The residue was suspended in Et_2O (20 mL), washed with H_2O (10 mL × 4). The combined aqueous layers were added $NaHCO_3$ until pH > 8, then CbzCl (1.3 g, 7.3 mmol) in EtOAc (10 mL) was added. The mixture was stirred at room temperature for 12 h, and then extracted with EtOAc (10 mL × 4). The combined organic layers were washed with H_2O, brine, dried over anhydrous Na_2SO_4, filtered and concentrated. The residue was purified by silica gel flash column chromatography (hexane/EtOAc = 4:1) to afford the titled compound as a colorless oil (1.3 g, 64% yield): ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.40 – 7.29 (m, 10H), 5.17 – 5.05 (m, 4H), 3.16 – 3.12 (m, 1H), 2.64 – 2.60 (m, 1H), 2.50 – 2.46 (m, 1H); ^{13}C NMR (126 MHz, $CDCl_3$) δ 168.1, 160.7, 135.4, 134.9, 128.7, 128.6, 128.6, 128.5, 128.5, 128.4, 68.6, 67.6, 35.0, 31.5; ESI-MS calcd for $C_{18}H_{17}NO_4Na$ 334.1 [M+Na^+], found 334.2.



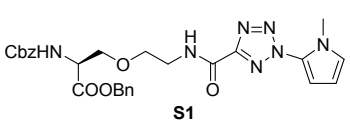
8

Benzyl 2-(((benzyloxy)carbonyl)amino)-3-(2-((tert-butoxy carbonyl)amino)ethoxy)propanoate (8): To a solution of **7** (375 mg, 1.2 mmol) and *tert*-butyl (2-hydroxyethyl)carbamate (291 mg, 1.8 mmol) in dichloromethane (5 mL) was added $BF_3{\cdot}Et_2O$ (0.1 mL, 0.81 mmol) at 0 °C. The mixture was stirred at room temperature for 3 h, washed with aqueous $NaHCO_3$ solution, brine, dried over anhydrous Na_2SO_4, filtered, and concentrated. The residue was purified by silica gel flash column chromatography (DCM/EtOAc = 10:1) to afford the titled compound as a colorless oil (215 mg, 38% yield): ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.40 – 7.28 (m, 10H), 5.71 (d, J = 7.6 Hz, 1H), 5.25 (d, J = 12.2 Hz, 1H), 5.17 (d, J = 12.1 Hz, 1H), 5.12 (s, 2H), 4.76 (s, 1H), 4.57 – 4.51 (m, 1H), 3.87 (d, J = 7.0 Hz, 1H), 3.71 (d, J = 7.3 Hz, 1H), 3.48 – 3.10 (m, 4H), 1.42 (s, 9H); ^{13}C NMR (126 MHz, $CDCl_3$) δ 170.1, 156.0, 155.9, 136.1, 135.3, 128.6, 128.5, 128.5, 128.2, 128.2, 128.1, 70.8, 70.7, 67.3, 67.1, 60.7, 54.5, 40.1, 28.4; ESI-MS calcd for $C_{25}H_{32}N_2O_7Na$ 495.2 [M+Na^+], found 495.3.



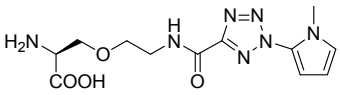
9

Benzyl 3-(2-aminoethoxy)-2-(((benzyloxy)carbonyl) amino) propanoate hydrochloride (9): To a solution of **8** (108 mg, 0.23 mmol) in dichloromethane (5 mL) was added 4 N HCl in dioxane (1 mL). The mixture was stirred at room temperature for 12 h, and then concentrated to afford the titled compound as colorless oil (90 mg, 96% yield): ^{1}H NMR (500 MHz, CD_3OD) δ 7.40 – 7.22 (m, 10H), 5.23 (d, J = 12.3 Hz, 1H), 5.16 (d, J = 12.3 Hz, 1H), 5.10 (s, 2H), 4.54 – 4.48 (m, 1H), 3.89 (dd, J = 9.5, 4.4 Hz, 1H), 3.79 (dd, J = 9.8, 3.1 Hz, 1H), 3.63 – 3.57 (m, 2H), 3.10 – 2.95 (m, 2H); ^{13}C NMR (126 MHz, CD_3OD) δ 170.4, 157.3, 136.6, 135.7, 128.2, 128.1, 128.0, 127.9, 127.7, 127.5, 72.1, 70.3, 60.8, 57.5, 54.4, 41.4; ESI-MS calcd for $C_{20}H_{25}N_2O_5$ 373.2 [M+H$^+$], found 373.2.



S1

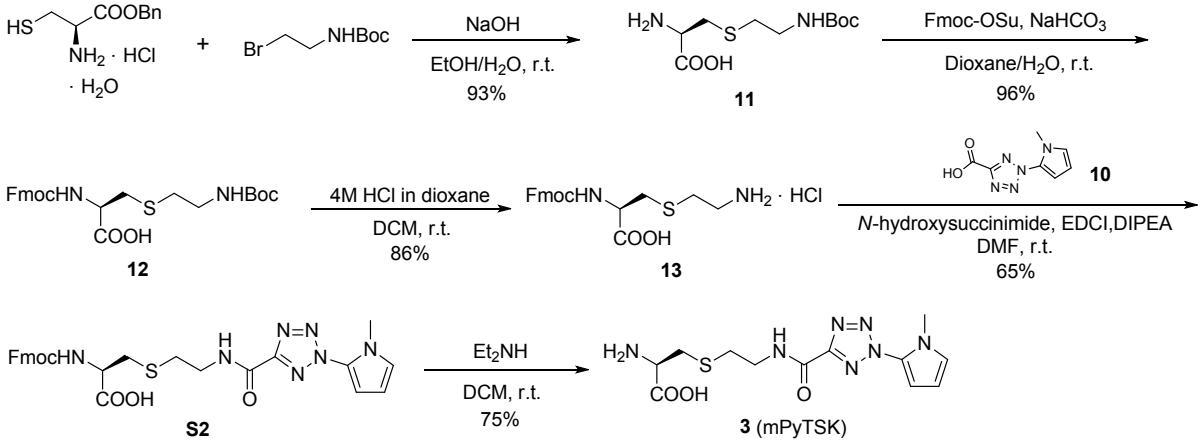
Benzyl 2-(((benzyloxy)carbonyl)amino)-3-(2-(2-(1- methyl-1*H*-pyrrol-2-yl)-2*H*-tetrazole-5-carboxamido)ethoxy)propanoate (S1): To a solution of **10**[S2] (43 mg, 0.22 mmol) in DMF (3 mL) was added **9** (90 mg, 0.22 mmol), HATU (100 mg, 0.26 mmol) and DIPEA (85 mg, 0.66 mmol). The mixture was stirred at 60 °C for 4 h. Then DMF was removed under reduced pressure. The residue was dissolved in EtOAc, washed with brine, dried over anhydrous Na_2SO_4, filtered, and concentrated. The residue was purified by silica gel flash column chromatography (hexanes/EtOAc = 1:1) to afford the titled compound as a yellow oil (70 mg, 58% yield): ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.45 (s, 1H), 7.36 – 7.27 (m, 10H), 6.70 (s, 1H), 6.62 (s, 1H), 6.21 (s, 1H), 5.74 (d, J = 8.0 Hz, 1H), 5.26 (d, J = 12.3 Hz, 1H), 5.18 (d, J = 12.1 Hz, 1H), 5.11 (s, 2H), 4.61 – 4.55 (d, J = 7.6 Hz, 1H), 3.93 (d, J = 7.4 Hz, 1H), 3.78 (d, J = 8.3 Hz, 1H), 3.68 (s, 3H), 3.66 – 3.51 (m, 4H); ^{13}C NMR (126 MHz, $CDCl_3$) δ 170.0, 159.2, 156.5, 156.0, 136.1, 135.2, 128.6, 128.5, 128.5, 128.2, 128.2, 128.0, 124.9, 123.8, 107.8, 106.2, 71.0, 69.7, 67.4, 67.1, 54.5, 39.2, 35.0; ESI-MS calcd for $C_{27}H_{29}N_7O_6Na$ 570.2 [M+Na$^+$], found 570.1.

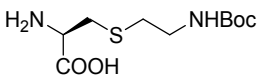


2 (mPyTOK)

2-Amino-3-(2-(2-(1-methyl-1*H*-pyrrol-2-yl)-2*H*-tetrazole-5-carboxamido)ethoxy)propanoic acid (2, mPyTOK): To a solution of **S1** (54 mg, 0.1 mmol) in MeOH (3 mL) was added Pd/C (6 mg, 10%). The mixture was filled with hydrogen and stirred at room temperature for 4 h. Pd/C was removed by filtered through celite. The filtrate was concentrated to afford the titled compound as a yellow solid (31 mg, 97% yield): ^{1}H NMR (500 MHz, CD_3OD) δ 6.91 – 6.86 (m, 1H), 6.61 – 6.55 (m, 1H), 6.25 – 6.19 (m, 1H), 3.95 – 3.89 (m, 1H), 3.88 – 3.82 (m, 1H), 3.82 – 3.77 (m, 1H), 3.73 – 3.65 (m, 7H); ^{13}C NMR (126 MHz, CD_3OD) δ 159.3, 157.8, 128.0, 124.7, 123.9, 107.1, 105.7, 69.4, 68.9, 54.88, 38.9, 33.4; HRMS (EI) calcd for $C_{12}H_{17}N_7O_4Na$ 346.1234 [M+Na$^+$], found 346.1248.

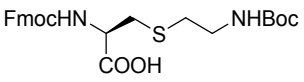
Scheme S2. Synthesis of mPyTSK





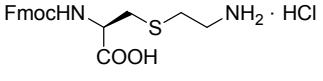
11

2-Amino-3-((2-((*tert*-butoxycarbonyl)amino)ethyl)thio)propanoic acid (11): To a solution of L-cysteine hydrochloride monohydrate (1 g, 5.7 mmol) in aqueous 2 N NaOH solution (12 mL) was added a solution of *tert*-butyl (2-bromoethyl)carbamate (1.4 g, 6.3 mmol) in EtOH (6 mL) at 0 °C. The mixture was stirred at room temperature for 12 h. 3 N HCl was added at 0 °C to neutralize the solution (pH = 7), and then white precipitate formed. The precipitate was filtered to afford the titled compound as white solid (1.4 g, 93% yield): ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.18 (s, 1H), 3.31 – 3.25 (m, 1H), 3.10 – 3.03 (m, 2H), 3.01 – 2.95 (m, 1H), 2.76 – 2.68 (m, 1H), 2.54 (t, J = 7.0 Hz, 2H), 1.36 (s, 9H); ESI-MS calcd for $C_{10}H_{20}N_2O_4SNa$ 287.1 [M+Na$^+$], found 287.1.



12

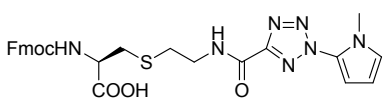
2-((((9*H*-fluoren-9-yl)methoxy)carbonyl)amino)-3-((2-((*tert*-butoxycarbonyl)amino)ethyl)thio)propanoic acid (12): To a solution of **11** (619 mg, 2.34 mmol) in saturated aqueous $NaHCO_3$ solution (10 mL) was added a solution of Fmoc-OSu (830 mg, 2.46 mmol) in dioxane (10 mL) at 0 °C. The mixture was stirred at room temperature for 12 h, then 10% citric acid was added to adjust pH = 2-3. The solution was extracted with EtOAc (8 mL×3). The combined organic layers were washed with H_2O, brine, dried over anhydrous Na_2SO_4, filtered and concentrated. The residue was purified by silica gel flash column chromatography (DCM/MeOH = 10:1) to afford the titled compound as a white solid (1.09 g, 96% yield): ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.88 (d, J = 7.4 Hz, 2H), 7.72 (d, J = 7.3 Hz, 2H), 7.54 (s, 1H), 7.40 (t, J = 7.3 Hz, 2H), 7.32 (t, J = 7.3 Hz, 2H), 6.89 (s, 1H), 4.32 – 4.18 (m, 3H), 4.12 – 4.02 (m, 1H), 3.12 – 3.03 (m, 2H), 2.95 (dd, J = 13.3, 3.5 Hz, 1H), 2.77 (dd, J = 12.7, 9.8 Hz, 1H), 2.60 – 2.51 (m, 2H), 1.35 (s, 9H); ^{13}C NMR (126 MHz, DMSO-d_6) δ 173.0, 156.4, 155.9, 144.3, 141.2, 128.1, 127.5, 125.7, 120.5, 78.1, 66.2, 55.3, 47.1, 34.6, 31.4, 28.7, 22.5; ESI-MS calcd for $C_{25}H_{30}N_2O_6SNa$ 509.2 [M+Na$^+$], found 509.2.



13

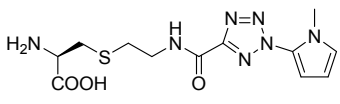
2-((((9*H*-fluoren-9-yl)methoxy)carbonyl)amino)-3-((2-aminoethyl)thio)propanoic acid hydrochloride (13): To a

solution of **12** (390 mg, 0.8 mmol) in dichloromethane (6 mL) was added 4 N HCl in dioxane (2 mL). The mixture was stirred at room temperature for 12 h, then concentrated to afford the titled compound as a white solid (289 mg, 86% yield): ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.89 (d, J = 7.5 Hz, 2H), 7.76 (d, J = 8.4 Hz, 1H), 7.72 (d, J = 7.4 Hz, 2H), 7.41 (t, J = 7.4 Hz, 2H), 7.32 (t, J = 7.3 Hz, 2H), 4.33 – 4.28 (m, 2H), 4.26 – 4.20 (m, 1H), 4.18 – 4.10 (m, 1H), 3.02 – 2.93 (m, 3H), 2.81 – 2.69 (m, 3H); ^{13}C NMR (126 MHz, DMSO-d_6) δ 172.6, 156.5, 144.2, 141.2, 128.1, 127.5, 125.7, 120.6, 66.2, 54.3, 47.1, 38.7, 32.9, 28.8; ESI-MS calcd for $C_{20}H_{23}N_2O_4S$ 387.1 [M+H$^+$], found 387.1.



S2

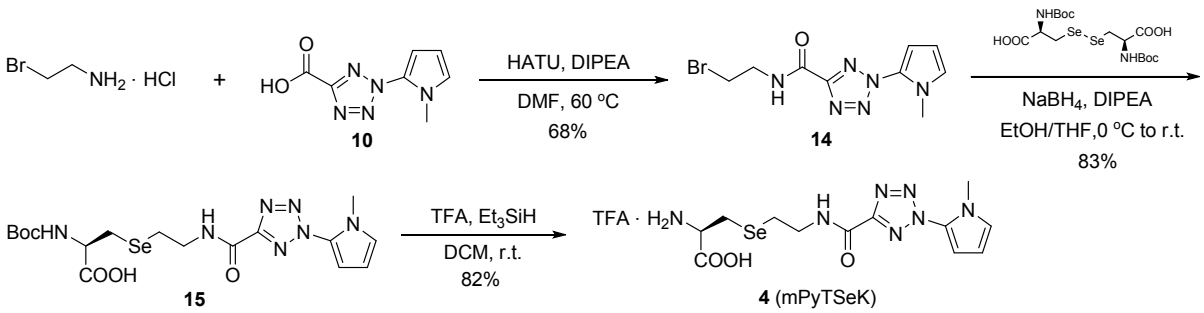
2-((((9*H*-fluoren-9-yl)methoxy)carbonyl)amino)-3-((2-(2-(1-methyl-1*H*-pyrrol-2-yl)-2*H*-tetrazole-5-carboxamido)ethyl)thio)propanoic acid (S2): To a solution of **10**[S2] (100 mg, 0.52 mmol) in DMF (5 mL) were added *N*-hydroxysuccinimide (66 mg, 0.57 mmol) and EDCI (120 mg, 0.62 mmol). The mixture was stirred at room temperature for 2 h before adding **13** (241 mg, 0.57 mmol) and DIPEA (222 mg, 1.72 mmol). The mixture was stirred at room temperature for 12 h. Then DMF was removed under reduced pressure. The residue was dissolved in EtOAc, washed with brine, dried over anhydrous Na_2SO_4, filtered and concentrated. The residue was purified by silica gel flash chromatography (DCM/MeOH = 10:1) to afford the titled compound as a yellow solid (190 mg, 65% yield): ^{1}H NMR (500 MHz, CD_3OD) δ 7.76 (d, J = 7.1 Hz, 2H), 7.66 (d, J = 7.1 Hz, 2H), 7.36 (t, J = 7.4 Hz, 2H), 7.28 (t, J = 7.5 Hz, 2H), 6.87 (s, 1H), 6.58 (s, 1H), 6.22 (s, 1H), 4.44 – 4.37 (m, 1H), 4.30 – 4.19 (m, 3H), 3.66 (s, 3H), 3.18 – 3.10 (m, 2H), 3.01 – 2.94 (m, 1H), 2.89 – 2.79 (m, 3H); ^{13}C NMR (126 MHz, $CDCl_3$) δ 175.5, 158.9, 157.0, 143.7, 141.0, 127.5, 127.0, 125.2, 125.2, 124.8, 123.7, 119.7, 107.6, 106.0, 64.6, 46.9, 34.9, 34.7, 31.6, 25.3, 22.6; ESI-MS calcd for $C_{27}H_{27}N_7O_5SNa$ 584.2 [M+Na$^+$], found 584.1.

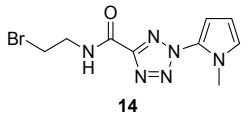


3 (mPyTSK)

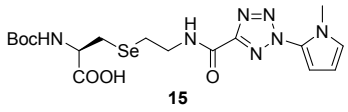
2-Amino-3-((2-(2-(1-methyl-1*H*-pyrrol-2-yl)-2*H*-tetrazole-5-carboxamido)ethyl)thio)propanoic acid (3, mPyTSK): To a solution of **S2** (120 mg, 0.21 mmol) in dichloromethane (5 mL) was added Et_2NH (2 mL). The mixture was stirred at room temperature for 12 h. Et_2O was added to produce white precipitate, which was filtered to afford the titled compound as a white solid (54 mg, 75% yield): ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.03 (s, 1H), 6.60 (s, 1H), 6.20 (s, 1H), 3.63 (m, 4H), 3.37 (m, 2H), 3.04 (m, 1H), 2.83 – 2.68 (m, 3H); ^{13}C NMR (126 MHz, DMSO-d_6) δ 159.9, 156.5, 124.7, 124.6, 110.0, 107.6, 106.5, 54.5, 34.6, 31.5, 28.6; HRMS (EI) calcd for $C_{12}H_{17}N_7O_3SNa$ 362.1006 [M+Na$^+$], found 362.1019.

Scheme S3. Synthesis of mPyTSeK[S3]

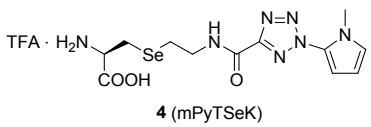




***N*-(2-bromoethyl)-2-(1-methyl-1*H*-pyrrol-2-yl)-2*H*-tetrazole-5-carboxamide (14):** To a solution of **10**[S2] (236 mg, 1.22 mmol) in DMF (5 mL) was added 2-bromoethanamine hydrochloride (275 mg, 1.34 mmol), HATU (510 mg, 1.34 mmol) and DIPEA (315 mg, 2.44 mmol). The mixture was stirred at 60 °C for 4 h. Then DMF was removed under reduced pressure. The residue was dissolved in EtOAc, washed with brine, dried over anhydrous Na_2SO_4, filtered, and concentrated. The residue was purified by silica gel flash column chromatography (hexanes/EtOAc = 1:2) to afford the titled compound as a yellow oil (247 mg, 68% yield): ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.65 (s, 1H), 6.74 (s, 1H), 6.65 (s, 1H), 6.24 (t, J = 3.3 Hz, 1H), 3.98 (q, J = 5.9 Hz, 2H), 3.75 (s, 3H), 3.62 (t, J = 5.9 Hz, 2H); ^{13}C NMR (126 MHz, $CDCl_3$) δ 158.9, 156.5, 124.9, 123.9, 107.9, 106.3, 41.3, 35.1, 31.3; ESI-MS calcd for $C_9H_{11}N_6OBrNa$ 321.0 [M+Na$^+$], found 321.0.



2-((*tert*-Butoxycarbonyl)amino)-3-((2-(2-(1-methyl-1*H*-pyrrol-2-yl)-2*H*-tetrazole-5-carboxamido)ethyl)selanyl)propanoic acid (15): To a solution of Boc-L-selenocystine (155 mg, 0.29 mmol) in EtOH (3 mL) was added $NaBH_4$ (109 mg, 2.89 mmol) at 0 °C. The mixture was stirred at 0 °C for 30 min, then a solution of **14** (191 mg, 0.64 mmol) in THF (3 mL) was added together with DIPEA (375 mg, 2.9 mmol). The mixture was stirred at room temperature for 12 h, then 1 N HCl was added to adjust pH < 5. The solution was extracted with EtOAc (5 mL×3). The combined organic layers were washed with H_2O, brine, dried over anhydrous Na_2SO_4, filtered, and concentrated. The residue was purified by silica gel flash column chromatography (DCM/MeOH = 7:1) to afford the titled compound as a white solid (235 mg, 83% yield): ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.14 (s, 1H), 6.69 (s, 1H), 6.57 (s, 1H), 6.17 (s, 1H), 5.96 (s, 1H), 5.75 (s, 1H), 4.46 (s, 1H), 3.80 – 3.72 (m, 2H), 3.70 (s, 3H), 3.17 – 2.95 (m, 2H), 2.91 – 2.74 (m, 2H), 1.38 (s, 9H); ^{13}C NMR (126 MHz, $CDCl_3$) δ 176.4, 159.1, 156.8, 155.9, 124.8, 123.8, 107.7, 106.2, 80.0, 54.6, 35.0, 31.5, 28.3, 25.3, 22.6; ESI-MS calcd for $C_{17}H_{25}N_7O_5SeNa$ 510.1 [M+Na$^+$], found 510.1.



2-Amino-3-((2-(2-(1-methyl-1*H*-pyrrol-2-yl)-2*H*-tetrazole-5-carboxamido)ethyl)selanyl)propanoic acid TFA salt (4, mPyTSeK): To a solution of **15** (185 mg, 0.38 mmol) in dichloromethane (5 mL) was added Et_3SiH (88 mg, 0.76 mmol) and TFA (1 mL). The mixture was stirred at room temperature for 1 h, then concentrated. The residue was suspended in MeOH/Et_2O to produce white precipitate, which was filtered to

afford the titled compound as a light yellow solid (151 mg, 82% yield): ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.32 (s, 1H), 7.04 (s, 1H), 6.61 (s, 1H), 6.21 (s, 1H), 3.64 (s, 3H), 3.60 – 3.52 (m, 2H), 3.51 – 3.44 (m, 1H), 3.04 (dd, J = 12.9, 4.1 Hz, 1H), 2.89 (dd, J = 12.9, 7.4 Hz, 1H), 2.82 – 2.73 (m, 2H); ^{13}C NMR (126 MHz, DMSO-d_6) δ 169.1, 159.9, 156.5, 124.8, 124.6, 107.6, 106.5, 54.5, 34.6, 25.2, 23.3; HRMS (EI) calcd for $C_{12}H_{17}N_7O_3SeNa$ 410.0450 [M+Na$^+$], found 410.0461.

Site-specific incorporation of mPyTXKs into sfGFP

200 μL of overnight cultured BL21(DE3) cells co-transformed with pET-sfGFP-Q204TAG and pEvol-mPyTKRS plasmids were diluted with fresh 20 mL LB broth containing 100 μg/mL ampicillin and 34 μg/mL chloramphenicol. The bacteria were grown at 37 °C to an OD600 ~0.6 before supplemented with 1 mM mPyTOK/mPyTSK/mPyTSeK and the protein expression was induced with 0.2% arabinose and 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). After incubation for 8 hours (37 °C, 280 rpm), the cells were pelletized in 50 mL conical tubes and resuspended with 4 mL binding buffer (10 mM imidazole, 300 mM NaCl in 50 mM Na_2HPO_4 pH 8.0), lysed by sonication on ice, then centrifuged. The lysates were transferred to 15 mL conical tubes containing 100 μL of Ni-NTA resin slurry (Thermo HisPur™) and incubated for 2 hours with gentle shaking. After removing the supernatant by centrifuge, the resin was washed three times with wash buffer (50 mM imidazole, 300 mM NaCl in 50 mM Na_2HPO_4, pH 8.0) and the proteins were eluted with 150 μL of elution buffer (250 mM imidazole, 300 mM NaCl in 50 mM Na_2HPO_4 pH 8.0). Protein yields were calculated based on the concentration of the protein quantified by Pierce™ BCA protein assay kit (Thermo). 4-5 μg of purified protein was mixed with equal volume of 2X SDS loading buffer, heated at 95 °C for 10 minutes before loading onto 4-12% SDS-PAGE gel (GenScript). The proteins were separated at 140 V for 50 minutes and detected with Coomassie blue staining. Three μg of the protein was analyzed by LC/ESI-MS and the molecular mass of the protein was obtained by deconvoluting its multiply charged protein spectra.

Site-specific incorporation of mPyTXK into glutathione *S*-transferase and subsequent photo-cross-linking reaction

Two hundred μL of overnight cultured BL21(DE3) cells co-transformed with pET28a-GST mutants and pEvol-mPyTKRS plasmids were diluted with fresh 20 mL LB broth containing 50 μg/mL kanamycin and 34 μg/mL chloramphenicol. Protein expression and purification were done using the same procedure as sfGFP. After elution, the solution was concentrated using Amicon Ultra-0.5 mL Centrifugal Filter Units-10,000 (Millipore) and buffer exchanged to PB pH 7.4 to a final volume of 150 μL. Protein yields and molecular mass were obtained using the same procedure as sfGFP. For photo-cross-linking reactions, 4-5 μg of purified GST-mPyTXK mutant proteins were transferred to a 96-well plate (BD Biosciences) and irradiated using handheld 302-nm UV lamp for 15 minutes on ice. Samples were incubated on ice for 20 minutes, then removed from the wells, analyzed by SDS-PAGE and detected with Coomassie blue staining.

H_2O_2–mediated cleavage study[S4, S5]

For oxidative cleavage of cross-linked GST dimer, 4-5 μg of purified GST-mPyTSeK protein was photoirradiated using handheld 302-nm UV lamp for 15 minutes on ice, then denatured with 0.5% SDS at 60 ˚C for 1 h and treated with 8 mM H_2O_2 at 30 ˚C for 1 h before subsequent SDS-PAGE analysis.

Site-specific incorporation of mPyTXK into mCherry-TAG-EGFP-HA in mammalian cells

Human Embryonic Kidney 293T (HEK293T) cells were seeded in 6-well plate and grown in DMEM supplemented with 10% FBS (HyClone™ GE Healthcare Life Sciences) and 10 μg/mL Gentamycin (Gibco) and 2.5 μg/mL plasmocin at 37 °C, 5% CO_2 until ~80% confluency. The medium was replaced with DMEM supplemented with 1 mM mPyTK/ mPyTOK/mPyTSK or 20 μM mPyTSeK and cells were transfected using polyethylenimine (Sigma-Aldrich) in Opti-MEM® (Gibco) with two plasmids: one encoding mPyTKRS/$tRNA_{CUA}$ pair and the other encoding mCherry-TAG-EGFP-HA. Control experiments were performed using DMEM without unnatural amino acids. After 48 hours, live cell images were recorded using Lionheart™ FX automated microscope (BioTek). The cells were lysed by modified RIPA buffer (25 mM Tris· HCl, pH 7.4, 150 mM NaCl, 1% NP-40, 1% sodium deoxycholate, 0.1% SDS, 1 mM EDTA, 1 mM PSMF). Twenty μL lysates were loaded to an SDS-PAGE gel, separated at 140 V for 50 minutes, and then transferred to a PVDF membrane (Thermo Fisher Scientific). The membrane was blocked in 1% casein in TBST (50 mM Tris, 150 mM NaCl, 0.05% Tween-20, pH 7.6) at 4 °C overnight, and then incubated with mouse anti-HA tag antibody (1:10000, Thermo Fisher Scientific) in TBST at room temperature for 1 h. The membrane was washed with TBST (6 x 5 min) before addition of the secondary goat anti-mouse horseradish peroxidase conjugate (1:5000, Santa Cruz Biotech). After 30 minutes, the membrane was washed with TBST (6 x 5 min), and incubated in 100 mM Tris buffer, pH 9.5 before addition of Pierce™ ECL Western Blotting Substrate (Thermo Fisher Scientific) and incubation for 5 min. The blot was exposed to an X-ray film (Phenix).

Site-Directed Mutagenesis to generate pET28b-Grb2-143TAG

The plasmid carrying Grb2 gene with the His tag sequence (HHHHHH) at the C-terminus (pET28b-Grb2) was purchased from Gene Universal Inc. The asparagine codon in position 143 of the Grb2 gene was mutated to amber codon TAG using Q5 Site-Direct Mutagenesis Kit (New England Biolabs) with the following primers (Forward: TGTTAGCCGCtagCAG CAGATTTTT; Reverse: CTGGTGCTGCGATGATAATC) to obtain pET28b-Grb2-143TAG.

Site-specific incorporation of mPyTK/mPyTSeK into Grb2

Two hundred μL of overnight culture of BL21(DE3) cells co-transformed with pET28b-Grb2-N143TAG and pEvol-mPyTKRS plasmids was diluted with fresh 20 mL LB broth containing 50 μg/mL kanamycin and 34 μg/mL chloramphenicol. Protein expression and purification were performed using the same procedure as sfGFP. After elution, the protein

solution was concentrated using Amicon Ultra-0.5 mL Centrifugal Filter Units-10,000 (Millipore) and buffer exchanged to PBS, pH 7.4, to a final volume of 150 μL. The proteins were analyzed by SDS-PAGE and detected with Coomassie blue staining. The protein concentrations were measured by Pierce™ BCA protein assay kit (Thermo).

Photo-cross-linking between Grb2-143mPyTK/mPyTSeK and EGFR in mammalian cell lysates and sequent oxidative cleavage

Human Embryonic Kidney 293T (HEK293T) cells were seeded using the same procedure as described earlier. After reaching ~80% confluency, cells were transfected with pCDNA3-EGFR-EGFP using polyethylenimine (Sigma-Aldrich) in Opti-MEM medium (Gibco). After 24 h, the medium was replaced with FBS-free DMEM and cells were starved for 12 h, treated with EGF at 100 ng/mL for 15 min, and lysed with modified RIPA buffer (25 mM Tris· HCl, pH 7.4, 150 mM NaCl, 1% NP-40, 1% sodium deoxycholate, 0.1% SDS, 1 mM EDTA, 1 mM PSMF). For photo-cross-linking, 20 μL of cell lysates were mixed with 0.5 μg wtGrb2/Grb2-143mPyTK/mPyTSeK, transferred to a 96-well plate (BD Biosciences) and irradiated using handheld 302-nm UV lamp for 15 minutes on ice. Samples were incubated on ice for 20 minutes, and then treated with protein tyrosine phosphatase 1B (Fitzgerald) at 30 °C for 1 h. For oxidative cleavage, the reaction mixture was denatured with 0.5% SDS at 60 ˚C for 1 h and then treated with 8 mM H_2O_2 at 30 ˚C for 1 h. The mixture was resolved by SDS-PAGE before being transferred to PVDF membrane (Thermo Fisher Scientific). The proteins on the membrane were probed with anti-His tag antibody or anti-EGFR antibody and detected with ECL Plus immunodetection system.

Reference:

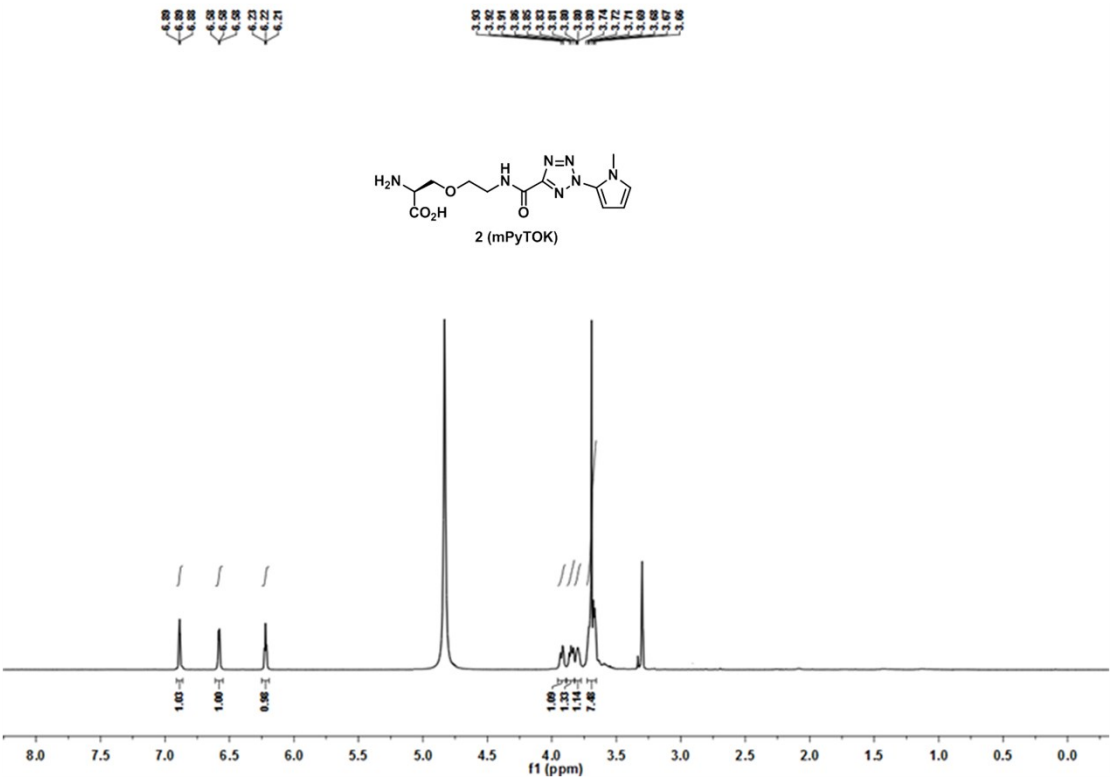
[S1] U. Larsson, R. Carlson, *Acta. Chem. Scand.,* 1994, **48**, 511-516.

[S2] Y. Tian, M. Jacinto, Y. Zeng, Z. Yu, J. Qu, W. Liu, Q. Lin, *J. Am. Chem. Soc.*, 2017, **139**, 6078−6081.

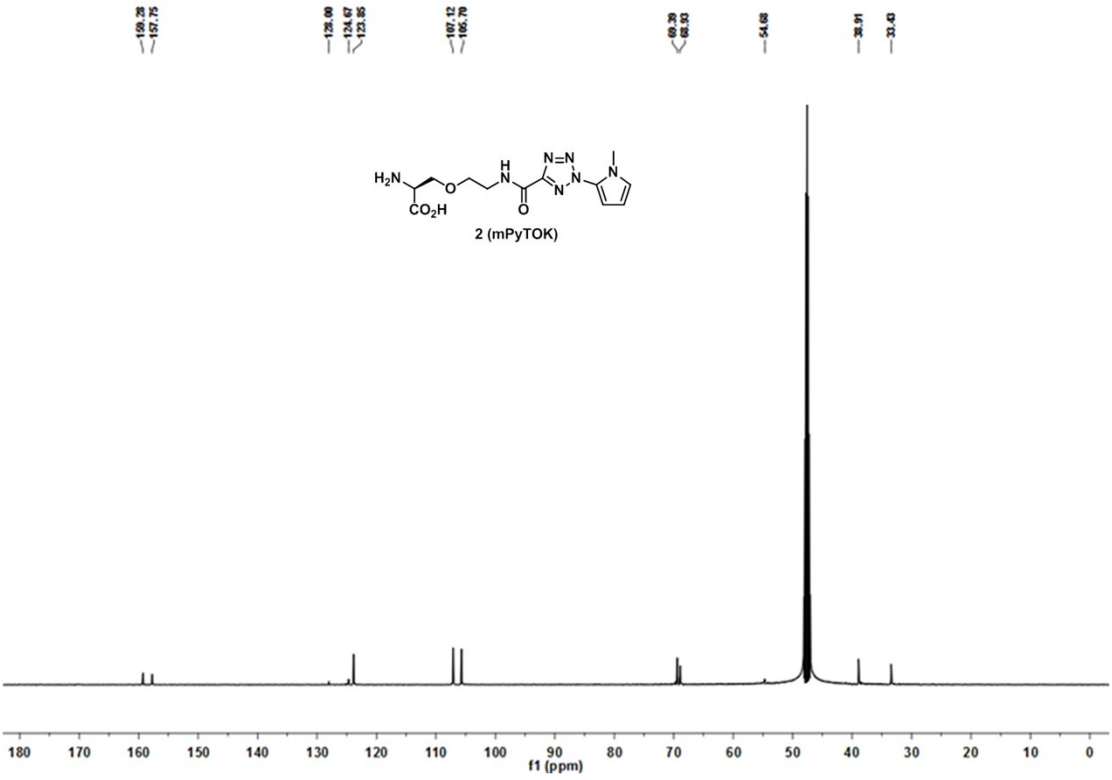
[S3] A. Siebum, W. Woo, J. Raap, J. Lugtenburg, *Eur. J. Org. Chem.*, 2004, **13**, 2905-2913.

[S4] S. Lin, D. He, T. Long, S. Zhang, R. Meng, P. R. Chen, *J. Am. Chem. Soc.*, 2014, **136**, 11860−11863.

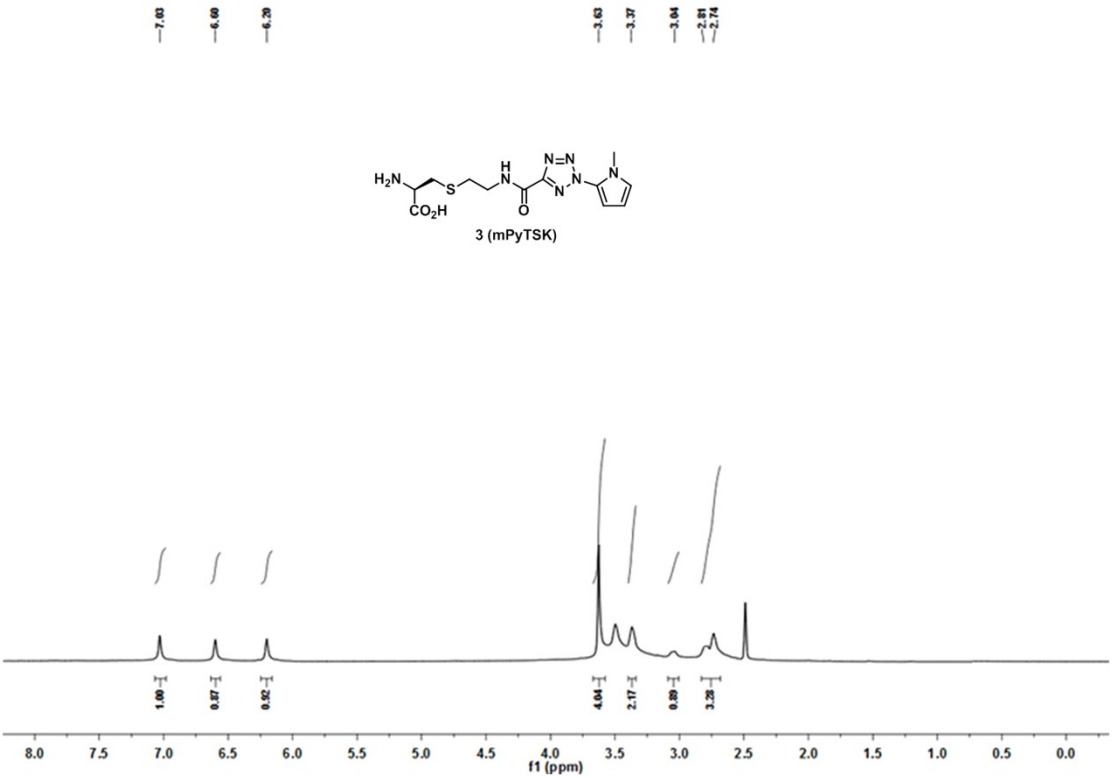
[S5] Y. Yang, H. Song, D. He, S. Zhang, S. Dai, S. Lin, R. Meng, C. Wang, P. R. Chen, *Nat. Commun.*, 2016, **7**, 12299.



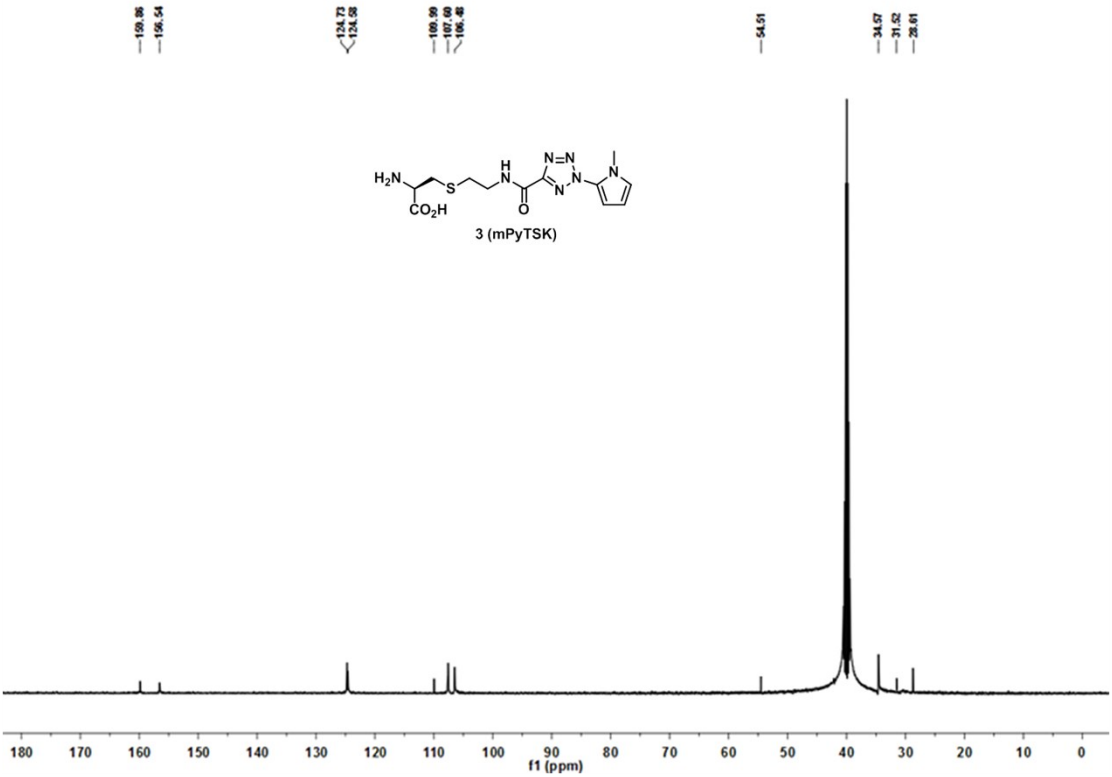
^{1}H NMR spectrum of compound **2** (mPyTOK)



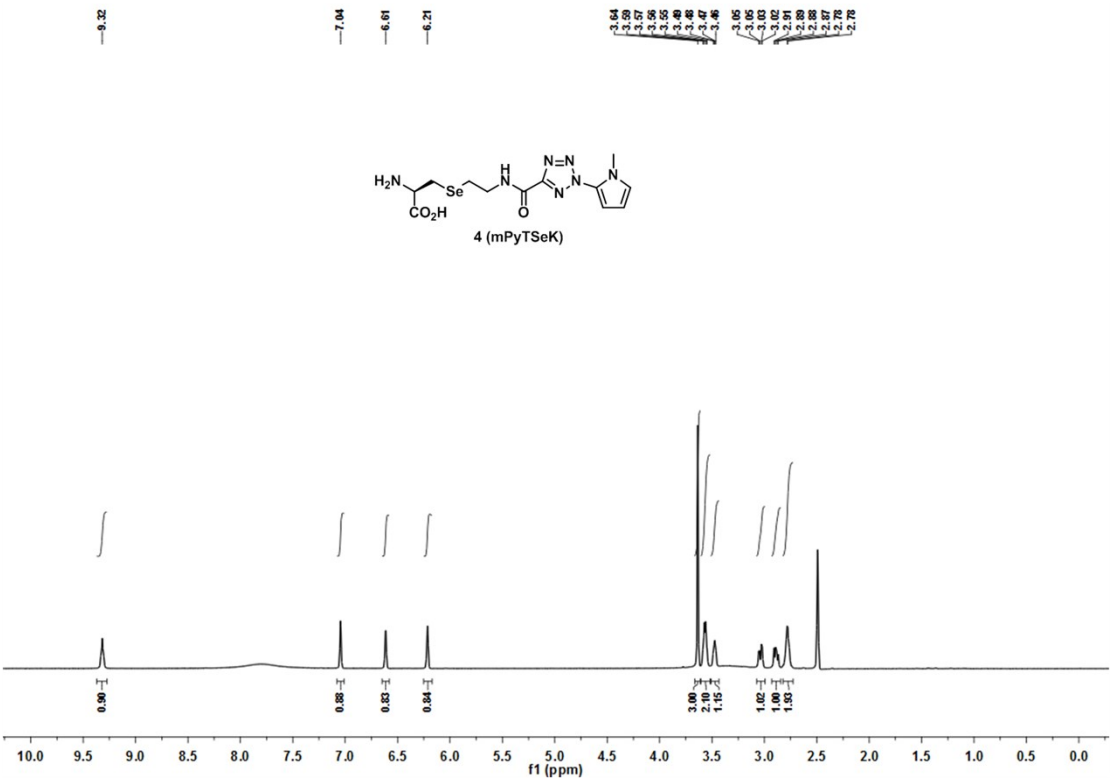
^{13}C NMR spectrum of compound **2** (mPyTOK)



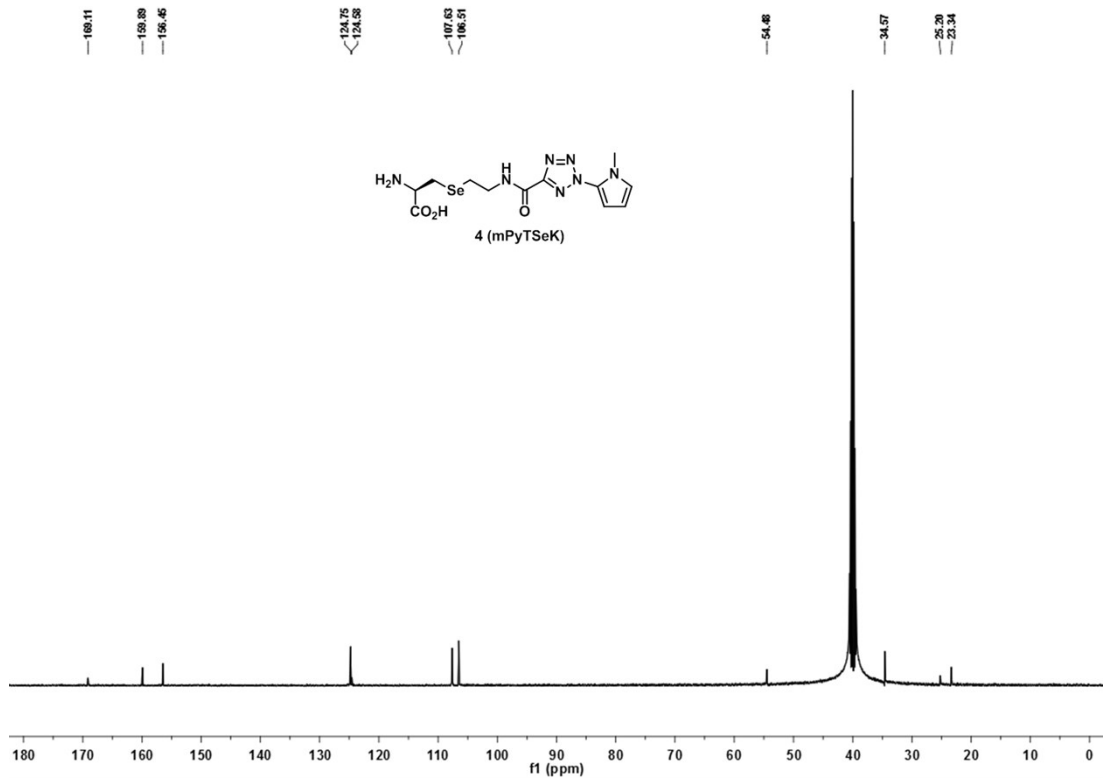
^{1}H NMR spectrum of compound **3** (mPyTSK)



^{13}C NMR spectrum of compound **3** (mPyTSK)



^{1}H NMR spectrum of compound **4** (mPyTSeK)



^{13}C NMR spectrum of compound **4** (mPyTSeK)