Supporting Information for

Access To Faster Eukaryotic Cell Labeling With Encoded Tetrazine Amino Acids

Hyo Sang Jang, Subhashis Jana, Robert J. Blizzard, Joseph C. Meeuwsen, Ryan A. Mehl*

[1]Department of Biochemistry and Biophysics, Oregon State University, 2011 Ag Life Sciences Bldg, Corvallis, OR, 97331-7305 USA

Correspondence to: ryan.mehl@oregonstate.edu

This PDF file includes:

Materials and Methods
Schemes S1–3
Figures S1 to S42
References 1–5

Materials and Methods

General Synthetic Methods:

All purchased chemicals were used without further purification**.** Anhydrous dichloromethane was used after overnight stirring with calcium hydride and distillation under argon atmosphere. Thin-layer chromatography (TLC) was performed on silica 60F-254 plates. The TLC spots of alkene were charred by potassium permanganate staining. Flash chromatographic purification was done on silica gel 60 (230-400 mesh size). ^{1}H NMR spectra were recorded at Bruker 400MHz and 700 MHz and ^{13}C NMR spectra were recorded at 175 MHz. Coupling constants (J value) were reported in hertz. The chemical shifts were shown in ppm and are referenced to the residual non-deuterated solvent peak $CDCl_3$ (δ =7.26 in ^{1}H NMR, δ = 77.23 in ^{13}C NMR), CD_3OD (δ =3.31 in ^{1}H NMR, δ = 49.2 in ^{13}C NMR), d_6-DMSO (δ =2.5 in ^{1}H NMR, δ = 39.5 in ^{13}C NMR) as an internal standard. Splitting patterns of protons are designated as follows: s-singlet, d-doublet, t-triplet, q-quartet, quin-quintet, sext-sextet, sept- septet, m-multiplet.

Synthetic Procedure

1. General Synthetic Procedure of Boc-protected Tet-v3.0 Derivatives (Scheme S1). The starting material Boc-protected 3-cyano phenylalanine, **4**, (1.0 equiv) was taken in a dried heavy walled reaction tube under argon atmosphere. Maintaining inert atmosphere inside the reaction vessel, catalyst $Ni(OTf)_2$ (0.5 equiv) and corresponding nitrile derivatives (10 equiv) were added. Then anhydrous hydrazine (50 equiv) was added slowly to reaction mixture using a glass syringe and under stirring condition. Purged argon for another 10 minutes and immediately sealed the tube. The reaction vessel was immersed into the preheated oil bath at 50-55 $^{\circ}$C for 24 hours. The reaction vessel was brought to room temperature and opened with care. The reaction mixture was poured into a beaker and added 10 equiv 2 M $NaNO_2$ aqueous solution and 5 mL water. The aqueous phase was washed with ethyl acetate (20 mL) to remove the homo coupled-tetrazine byproduct. Then the aqueous phase was acidified with 4 M HCl (pH~2) under ice-cold condition with homogeneous mixing and extracted with ethyl acetate (3x 20 mL). The combined organic layer was washed with brine, dried with anhydrous Na_2SO_4 and concentrated under reduced pressure. Silica gel flash column chromatography (20-25% ethyl acetate in hexanes with 1% acetic acid) yielded Boc-protected tetrazine derivatives, **2**, in the form of a pinkish red gummy material.

2. General Procedure of Boc-deprotection of Tet 3.0 Derivatives (Scheme 1). In a dry flask, Boc-protected Tet-v3.0 derivatives, **2**, were dissolved in 3 mL ethyl acetate and 1 mL HCl gas saturated 1,4 dioxane was added to the solution under argon atmosphere. The reaction mixture was stirred at room temperature until the starting materials were consumed monitoring by TLC (normally 3 to 4 hr). After completion, remove the solvent under reduced pressure and dissolved in ethyl acetate (2x 10 mL) and similarly concentrated to remove excess HCl gas. Finally, added 5 mL pentane and dried resulting in a pink color solid chloride salt of **1, Tet-v3.0,** derivatives in near quantitative yield (~ 97-98%).

***(S)-2-((tert-butoxycarbonyl)amino)-3-(3-(6-methyl-1,2,4,5-tetrazin-3-yl)phenyl)propanoic acid* (2a):** Following the general synthetic procedure **1**, starting from 0.30 g (1.03 mmol) of Boc-protected 3-cyano phenyl alanine **4** and 0.53 mL (10.3 mmol) of acetonitrile afforded 0.274 g (0.76 mmol) of the title compound **2a** as a pink gummy material. Yield 74%. ^{1}H NMR (400MHz, $CDCl_3$)

δ 8.41 (2H, t, J = 7.2 Hz), 7.51-7.44 (2H, m), 5.17 (1H, d, J = 7.4), 4.71 (1H, d, J = 5.2), 3.33 (1H, dd, J = 13.6, 5.2 Hz), 3.21 (1H, dd, J = 13.2, 6.4 Hz), 3.07 (3H, s), 1.39 (9H, s). ^{13}C NMR (175MHz, $CDCl_3$) δ 175.6, 167.3, 164.1, 155.5, 137.7, 133.8, 132.1, 129.5, 129.1, 126.7, 80.4, 54.4, 38.1, 28.4, 21.1.

Chloride salt of (S)-2-amino-3-(3-(6-methyl-1,2,4,5-tetrazin-3-yl)phenyl)propanoic acid **(1a):** Following the general procedure **2**, 0.274 g (0.76 mmol) of Boc-protected methyl derivative of Tet -v3.0 **2a** yielded 0.218 g (0.74 mmol) of the title compound **1a.** Yield 97%. ^{1}H NMR (400MHz, CD_3OD) δ 8.52-8.49 (2H, m), 7.66-7.60 (2H, m), 4.38 (1H, dd, J =7.2, 6 Hz), 3.46 (1H, dd, J = 14.4, 5.6 Hz), 3.35 (1H, dd, J = 14.4, 7.2 Hz), 3.05 (3H, s). ^{13}C NMR (175 MHz, CD_3OD) δ 171.1, 169.1, 165.2, 137.1, 134.7, 134.4, 131.2, 129.8, 128.4, 55.1, 37.3, 21.1. ESI-MS calculated for $C_{12}H_{14}N_5O_2$ ($[M + H]^+$) 260.1142, found 260.1133.

(S)-2-((tert-butoxycarbonyl)amino)-3-(3-(6-ethyl-1,2,4,5-tetrazin-3-yl)phenyl)propanoic acid **(2b):** Following the general synthetic procedure **1**, starting from 0.30 g (1.03 mmol) of Boc-protected 3-cyano phenyl alanine **4** and 0.72 mL (10.3 mmol) of propionitrile afforded 0.294 g (0.79 mmol) of the title compound **2b** (ethyl derivative of Tet -v3.0**)** as a pink gummy material. Yield 72%. %. ^{1}H NMR (400MHz, $CDCl_3$) δ 8.45 (1H, d, J = 7.6 Hz), 8.41 (1H, s), 7.50 (1H, t, J = 7.6 Hz), 7.45 (1H, d, J = 7.6 Hz), 5.14 (1H, d, J = 7.6 Hz), 4.70 (1H, d, J = 5.6 Hz), 3.38 (2H, q, J = 7.6 Hz), 3.32 (1H, d, J = 4.4 Hz), 3.19 (1H, dd, J = 12.8, 5.6 Hz), 1.53 (3H, t, J = 7.6 Hz), 1.39 (9H, s). ^{13}C NMR (175MHz, $CDCl_3$) δ 176.1, 171.1, 164.3, 155.5, 137.6, 133.8, 132.3, 129.6, 129.1, 126.8, 80.5, 54.5, 38.1, 28.5, 28.4, 12.4.

Chloride salt of (S)-2-amino-3-(3-(6-ethyl-1,2,4,5-tetrazin-3-yl)phenyl)propanoic acid **(1b):** Following the general procedure **2**, 0.270 g (0.72 mmol) of Boc-protected ethyl derivative of Tet -v3.0 **2b** yielded 0.220 g (0.71 mmol) of the title compound **1b.** Yield 98%. ^{1}H NMR (400MHz, CD_3OD) δ 8.52 (1H, s), 8.51 (1H, d, J = 1.6 Hz), 7.66-7.61 (2H, m), 4.38 (1H, d, J = 6 Hz), 3.47 (1H, dd, J = 14.8, 5.6 Hz), 3.37 (3H, q, J = 7.6 Hz), 1.53 (3H, t, J = 7.6 Hz). ^{13}C NMR (175 MHz, CD_3OD) δ 172.3, 165.4, 137.2, 134.7, 134.3, 131.2, 129.9, 128.3, 55.3, 37.3, 29.2, 12.4. ESI-MS calculated for $C_{13}H_{16}N_5O_2$ ($[M + H]^+$) 274.1299, found 274.1290.

(S)-2-((tert-butoxycarbonyl)amino)-3-(3-(6-isopropyl-1,2,4,5-tetrazin-3-yl)phenyl)propanoic acid **(2c):** Following the general synthetic procedure **1**, starting from 0.20 g (0.69 mmol) of Boc-protected 3-cyano phenyl alanine **4** and 0.62 mL (6.9 mmol) of isobutyronitrile afforded 0.152 g (0.39 mmol) of the title compound **2c** (isopropyl derivative of Tet -v3.0**)** as a pink gummy material. Yield 57%. ^{1}H NMR (400MHz, $CDCl_3$) δ 8.47 (1H, d, J = 7.6 Hz), 8.42 (1H, s), 7.52 (1H, t, J = 7.6 Hz), 7.45 (1H, d, J = 7.6 Hz), 5.08 (1H, d, J = 6.4 Hz), 4.69 (1H, d, J = 4 Hz), 3.68 (1H, sept, J = 6.8 Hz), 3.35 (1H, dd, J = 13.2, 4.4 Hz), 3.2 (1H, dd, J = 12.8, 5.6 Hz), 1.55 (6H, d, J = 7.2 Hz), 1.39 (9H, s). ^{13}C NMR (175MHz, $CDCl_3$) δ 175.8, 173.8, 164.3, 155.5, 137.5, 133.7, 132.4, 129.7, 129.1, 126.7, 54.5, 38.1, 34.4, 28.4, 21.4.

Chloride salt of (S)-2-amino-3-(3-(6-isopropyl-1,2,4,5-tetrazin-3-yl)phenyl)propanoic acid (1c): Following the general procedure **2**, 0.150 g (0.39 mmol) of Boc-protected isopropyl derivative of Tet -v3.0 **2c** yielded 0.122 g (0.38 mmol) of the title compound **1c.** Yield 97.5%. ^{1}H NMR (400MHz, CD_3OD) δ 8.52 (1H, s),8.51 (1H, d, J = 6.4 Hz), 7.66-7.62 (2H, m), 4.36 (1H, bs), 3.65 (1H, sept, J = 6.8 Hz), 3.46 (1H, dd, J = 14, 4.4 Hz), 3.4 (1H, dd, J = 12.8, 4 Hz), 1.54 (6H, d, J =

7.2 Hz). ^{13}C NMR (175MHz, CD_3OD) δ 175.1, 165.4, 137.3, 134.7, 134.4, 131.2, 129.9, 128.3, 55.6, 37.4, 35.5, 21.5. ESI-MS calculated for $C_{14}H_{18}N_5O_2$ ($[M + H]^+$) 288.1455, found 288.1451.

***(S)-2-((tert-butoxycarbonyl)amino)-3-(3-(6-butyl-1,2,4,5-tetrazin-3-yl)phenyl)propanoic acid* (2d):** Following the general synthetic procedure **1**, starting from 0.40 g (1.37 mmol) of Boc-protected 3-cyano phenyl alanine **4** and 1.4 mL (13.3 mmol) of valeronitrile afforded 0.43 g (1.07 mmol) of the title compound **2d** (butyl derivative of Tet -v3.0) as a pink gummy material. Yield 78%. ^{1}H NMR (400MHz, $CDCl_3$)) δ 8.46 (1H, d, J = 7.6 Hz), 8.42 (1H, s), 7.51 (1H, t, J = 7.6 Hz), 7.46 (1H, d, J = 7.6 Hz), 5.12 (1H, d, J = 6.8 Hz), 4.7 (1H, d, J = 4 Hz), 3.34 (3H, t, J = 7.6 Hz), 3.20 (1H, t, J = 6.4 Hz), 1.95 (2H, quin, J = 7.6 Hz), 1.47 (2H, sext, J = 7.6 Hz), 1.39 (9H, s), 0.98 (3H, t, J = 7.2 Hz). ^{13}C NMR (175MHz, $CDCl_3$) δ 175.8, 170.4, 164.2, 155.5, 137.7, 133.7, 132.3, 129.7, 129.1, 126.8, 80.5, 54.5, 38.2, 34.6, 30.4, 28.4, 22.4, 13.8.

***Chloride salt of (S)-2-amino-3-(3-(6-butyl-1,2,4,5-tetrazin-3-yl)phenyl)propanoic acid* (1d):** Following the general procedure **2**, 0.400 g (0.99 mmol) of Boc-protected butyl derivative of Tet -v3.0 **2d** yielded 0.326 g (0.97 mmol) of the title compound **1d.** Yield 98%. ^{1}H NMR (400MHz, CD_3OD) δ 8.52 (1H, s),8.51 (1H, d, J = 6 Hz), 7.66-7.61 (2H, m), 4.38 (1H, t, J = 6.4 Hz), 3.46 (1H, dd, J = 8.8, 5.6 Hz), 3.35 (3H, dd, J = 15.6, 8 Hz), 1.96 (2H, quin, J = 7.6 Hz), 1.47 (2H, sext, J = 7.2 Hz), 1.02 (3H, t, J = 7.8 Hz). ^{13}C NMR (175MHz, CD_3OD) δ 171.7, 165.3, 137.2, 134.7, 134.3, 131.2, 129.9, 128.4, 55.2, 37.3, 35.4, 31.3, 23.4, 14.1. ESI-MS calculated for $C_{15}H_{20}N_5O_2$ ($[M + H]^+$) 302.1612, found 302.1604.

***(Z)-bicyclo[6.1.0]non-4-en-9-ylmethanol* (sCCO, 7):** Synthetic procedure, O'Brien et al.[1], ^{1}H NMR (700MHz, CD_3OD) δ 5.64-5.60 (2H, m), 3.38 (2H, d, J = 7 Hz), 2.30-2.26 (2H, m), 2.19-2.15 (2H, m), 2.09-2.04 (2H, m), 1.73-1.43 (2H, m), 0.79-0.74 (2H, m), 0.58-0.55 (1H, m).

***(E)-bicyclo[6.1.0]non-4-en-9-ylmethanol* (sTCO, 8):** Synthetic procedure, Royzen et al.[2], ^{1}H NMR (700MHz, CD_3OD) δ 5.89-5.84 (1H, m), 5.15-5.10 (1H, m), 3.44-3.39 (2H, m), 2.37 (1H, d, J = 13.3 Hz), 2.27 (1H, dt, J = 12.6, 4.2 Hz), 2.25-2.23 (1H, m), 2.19-2.15 (1H, m), 1.94-1.87 (2H, m), 0.92-0.87 (1H, m), 0.64-0.58 (1H, m), 0.50-0.46 (1H, m), 0.37-0.31 (2H, m).

***(E)-bicyclo[6.1.0]non-4-en-9-ylmethyl (4-nitrophenyl) carbonate* (9):** In a dry round-bottom flask, sTCO **8** (0.3 gm, 1.97 mmol) was dissolved in anhydrous dichloromethane (DCM) under inert atmosphere. Subsequently, trimethylamine (Et_3N) (650 µL, 4.9 mmol) and 4-Nitrophenyl chloroformate (0.43 gm, 2.16 mmol) were added to the solution and stirred at room temperature for 2-3 hrs. After consumption of all starting material (monitored by TLC), added 15 mL DCM to the reaction mixture and washed with water. The aqueous layer was re-extracted twice with DCM. The organic layers were combined, dried with anhydrous Na_2SO_4, concentrated using rotary evaporator. Purification was done using silica gel flash column chromatography (5% ethyl acetate in hexane) yielded yellowish white solid material **9** (0.51 gm, 1.6 mmol). Yield 81%. ^{1}H NMR (400MHz, $CDCl_3$) δ 8.27 (2H, d, J = 9.6 Hz), 7.37 (2H, d, J = 9.6 Hz), 5.88-5.82 (1H, m), 5.18-5.14 (1H, m), 4.18 (2H, d, J = 7.2 Hz), 2.43-2.39 (1H, m), 2.35-2.22 (3H, m), 1.96-1.90 (2H, m), 0.94-0.83 (1H, m), 0.69-0.64 (1H, m), 0.62-0.49 (3H, m).

3. General synthetic procedure of TAMRA linked strained alkenes:
A dry 10 mL round-bottom flask was charged with Tetramethylrhodamine 5 - (and - 6) - carboxamide cadaverine (TAMRA) and activated ester of strained alkene (**9** or **10**) (1.5 equiv) under N_2 atmosphere. Anhydrous dichloromethane (2 mL) and N,N-diisopropylethylamine (DIPEA) (3 equiv) were added to the reaction mixture and allowed to stirrer at room temperature for 18 hours. After that, solvent was concentrated onto silica gel under reduced pressure and directly loaded on the silica gel column chromatography for purification. Using the solvent gradient 30-35% methanol in dichloromethane isolated desired molecules as a red solid material. Yield 50-57%.

TAMRA linked sCCO (5): Using general procedure **3**, 5 mg (9.7 µmol) of 5(6) - TAMRA cadaverine yielded 3.4 mg (4.9 µmol) of the title compound **(5).** Yield 51%. ^{1}H NMR (400MHz, CD_3OD) δ 8.14 (1H, d, J = 7.2 Hz), 8.07 (1H, dd, J = 8.0, 1.6 Hz), 7.69 (1H, d, J = 2.0 Hz), 7.27 (1H, s), 7.25 (1H, s), 7.03 (1H, d, J = 2.4 Hz), 7.01 (1H, d, J = 2.4 Hz), 6.92 (2H, d, J = 2.4 Hz), 5.62-5.58 (2H, m), 3.82 (2H, d, J = 7.6 Hz), 3.38 (2H, t, J = 7.2 Hz), 3.28 (12H, s), 3.08 (2H, t, J = 7.2 Hz), 2.27-2.22 (2H, m), 2.16-2.02 (3H, m), 1.67-1.59 (2H, m), 1.56-1.49 (2H, m), 1.42-1.34 (5H, m), 0.83-0.78 (2H, m), 0.64-0.59 (1H, m).

TAMRA linked sTCO (6): Using general procedure **3**, 10 mg (0.019 mmol) of 5(6) - TAMRA cadaverine yielded 7.5 mg (0.010 mmol) of the title compound **(6).** Yield 57%. ^{1}H NMR (400MHz, CD_3OD) δ 8.50 (1H, s), 8.05 (1H, t, J = 8.4 Hz), 7.35 (1H, d, J = 7.2 Hz), 7.25 (2H, dd, J = 8.8, 4.0 Hz), 7.01 (2H, d, J = 9.2 Hz), 6.91 (2H, s), 5.89-5.78 (1H, m), 5.16-5.04 (1H, m), 3.90 (1H, bs), 3.61 (1H, s), 3.49-3.42 (2H, m), 3.27 (12H, s), 3.16-3.05 (2H, m), 2.37-2.29 (1H, m), 2.26-2.16 (2H, m), 1.95 (2H, s), 1.92-1.87 (1H, m), 1.74-1.43 (5H, m), 1.35 (3H, d, J = 6 Hz), 0.93-0.82 (1H, m), 0.62-0.50 (1H, m), 0.46-0.36 (1H, m).

dh-sTCO (11): In 3 mL anhydrous dichloromethane, Amino-PEG192-Amine (15 mg, 0.07 mmol), the activated ester of sTCO **9** (86 mg, 0.27 mmol) and followed by triethylamine (30 µL, 0.212 mmol) were added under argon atmosphere. The reaction mixture was stirred at room temperature for 24 hours. The solution was concentrated onto silica gel under reduced pressure and purified our desired molecule **11** (25 mg, 0.045 mmol) by silica gel column chromatography (10-15% methanol in dichloromethane). Yield 64%. ^{1}H NMR (400MHz, CD_3OD) δ 5.91-5.83 (2H, m), 5.18-5.10 (2H, m), 3.93(4H, d, J = 6.4 Hz), 3.66-3.62 (8H, m), 3.54 (4H, t, J = 5.6 Hz), 3.29 (4H, q, J = 5.6 Hz), 2.36 (2H, dd, J = 14, 2 Hz), 2.30-2.17 (6H, m), 1.98-1.87 (4H, m), 0.94-0.85 (2H, m), 0.67-0.55 (4H, m), 0.49-0.44 (4H, m).

sTCO-PEG5000 (12): Following the above procedure, using 130 mg (0.026 mmol) of mPEG5000-Amine, 10 mg of the activated ester of sTCO **9** (0.031 mmol) and 10 µL triethylamine (0.052 mmol) produced 104 mg of the title molecule **12** (0.019 mmol). Yield 74%. Compound was purified by silica gel column chromatography (5% methanol in dichloromethane). ^{1}H NMR (400MHz, CD_3OD) δ 5.90-5.82 (1H, m), 5.17-5.09 (1H, m), 3.91(2H, d, J = 6.8 Hz), 3.82-3.79 (3H, m), 3.63 (411H, bs), 3.54-3.50 (4H, m), 3.45 (2H, t, J = 4.8 Hz), 3.35 (3H, s), 3.26 (2H, t, J = 5.6 Hz), 3.16 (2H, q, J = 7.6 Hz), 2.35 (1H, d, J = 15.2 Hz), 2.27-2.15 (3H, m), 1.97-1.88 (2H, m), 0.95-0.84 (1H, m), 0.66-0.54 (2H, m), 0.48-0.41 (2H, m).

Tetrazine amino acid rate constants. The solutions of tetrazine amino acids (0.2 mM) and sTCO (1.0-6.0 mM) were made in PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.4) with 2% methanol to equally solubilize all compounds. To measure the rate constants the loss tetrazine signal at 270 nm was measured over time using pseudo first order conditions with sTCO in 5 to 60 fold excess of tetrazine amino acids at 25 ºC. All measurements were performed in triplicate and the resulted decay curves were fit to a single exponential equation. The mean value of pseudo first order rate constant (k') were plotted against different concentration of sTCO to obtain second order rate constant (k) from the slope of the plot (Figure S3).

Selection of Tet-v3.0 Aminoacyl-tRNA Synthetases. The library-1 (Figure S1) from the *Methanosarcina barkeri* (*Mb*) system was chosen for its previously demonstrated ability to incorporate large aromatic amino acids into proteins following the same alternating rounds of life (positive) and death (negative) selection in the presence and absence of Tet-v3.0Me.[3] The remaining pBK plasmid library members after the first negative and second negative selection steps were evaluated for their ability to suppress TAG codons in the GFP gene on the pALS reporter plasmid. This plasmid contains the GFP reporter with a TAG codon at residue 150 as well as pyrrolysyl-$tRNA_{CUA}$. When a pBK plasmid with a functional synthetase is transformed with the pALS plasmid and the cells are grown in the presence of the appropriate amino acid on autoinduction agar, sfGFP is expressed and the colonies are visibly green. The remaining pBK plasmid library (70 ng) was transformed into 100 µL of pALS-containing DH10B cells. The cells were rescued for 1 hour in 1 mL of SOC (37 °C, 250 rpm). Both 250 µL and 50 µL aliquots of cells were plated on autoinducing agar plates with 25 µg/mL kanamycin and 50 µg/mL tetracycline. The plates were further divided by the presence or absence of amino acid (1 mM Tet-v3.0). Plates were grown at 37 °C for 24 hours and then grown on the bench top, at room temperature, for an additional 24 hours.

A total of 72 visually green colonies from the two 1 mM Tet-v3.0 plates and 24 visually white colonies from the two plates without Tet-v3.0 were used to inoculate a 96-well plate containing 0.5 mL per well non-inducing media (NIM)[3-4] containing 50 µg/mL kanamycin and 25 µg/mL tetracycline. After 24 hours of growth (37 °C, 250 rpm), 50 µL of these non-inducing samples were used to inoculate two 96-well plates with 0.5 mL autoinducing media (AIM)[3-4] containing 50 µg/mL kanamycin, 25 µg/mL tetracycline. One 96-well plate was created with and one 96-well plate was created without 1 mM Tet-v3.0. Fluorescence measurements of the cultures were collected 24, 48, and 72 hours after inoculation using a BIOTEK® Synergy 2 Microplate Reader. The emission from 528 nm (20 nm bandwidth) was summed with excitation at 485 nm (20 nm bandwidth). Samples were prepared by diluting suspended cells directly from culture 4-fold with sterile water. Fifteen colonies were selected for their high fluorescence with Tet-v3.0 present and low fluorescence in the absence of ncAAs. Selected hits were grown overnight (37 °C for 24 hours) in LB media containing 50 µg/mL kanamycin and 25 µg/mL tetracycline. Grown media (800 µL) was mixed with 80% (v/v) glycerol (200 µL) and stored at -80 °C. The sequencing of the fifteen pBK-libraries members resulted in six unique synthetases (Figure S4b).

Characterization of Efficiency and Fidelity. After the completion of the selections, efficiency and fidelity were measured in larger, more aerated cultures. Cell stocks were used to inoculate 5 mL of NIM containing 50 µg/mL kanamycin and 25 µg/mL tetracycline and allowed to grow overnight (37 °C, 18 hours). Saturated NIM (50 µL) was then used to inoculate AIM containing 50 µg/mL kanamycin, 25 µg/mL tetracycline , and Tet-v3.0 (1 mM). Fluorescence was measured

every 12 hours for 48 hours. The six unique synthetases are compared in (Figure S4a). Fluorescence measurements were collected using a Turner Biosystems Picofluor fluorimeter diluting 100 µL cell culture in 1.9 mL water.

Permissivity Screen of Select Synthetases. Cell stocks stored at -80 °C containing Tet-v3.0 selection hits in the pBK/pALS system, were used to inoculate 5 mL of NIM containing kanamycin (50 µg/mL) and tetracycline (25 µg/mL). Cells were grown for 16 hours at 37 °C shaking at 250 rpm. The grown NIM cultures were used to inoculate 5 mL cultures of AIM (50 µL inoculate volume) containing kanamycin (50 µg/mL) and tetracycline (25 µg/mL). A separate culture was inoculated for each Tet-v3.0 amino acid tested and 1 mM amino acid was introduced to those cultures from a 100 mM DMF stock solution. Cultures were grown for 30-36 hours at 37 °C and 250 rpm. Fluorescence was assessed every 12 hours by removing 100 µL of media and diluting it to 2 mL total volume. Fluorescence was measured using a Turner Biosystems Picofluor fluorimeter.

Expression and purification of GFP-TAG150-Tet-v3.0. A cell stock of DH10B cells cotransformed with the R2-84 pBK plasmid, and pALS plasmid was used to inoculate a 5 mL culture of NIM containing kanamycin (50 µg/mL) and tetracycline (25 µg/mL), which was then grown 16 hours at 37 °C shaking at 250 rpm. A 50 mL AIM culture containing kanamycin (50 µg/mL) and tetracycline (25 µg/mL) was inoculated with 0.5 mL of the grown NIM. The AIM was supplemented with 1 mM Tet-v3.0 from a 100 mM DMF stock solution. The AIM culture was allowed to grow shaking at 250 rpm for 48 hours at 37 °C. Additionally, cultures without Tet-v3.0 and containing only the WT pALS plasmid were grown simultaneously. All cells were harvested by centrifugation 5000 rcf for 5 min. Supernatant was decanted and cell pellets were stored at -80 °C. To purify, cells were resuspended in lysis buffer (NaCl 300 mM, NaH_2PO_4 50 mM, pH 7.0). Cells were lysed using a Microfluidics M-110P microfluidizer (18,000 psi) and the lysate was collected in lysis buffer. The lysate was clarified by centrifugation (21036 rcf, 1 hour) and to the decanted supernatant was added 100 µL bed volume TALON resin. Lysate was incubated with the resin for 1-2 hours gently rocking at 4° C. Resin and lysate were applied to a column and flow through was discarded. Resin was washed with 3 x 10 mL lysis buffer. Protein was eluted with 4 x 250 µL elution buffer (NaCl 300 mM, NaH_2PO_4 50 mM, imidazole 250 mM, pH 7.0). Protein concentration was measured using a Bradford assay. Protein purity was assessed using SDS-PAGE (Figure S5).

Mass spectra of GFP-Tet-v3.0. Purified GFP-TAG150-Tet-v3.0 was diluted to 10 µM and desalted on C_4 zip tips and analyzed using an FT LTQ mass spectrometer at the Oregon State University mass spectrometry facility (Figure S6).

A 6545XT Q-TOF mass spectrometer interfaced with a 1290 Infinity II HPLC (Agilent Technologies, Santa Clara, CA) was used to verify reaction of purified GFP-TAG150-Tet-v3.0 with sTCO. Intact proteins were desalted and concentrated using a Poroshell 300SB-C18 column prior to electrospray. Spectra were deconvoluted using the Maximum Entropy deconvolution algorithm in BioConfirm.

Tet-v3.0 concentration dependent expression of GFP-Tet-v3.0. Auto-induction media (3 mL, tetracycline 25 ng/µL, kanamycin 50 ng/µL) was inoculated using 30 µL of non-induction media cultures of DH10B cells cotransformed with the pALS plasmid and a pBK-R2-84 or pBK-R2-74

plasmid. Autoinduction media contained various amounts of the amino acid of interest ranging from 0-1.0 mM. Cultures were grown at 37 °C and shaking at 250 rpm. OD_{600} and fluorescence measurements were taken every 12 hours and the fluorescence was normalized to the OD.

Measuring reaction rates of GFP-Tet-v3.0 with sTCO. Fluorescence of 1.9 nmol of purified GFP-Tet-v3.0 diluted in 3 mL of PBS was measured (488 nm excitation, 509 nm emission, 5 points/second) for 60 seconds prior to the addition of various quantities (2 nmol-200 nmol) of sTCO. Fluorescence was measured until no fluorescence increase was observed. Curves were fit using the curve-fitting program Igor to determine kinetic constants.

Cloning into pDule1/2. The pDule forward and reverse primers were used to amplify the synthetase gene from the pBK plasmid. The amplified fragment was gel purified and cloned into the pDule1/pDule2 plasmid using a SLICE reaction.
pDule Forward primer:
5'-GAGTTTACGCTTTGAGGAATCCCCCATGGATAAAAAACCGCTGGATG-3'
pDule Reverse primer:
5'-CCTCTTCTGAGATGAGTTTTTGTTCTTACAGGTTCGTGCTAATGC-3'

Cloning Tet-3.0 RSs into pAcBac1. Mammalian codon optimized pylRS synthetases genes for R2-74-RS and R2-84-RS were amplified with pACBac forward-RS and pACBac reverse-RS primers. The PCR products and pUC plasmid backbone were digested with the restriction enzymes EcoR1 and Xho1. The desired synthetase genes fragments were gel purified and ligated into pUC backbone. The ligated product was sequenced to confirm identity. The pUC plasmids containing both the R2-74 and R2-84 synthetases, as well as pAcBac1 plasmid[5] with no insert were digested using the restriction enzymes Nhe1 and EcoR1. The fragments containing the synthetase gene from the pUC plasmids and the fragments containing the pAcBac1 backbone were gel purified. A ligation reaction was performed between individual synthetase genes and the pAcBac1 backbone. Identity of ligation products was confirmed by sequencing.
pACBac forward-RS:
5'-CTTCCTGGAAATCAAGAGCCCCATCCT-3'
pACBac reverse-RS:
5'-GTTCCAGGTCGCCGTGCATGATGT-3'

Toxicity Screen of Tetrazines. HEK293T cells plated in a 96-well plate at approximately 10% confluency. Cells were incubated for 48 h with Tet-v3.0 or 1% DMSO, and the cell viability was measured using CellTiter Glo assay kit (Promega) according to the manufacturer's instruction. Briefly, 25 µl of CellTiter Glo reagent was added to each well and incubated for 10 min at RT. The signal was measured for 1 sec using TR717 microplate luminometer (Berthold, Germany) and WinGlow software version 1.25 (Berthold Technologies). The data was normalized to vehicle control and fitted to a curve using non-linear regression method using GraphPad Prism 5. n = 3 ± 95% CI

Transfection of HEK cells. HEK293T cells were plated in a 24-well plate at about 40% confluency so that they reach 70% ~ 90% confluency at the time of transfection. Cells were transfected using Lipofectamine 2000 (Thermo Fisher) using the manufacturer's protocol with minor modification. Briefly, 600 ng of plasmid DNA was diluted in 25 µl of serum free DMEM

and 1.8 µl of Lipofectamine 2000 reagent was diluted in 25 µl of serum free DMEM. They were combined and incubated for 10 min at RT before adding to cells. Tet-v3.0 amino acid was added to the cells immediately and incubated for 24 h ~ 48 h.

Flow Cytometry Assessment. HEK293T cells were transfected as described above with pAcBac1-R2-84-RS or pAcBac1-NES-R2-84-RS and pAcBac1-sfGFP-150TAG for 24 h. GFP expression was confirmed by fluorescence microscope. Adherent cells were washed once with PBS and dissociated into single cells by 0.05% trypsin/0.53 mM EDTA incubation. Cells were washed with PBS and 10,000-20,000 events were collected and analyzed by flow cytometry using CytoFLEX flow cytometer and CytExpert software version 2.2 (Beckman Coulter). GFP signal was collected on FL1-H FITC channel. The basal level of GFP signal was determined from untransfected cells or no ncAA control cells. Any cells that have higher level of GFP than the background level were defined to be a GFP expressing cell. The mean fluorescence intensity (MFI) of GFP was calculated by using CytExpert software and compared between R2-84-RS and NES-R2-84-RS to estimate the effect of the nuclear export signal (NES) on the suppression efficiency.

Confocal Microscopy. HEK293T cells were seeded in a 4-chamber 35 mm glass-bottom dish (Cellvis) and transfected as described above with pAcBac1-NES-R2-84-RS and pAcBac1-sfGFP-150TAG or pAcBac1-WT-sfGFP for 24 h. Cell culture media was replaced three times, and cells were incubated for 30 min with 1 µM TAMRA-sTCO, 1 µM TAMRA-sCCO, 1 µM TAMRA-acid, or 0.1% DMSO at 37 °C. Free label was quenched by Tet-v2.0 Me and nuclei were stained by NucBlue Live ReadyProbes reagent (Hoechst 33342) (ThermoFisher) according to the manufacturer's protocol. Fluorescent images and differential interference contrast (DIC) images were captured using Zeiss LSM 780 confocal microscope system with Plan-Apochromat 40x/1.4 Oil DIC M27 objective lens and ZEN software 2012 SP1 black edition. The captured images were processed and analyzed using ZEN software 3.1 blue edition. The GFP signals were collected from 491 nm to 553 nm using 488 nm laser at 1–2% power. The TAMRA signals were collected from 562 nm to 624 nm using 561 nm laser at 8% power. The Hoechst signals were collected from 410 nm to 495 nm using 405 nm laser at 7% power. The detector gain for GFP, TAMRA, and Hoechst was 500, 500, and 750, respectively.

In-gel Fluorescence Analysis of *in vivo* Protein Labeling. To examine the specific labeling of sfGFP-150-Tet-v3.0Bu protein in HEK293T cells, transfected cells were washed three times with culture media and incubated for 30 min either with 100 nM TAMRA-sTCO, 100 nM TAMRA-sCCO, 100 nM TAMRA-sTCO-Tet-v2.0Me, or 0.1% DMSO (Figure 3D), with increasing concentrations of sTCO (Figure 4A, 4B), or with TAMRA dyes as indicated (Figure 4C, 4D). Unreacted dyes were quenched by at least three molar excess of Tet-v2.0Me. Cells were washed with PBS and lysed in beta-mercaptoethanol-free Laemmli buffer. Without thermal denaturation, the lysates were separated by 12% SDS-PAGE. The fluorescence images were captured using ChemiDoc MP imaging system (Bio-Rad). UV transillumination and the emission filter of 580/120 nm were used to visualized GFP, and green epi illumination and the emission filter of 605/50 nm were used to detect TAMRA labeled protein band.

Densitometry Analysis. The images were analyzed using ImageJ software 1.52a (NIH) to detect and quantify the band. Briefly, the lane was manually defined using rectangle tool and the densitogram was generated using the same tool. The specific protein peak above the base line was

manually set using a straight-line tool. The band intensity was calculated by the software. The signal was normalized to the vehicle control.

GFP-Tet-v3.0Bu Dimerization *in cellulo*: HEK293T cells were transfected for 24 h as described in the transfection section. To evaluate the feasibility of protein dimerization, transfected cells were washed three times with DMEM supplemented with 10% FBS and penicillin/streptomycin followed by 30 min incubation with dh-sTCO. Unreacted dimerization agent was quenched by incubating with excessive amount of Tet-v2.0Me for 10 min at 37 °C. Crude cell lysates were prepared and analyzed as described in the in-gel fluorescence assay section. Equal amount of protein loading was shown by staining the gel with Coomassie Brilliant Blue R-250.

Gene and Protein Sequence:

GFP-wt (protein):
MVSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTLKFICTTGKLPVPWP
TLVTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGTYKTRAEVKFEG
DTLVNRIELKGIDFKEDGNILGHKLEYNFNSHNVYITADKQKNGIKANFKIRHNVEDGSV
QLADHYQQNTPIGDGPVLLPDNHYLSTQSVLSKDPNEKRDHMVLLEFVTAAGITHGMD
ELYKGSHHHHHH
GFP-wt (DNA):
ATGGTTAGCAAAGGTGAAGAACTGTTTACCGGCGTTGTGCCGATTCTGGTGGAACTG
GATGGTGATGTGAATGGCCATAAATTTAGCGTTCGTGGCGAAGGCGAAGGTGATGC
GACCAACGGTAAACTGACCCTGAAATTTATTTGCACCACCGGTAAACTGCCGGTTCC
GTGGCCGACCCTGGTGACCACCCTGACCTATGGCGTTCAGTGCTTTAGCCGCTATCC
GGATCATATGAAACGCCATGATTTCTTTAAAAGCGCGATGCCGGAAGGCTATGTGCA
GGAACGTACCATTAGCTTCAAAGATGATGGCACCTATAAAACCCGTGCGGAAGTTA
AATTTGAAGGCGATACCCTGGTGAACCGCATTGAACTGAAAGGTATTGATTTTAAAG
AAGATGGCAACATTCTGGGTCATAAACTGGAATATAATTTCAACAGCCATAATGTGT
ATATTACCGCCGATAAACAGAAAAATGGCATCAAAGCGAACTTTAAAATCCGTCAC
AACGTGGAAGATGGTAGCGTGCAGCTGGCGGATCATTATCAGCAGAATACCCCGAT
TGGTGATGGCCCGGTGCTGCTGCCGGATAATCATTATCTGAGCACCCAGAGCGTTCT
GAGCAAAGATCCGAATGAAAAACGTGATCATATGGTGCTGCTGGAATTTGTTACCG
CCGCGGGCATTACCCACGGTATGGATGAACTGTATAAAGGCAGCCACCATCATCAT
CACCAT

R2-84 PylRS (Protein):
MDKKPLDVLISATGLWMSRTGTLHKIKHHEVSRSKIYIEMACGDHLVVNNSRSCRTAR
AFRHHKYRKTCKRCRVSDEDINNFLTRSTESKNSVKVRVVSAPKVKKAMPKSVSRAPK
PLENSVSAKASTNTSRSVPSPAKSTPNSSVPASAPAPSLTRSQLDRVEALLSPEDKISLNM
AKPFRELEPELVTRRKNDFQRLYTNDREDYLGKLERDITKFFVDRGFLEIKSPILIPAEYV
ERMGINNDTELSKQIFRVDKNLCLCPMLAPTGYNYLRKLDRILPGPIKIFEVGPCYRKES
DGKEHLEEFTMVGFAQMGSGCTRENLEALIKEFLDYLEIDFEIVGDSCMVYGDTLDIMH
GDLELSSAVVGPVSLDREWGIDKPWIGAGFGLERLLKVMHGFKNIKRASRSESYYNGIS
TNL

R2-84 PylRS (DNA):

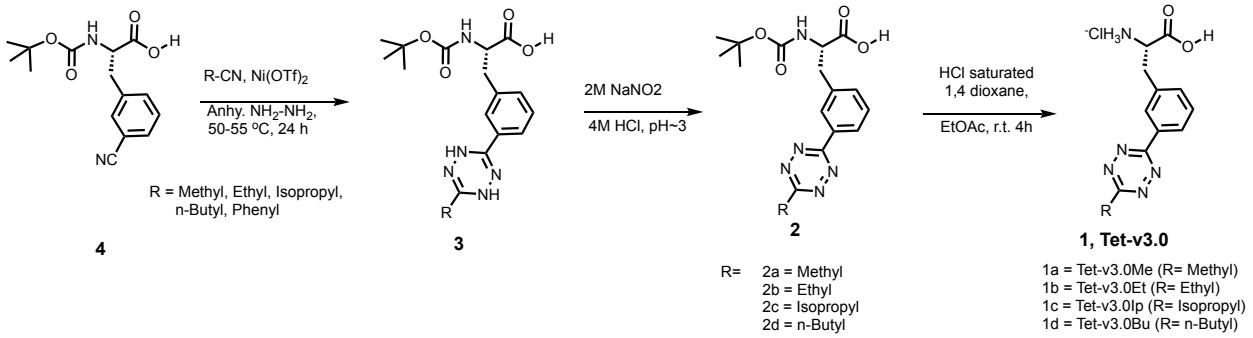
ATGGATAAAAAACCGCTGGATGTGCTGATTAGCGCGACCGGCCTGTGGATGAGCCG
TACCGGCACCCTGCATAAAATCAAACATCATGAAGTGAGCCGCAGCAAAATCTATA
TTGAAATGGCGTGCGGCGATCATCTGGTGGTGAACAACAGCCGTAGCTGCCGTACC
GCGCGTGCGTTTCGTCATCATAAATACCGCAAAACCTGCAAACGTTGCCGTGTGAGC
GATGAAGATATCAACAACTTTCTGACCCGTAGCACCGAAAGCAAAAACAGCGTGAA
AGTGCGTGTGGTGAGCGCGCCGAAAGTGAAAAAAGCGATGCCGAAAAGCGTGAGC
CGTGCGCCGAAACCGCTGGAAAATAGCGTGAGCGCGAAAGCGAGCACCAACACCA
GCCGTAGCGTTCCGAGCCCGGCGAAAAGCACCCCGAACAGCAGCGTTCCGGCGTCT
GCGCCGGCACCGAGCCTGACCCGCAGCCAGCTGGATCGTGTGGAAGCGCTGCTGTC
TCCGGAAGATAAAATTAGCCTGAACATGGCGAAACCGTTTCGTGAACTGGAACCGG
AACTGGTGACCCGTCGTAAAAACGATTTTCAGCGCCTGTATACCAACGATCGTGAAG
ATTATCTGGGCAAACTGGAACGTGATATCACCAAATTTTTTGTGGATCGCGGCTTTC
TGGAAATTAAAAGCCCGATTCTGATTCCGGCGGAATATGTGGAACGTATGGGCATTA
ACAACGACACCGAACTGAGCAAACAAATTTTCCGTGTGGATAAAAACCTGTGCCTG
TGTCCGATGCTGGCCCCGACCGGTTATAACTATTTGCGTAAACTGGATCGTATTCTG
CCGGGTCCGATCAAAATTTTTGAAGTGGGCCCGTGCTATCGCAAAGAAAGCGATGG
CAAAGAACACCTGGAAGAATTCACCATGGTTGGTTTTGCTCAAATGGGCAGCGGCT
GCACCCGTGAAAACCTGGAAGCGCTGATCAAAGAATTCCTGGATTATCTGGAAATC
GACTTCGAAATTGTGGGCGATAGCTGCATGGTGTATGGCGATACCCTGGATATTATG
CATGGCGATCTGGAACTGAGCAGCGCGGTGGTGGGTCCGGTTAGCCTGGATCGTGA
ATGGGGCATTGATAAACCGTGGATTGGCGCGGGTTTTGGCCTGGAACGTCTGCTGAA
AGTGATGCATGGCTTCAAAAACATTAAACGTGCGAGCCGTAGCGAAAGCTACTATA
ACGGCATTAGCACGAACCTGTAA

NES-R2-84 PylRS (protein):
MACPVPLQLPPLERLTLDDYKDDDDKDKKPLDVLISATGLWMSRTGTLHKIKHHEVSR
SKIYIEMACGDHLVVNNSRSCRTARAFRHHKYRKTCKRCRVSDEDINNFLTRSTESKNS
VKVRVVSAPKVKKAMPKSVSRAPKPLENSVSAKASTNTSRSVPSPAKSTPNSSVPASAP
APSLTRSQLDRVEALLSPEDKISLNMAKPFRELEPELVTRRKNDFQRLYTNDREDYLGKL
ERDITKFFVDRGFLEIKSPILIPAEYVERMGINNDTELSKQIFRVDKNLCLRPMLAPTGYN
YLRKLDRILPGPIKIFEVGPCYRKESDGKEHLEEFTMVGFAQMGSGCTRENLEALIKEFL
DYLEIDFEIVGDSCMVYGDTLDIMHGDLELSSAVVGPVSLDREWGIDKPWIGAGFGLER
LLKVMHGFKNIKRASRSESYYNGISTNL

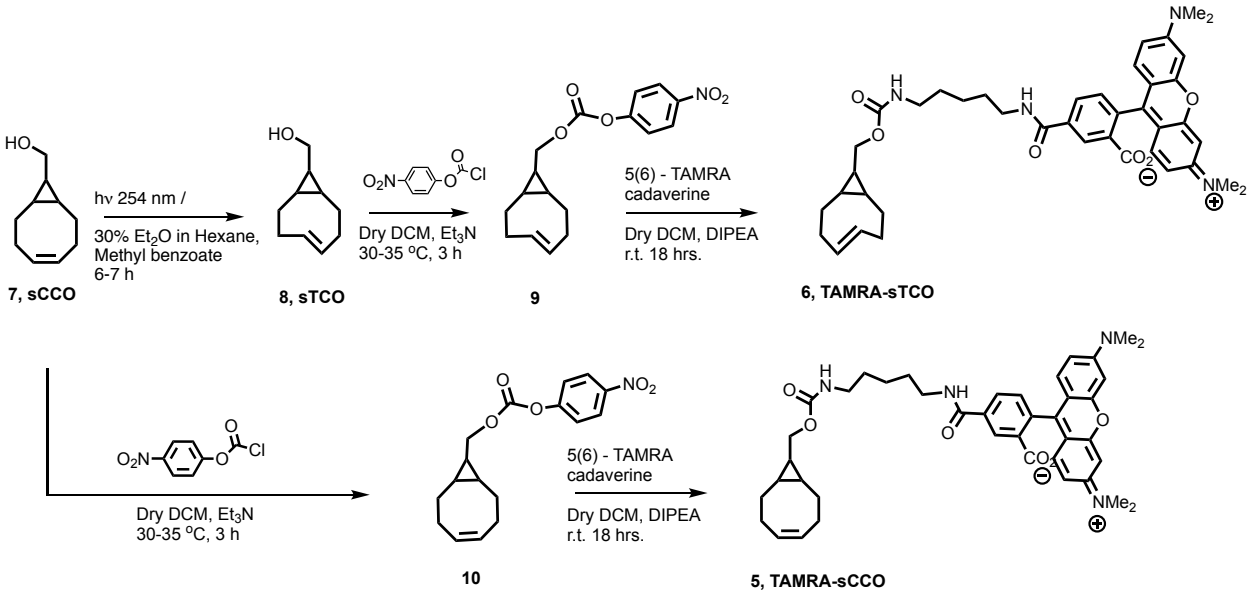
NES-R2-84 PylRS (DNA):
ATGGCGTGTCCGGTTCCTTTGCAGTTGCCTCCACTGGAGCGCCTCACACTCGACGAC
TACAAGGACGACGACGACAAGGACAAGAAACCCCTGGACGTGCTGATCAGCGCCAC
CGGCCTGTGGATGAGCCGGACCGGCACCCTGCACAAGATCAAGCACCACGAGGTGT
CAAGAAGCAAAATCTACATCGAGATGGCCTGCGGCGACCACCTGGTGGTGAACAAC
AGCAGAAGCTGCCGGACCGCCAGAGCCTTCCGGCACCACAAGTACAGAAAGACCTG
CAAGCGGTGCCGGGTGTCCGACGAGGACATCAACAACTTTCTGACCAGAAGCACCG
AGAGCAAGAACAGCGTGAAAGTGCGGGTGGTGTCCGCCCCCAAAGTGAAGAAAGC
CATGCCCAAGAGCGTGTCCAGAGCCCCCAAGCCCCTGGAAAACAGCGTGTCCGCCA
AGGCCAGCACCAACACCAGCCGCAGCGTGCCCAGCCCCGCCAAGAGCACCCCCAAC
AGCTCCGTGCCCGCCTCTGCTCCTGCTCCCAGCCTGACACGGTCCCAGCTGGACAGA
GTGGAGGCCCTGCTGTCCCCCGAGGACAAGATCAGCCTGAACATGGCCAAGCCCTT

CCGGGAGCTGGAACCCGAGCTGGTGACCCGGCGGAAGAACGACTTCCAGCGGCTGT
ACACCAACGACCGGGAGGACTACCTGGGCAAGCTGGAACGGGACATCACCAAGTTC
TTCGTGGACCGGGGCTTCCTGGAAATCAAGAGCCCCATCCTGATCCCCGCCGAGTAC
GTGGAGCGGATGGGCATCAACAACGACACCGAGCTGTCCAAGCAGATTTTCCGGGT
GGACAAGAACCTGTGCCTGCGGCCTATGCTGGCCCCCACCGGCTACAACTACCTGCG
GAAACTGGACAGAATCCTGCCTGGCCCCATCAAGATTTTCGAAGTGGGACCCTGCTA
CCGGAAAGAGAGCGACGGCAAAGAGCACCTGGAAGAGTTTACAATGGTGGGCTTTG
CCCAGATGGGCAGCGGCTGCACCCGGGAGAACCTGGAAGCCCTGATCAAAGAGTTC
CTGGATTACCTGGAAATCGACTTCGAGATCGTGGGCGACAGCTGCATGGTGTACGGC
GACACCCTGGACATCATGCACGGCGACCTGGAACTGAGCAGCGCCGTGGTGGGACC
CGTGTCCCTGGACCGGGAGTGGGGCATCGACAAGCCCTGGATCGGAGCCGGCTTCG
GCCTGGAACGGCTGCTGAAAGTGATGCACGGCTTCAAGAACATCAAGCGGGCCAGC
AGAAGCGAGAGCTACTACAACGGCATCAGCACCAACCTGTGA

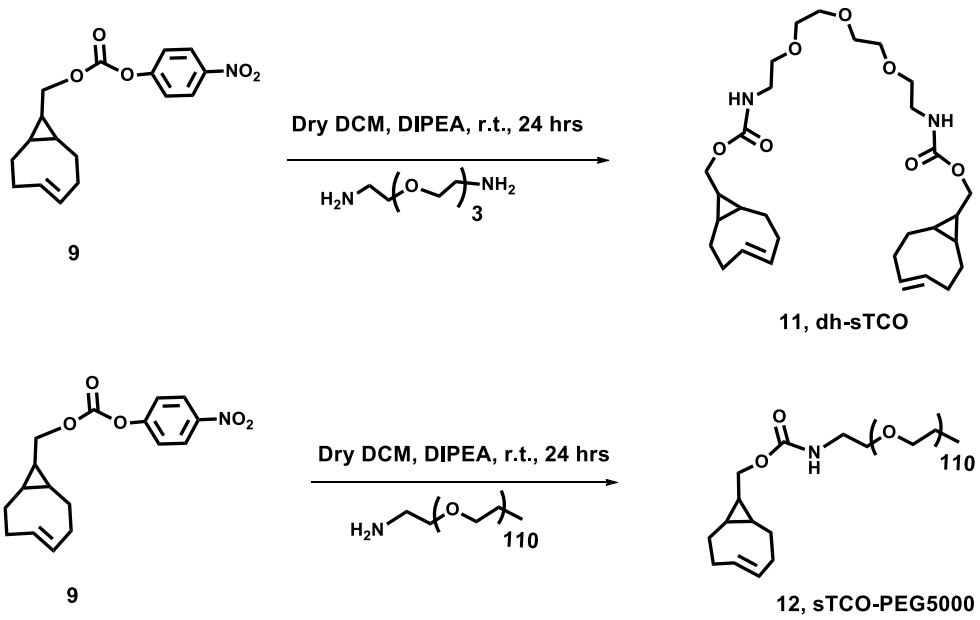
Schemes S1-S3



Scheme S1. Synthesis of s-tetrazine derivatives of phenylalanine (Tet-v3.0).

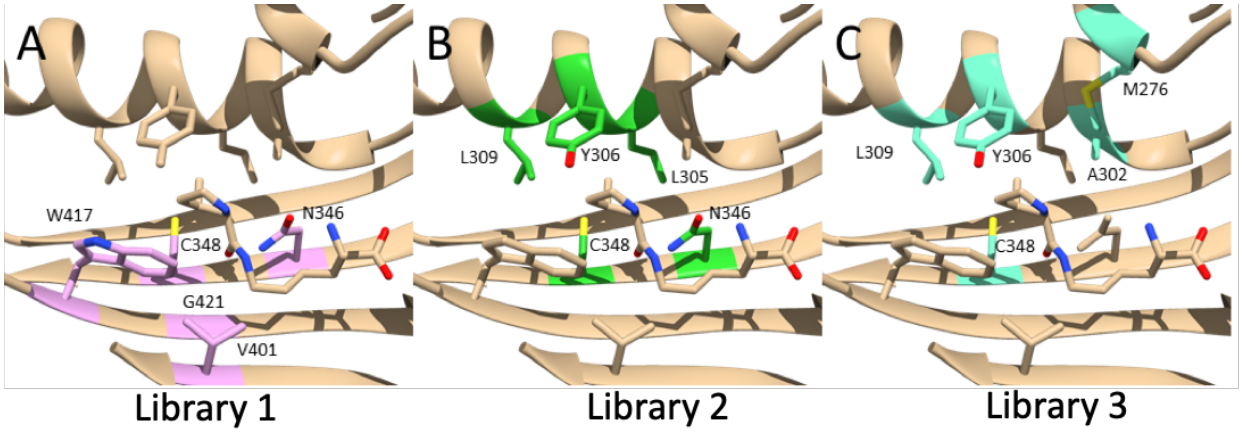


Scheme S2. Synthesis of TAMRA linked strained alkenes (TAMRA-sCCO and TAMRA-sTCO).



Scheme S3. Synthesis of dh-sTCO and sTCO-Peg-5000.

Supplemental Figures S1-S43



D

Library	Library Residues (M. mazei / M. barkeri)				
1	N346/N311	C348/C313	V401/V366	W417/W382	G421/G386
2	L305/L270	Y306/Y271	L309/L274	N346/N311	C348/C313
3	M276/M241	A302/A267	Y306/Y271	L309/L274	C348/C313

Figure S1. Amino acid libraries of PylRS active-sites. A) Structure of *M. mazei* PylRS with mutants of library 1. B) Structure of *M. mazei* PylRS with mutants of library 2. C) Structure of *M. mazei* PylRS with mutants of library 3 (D) mutants of the PylRS libraries showing amino acid

numbering for *M. mazei / M. barkeri.* The active-site library sites were selected for maximum direct interaction with the side chain group of the Tet-ncAA.

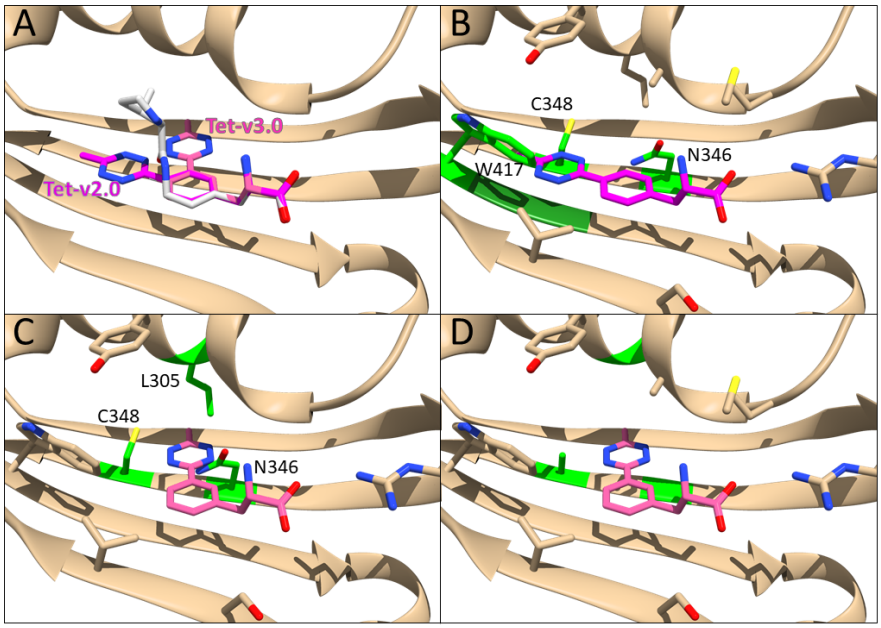
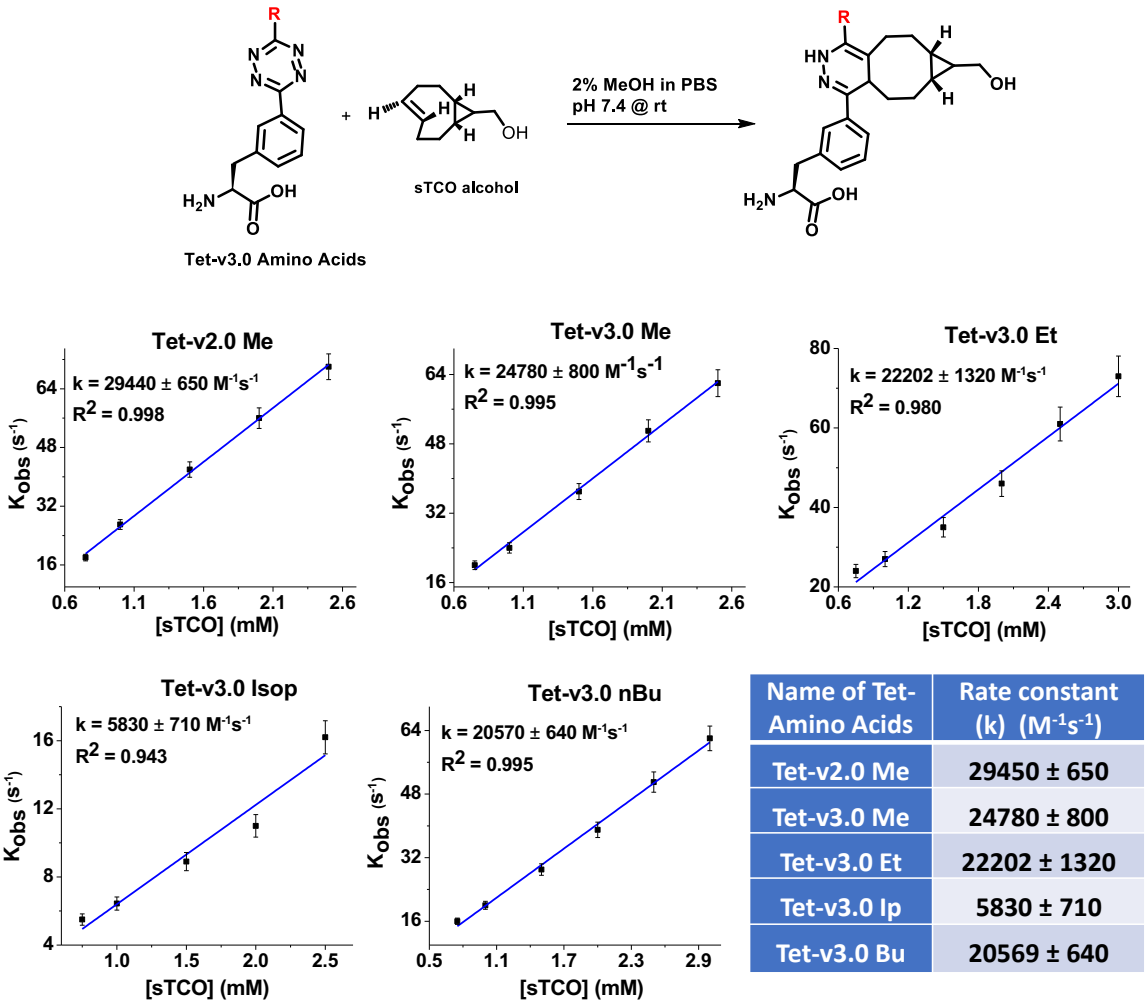
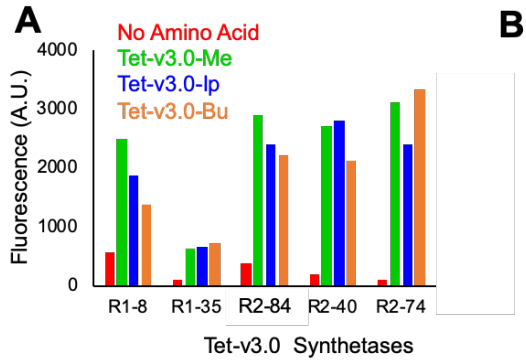


Figure S2. Assessment of how different Tetrazine Amino Acids will be accommodated by PylRS active-site. A) Tet-v2.0Me and Tet-v3.0Me superimposed upon Pyrrolysine (white) in the PylRS active-site (2ZCE). B) Potential clashes between Tet-v2.0Me and the side chains of active site residues (light green) as well as potential clashes with the backbone of active site residues (dark green). C) Potential clashes between Tet-v3.0Me and the side chains of active site residues (light green). D) Mutations found in R2-84 synthetase (light green) open the active site to accommodate Tet-v3.0Me.



Name of Tet-Amino Acids	Rate constant (k) ($M^{-1}s^{-1}$)
Tet-v2.0 Me	29450 ± 650
Tet-v3.0 Me	24780 ± 800
Tet-v3.0 Et	22202 ± 1320
Tet-v3.0 Ip	5830 ± 710
Tet-v3.0 Bu	20569 ± 640

Figure S3. **Measurement of reaction rates for Tet-v3.0 with sTCO.** The loss tetrazine signal at 270 nm was measured over time after the addition of sTCO at different concentrations. To measure the kinetics, the solutions of tetrazine amino acids and sTCO were made in phosphate buffer saline 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.4) with 2% methanol to ensure all compounds were completely soluble. All measurements were performed in triplicate and the resulted decay curves were fit to a single exponential equation. Plot of pseudo first order rate constant (k´) against concentration of sTCO to determine the second order rate constant (k) of cycloaddition reaction of Tet-v3.0 derivatives with trans-cyclooctene (sTCO).



B

Site (Mb/Mm)	R1-8	R1-35	R2-84	R2-40	R2-74
L270/L305	-	-	G	-	G
Y271/Y306	-	-	-	-	
L274/L309	-	-	-	-	-
N311/N346	A	A	G	G	G
C313/C348	S	V	A	A	S
Other (Mb)	-	-	-	K296R	-

Figure S4. Standard selection methods with Tet-v3.0Me resulted in six unique top RS/tRNA pairs. A) Efficiency and fidelity of the selected RSs towards the Tet-v3.0 amino acids as measured by fluorescence of TAG-suppressed GFP in the presence or absence of supplemented Tet-v3.0. at 1.0 mM Tet-v3.0 B) Sequence differences of the selected RSs. R2-74 and R2-84 were chosen for future use.

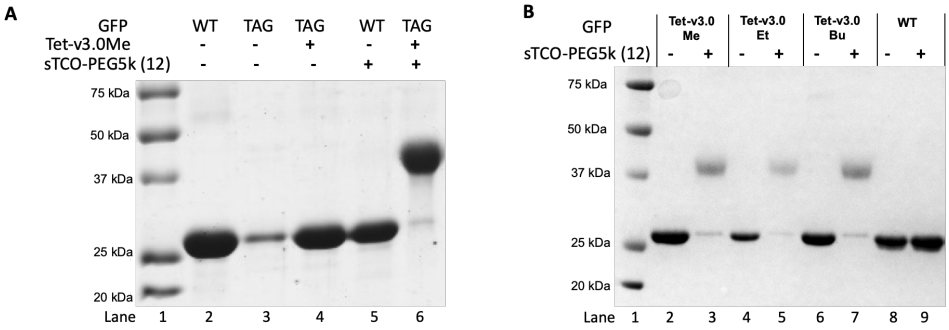


Figure S5. Assessment of purified GFP-Tet-v3.0. Purification of his-tag protein from DH10B cells grown in autoinduction media cotransformed with the pDule-R2-84 and pBad-GFP-150(TAG) or pBad-GFP(WT) plasmids. A) Samples GFP-wt and minus Tet3.0 were grown in the absence of Tet-v3.0 and plus Tet3.0 contained 1.0 mM Tet-v3.0Me. For lanes 5 and 6, purified GFP-wt and GFP-Tet-v3.0Me were incubated with 5 eq. of sTCO-PEG5000, **12**, for 5 minutes in PBS prior to adding SDS load buffer. B) Tet3.0 ncAA samples were grown in the presence of 1.0 mM Tet-v3.0-ncAA, GFP-wt was grown in the absence of ncAA. For lanes 3, 5, 7 and 9, purified GFP-Tet-v3.0-ncAA and GFP-wt protein were incubated with 5 eq. of sTCO-PEG5000, **12**, for 5 minutes in PBS prior to adding SDS load buffer. The clear mobility shift of the Tet-v3.0 containing protein and not GFP-wt confirms reactivity of encoded Tet-v3.0 ncAAs.

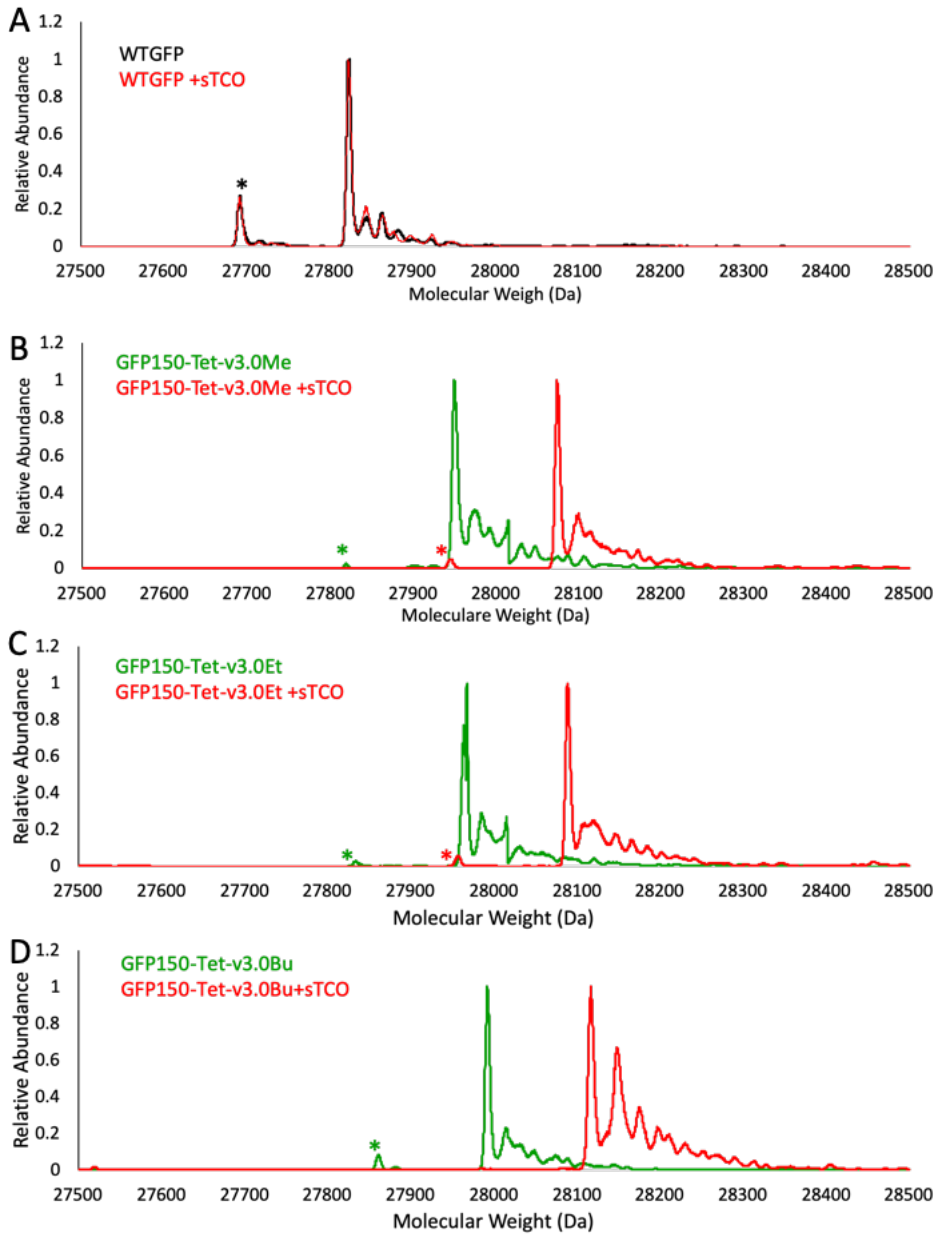


Figure S6. ESI mass spectroscopy analysis of GFP-Tet-v3.0 and reactions with sTCO. Purified GFP-wt (black) and GFP-Tet-v3.0 (green) (Figure S5) show the appropriate mass increase for the site-specific incorporation of Tet-v3.0Me (B), Tet-v3.0Et (B), Tet-v3.0Bu (D). Purified proteins were exposed to a 10-fold molar excess of sTCO for 5 minutes (red). GFP-wt undergoes no mass change when exposed to sTCO (A) and each Tet-v3.0 protein showed the expected 124 Da increase in mass corresponding to the addition of sTCO and loss of molecular nitrogen. No unreacted GFP150-Tet-v3.0 was detected, verifying the reaction of genetically encoded Tet3.0 amino acids with sTCO was quantitative. The lower mass peak labeled with * is a loss of n-terminal methionine and upper mass peaks are salt sodium and potassium adducts. GFP-wt predicted: 27827.02 Da avg, observed 27827.49 Da avg; GFP-Tet-v3.0Me predicted: 27954.17 Da avg, observed: 27954.54 Da avg; GFP-Tet-v3.0Me+sTCO predicted: 28078.39 Da avg, observed:

28078.61 Da avg). GFP-Tet-v3.0Et predicted: 27968.20 Da avg, observed: 27967.2 Da avg; GFP-Tet-v3.0Et+sTCO predicted: 28092.42 Da avg, observed: 28099.5 Da avg). GFP-Tet-v3.0Bu predicted: 27996.25 Da avg, observed: 27996.3 Da avg; GFP-Tet-v3.0Bu+sTCO predicted: 28120.47 Da avg, observed: 28121.3 Da avg).

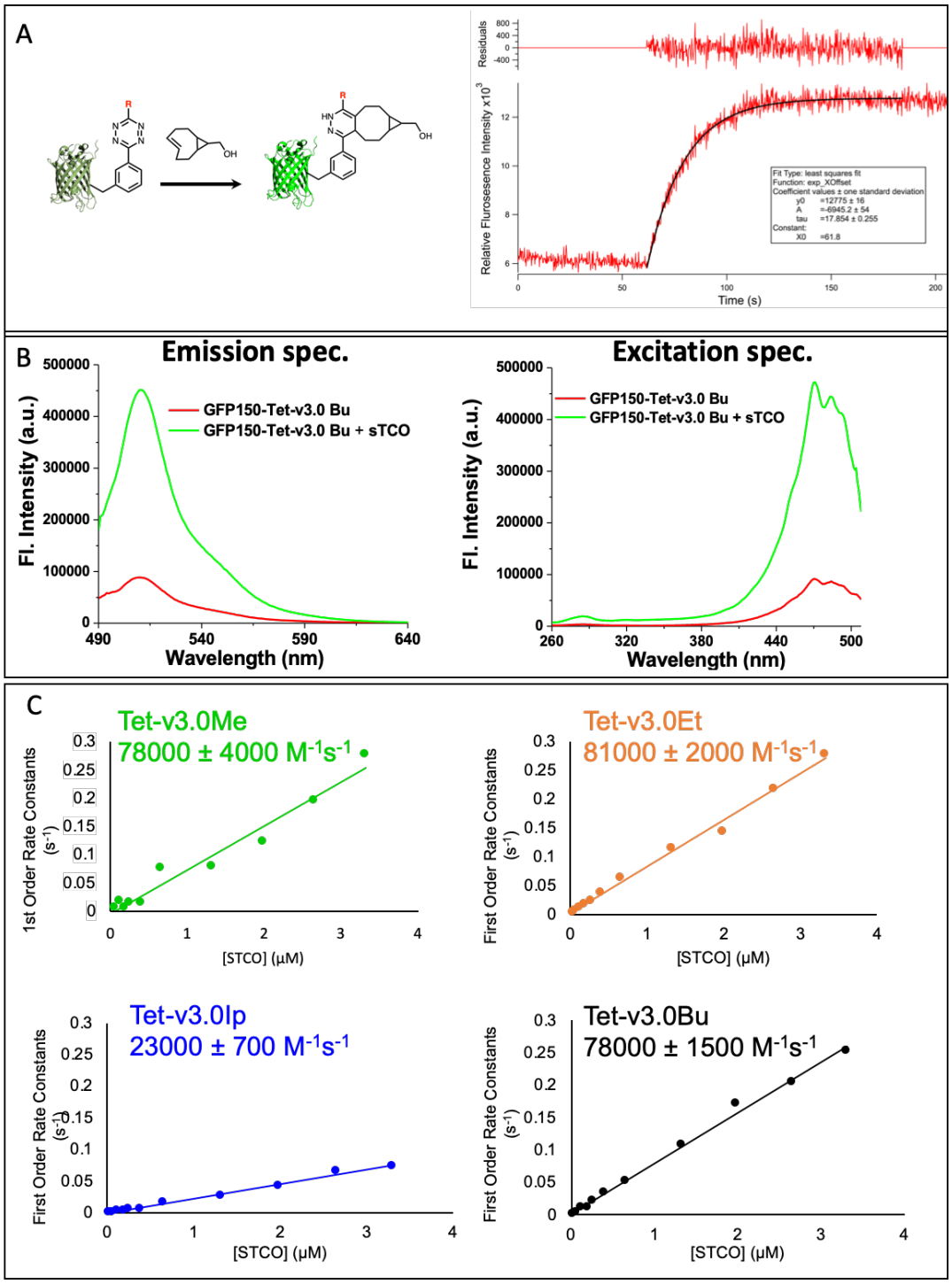


Figure S7. Measurement of reaction rates for GFP-Tet-v3.0 with sTCO. A) the GFP fluorescence is quenched by the addition of tetrazine ncAA at site 150 and the GFP fluorescence return upon reaction with sTCO. B) Fluorescence emission (λ ex 488 nm) and excitation (λ em 510 nM) spectra of the GFP-Tet-v3.0Bu before and after the tetrazine ligation reaction (50 nM GFP-Tet protein reacted with excess (10 eqv.) sTCO alcohol in phosphate buffer saline 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.4). C) Plot of pseudo first order rate

constant (k′) against concentration of sTCO to determine the second order rate constant (k) for reaction of GFP-v3.0Tet with sTCO.

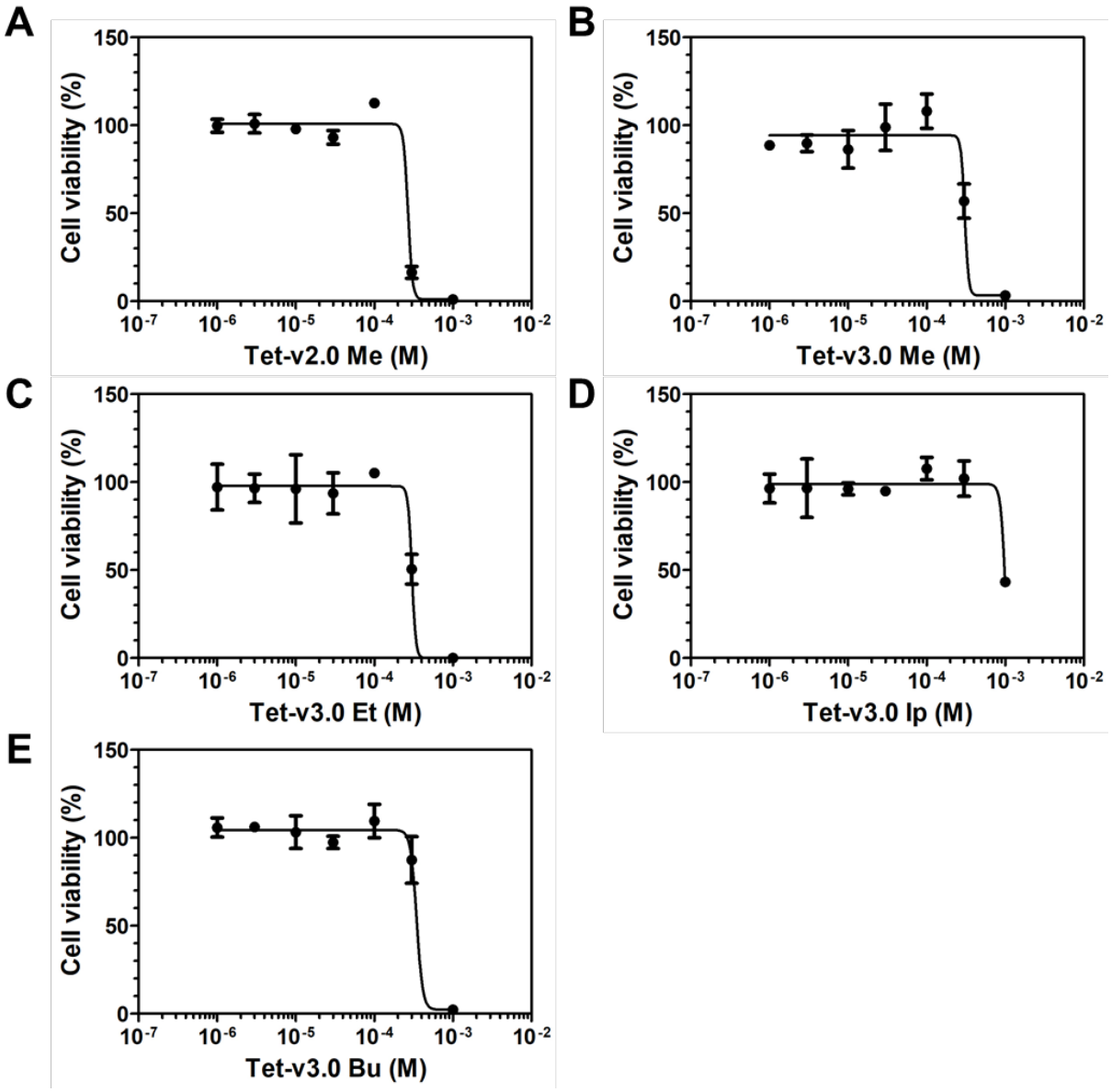
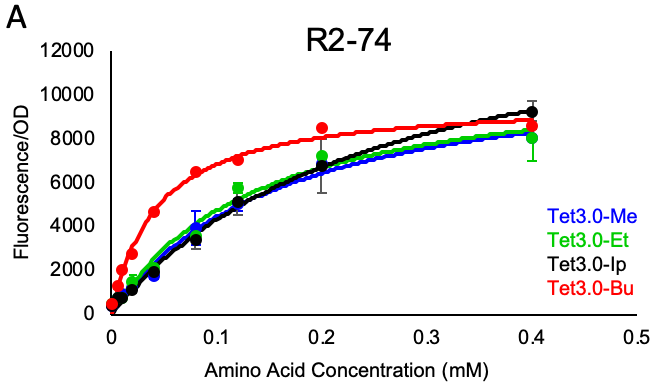


Figure S8. The effect of tetrazine amino acid concentration on mammalian cell viability was measured. HEK293T cells were incubated for 48 h with (A) Tet-v2.0 Me, (B) Tet-v3.0 Me, (C) Tet-v3.0 Et, (D) Tet-v3.0 Ip, (E) Tet-v3.0 Bu, or 1% dimethyl sulfoxide. Cell viability was compromised above 0.1 mM tetrazine amino acid. The cell viability was examined by using the CellTiter Glo assay kit, and the data was normalized using the vehicle control. The data was fit to a curve by nonlinear regression using GraphPad Prism software version 5. n=3 with error bars showing the 95% confidence interval.

A



B

Amino Acid	R2-84 (μM)	R2-74 (μM
Tet-v3.0Me	25.7	163.5
Tet-v3.0Et	17.4	140.0
Tet-v3.0Ip	32.4	251.0
Tet-v3.0Bu	13.5	43.5

B

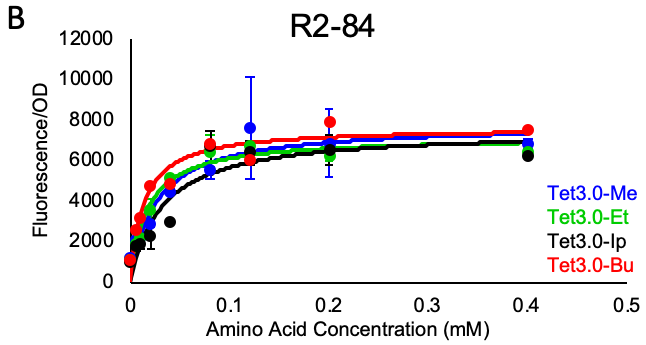


Figure S9: Tet-v3.0 concentration depend expression of GFP-Tet-v3.0. A) Expression of GFP-Tet-v3.0 with the R2-74-RS/tRNA pair in *E. coli*. B) Expression of GFP-Tet-v3.0 with the R2-84-RS/tRNA pair in *E. coli*. C) The concentration of amino acid to reach one half maximum protein yield was calculated for each Tet-v3.0 amino acid. The R2-84 synthetase was found to be able to produce more GFP at lower concentrations of Tet-v3.0-Bu than the other Tet-v3.0 amino acids.

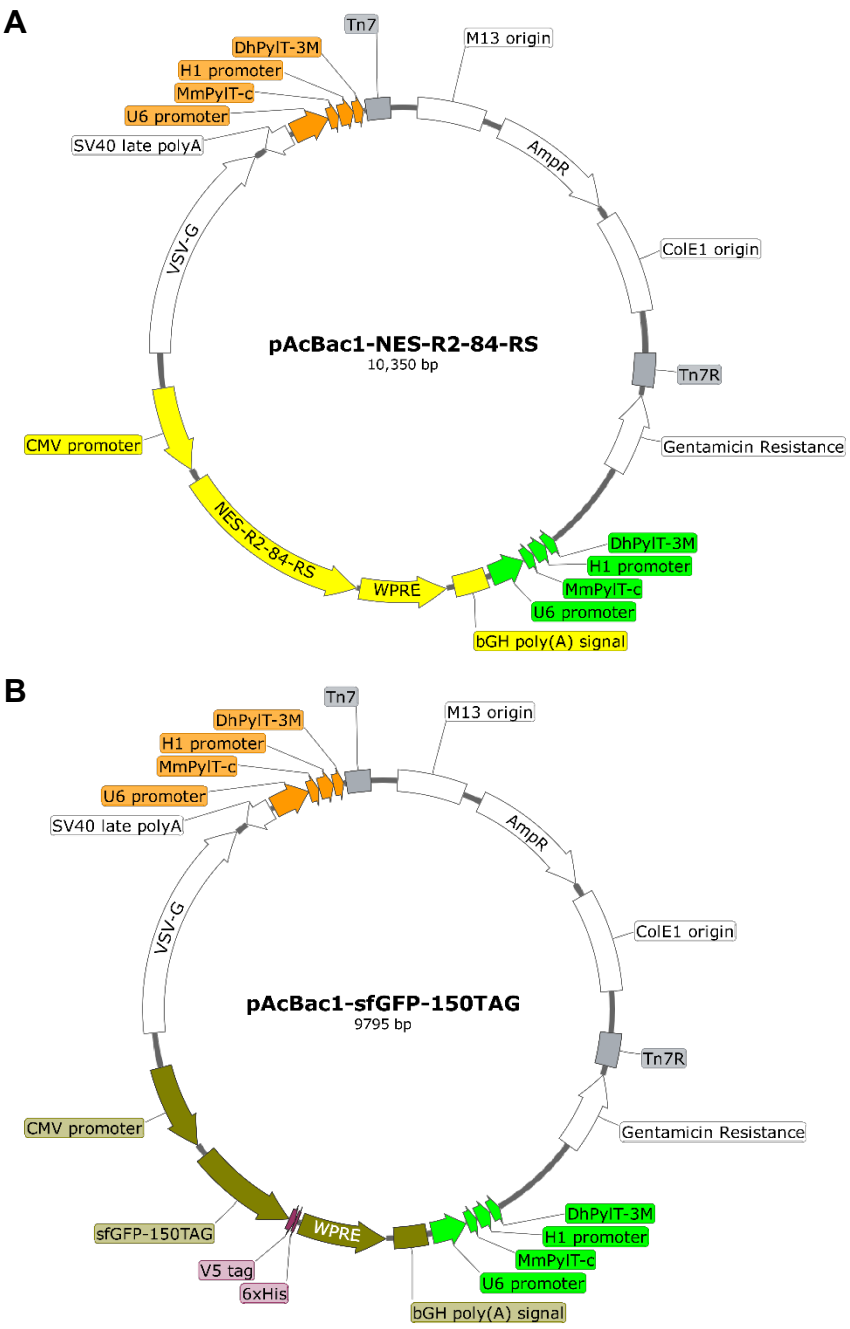


Figure S10. The plasmid maps of mammalian expression systems. (A) pAcBac1-NES-R2-84-RS (C) pAcBac1-sfGFP-150TAG.

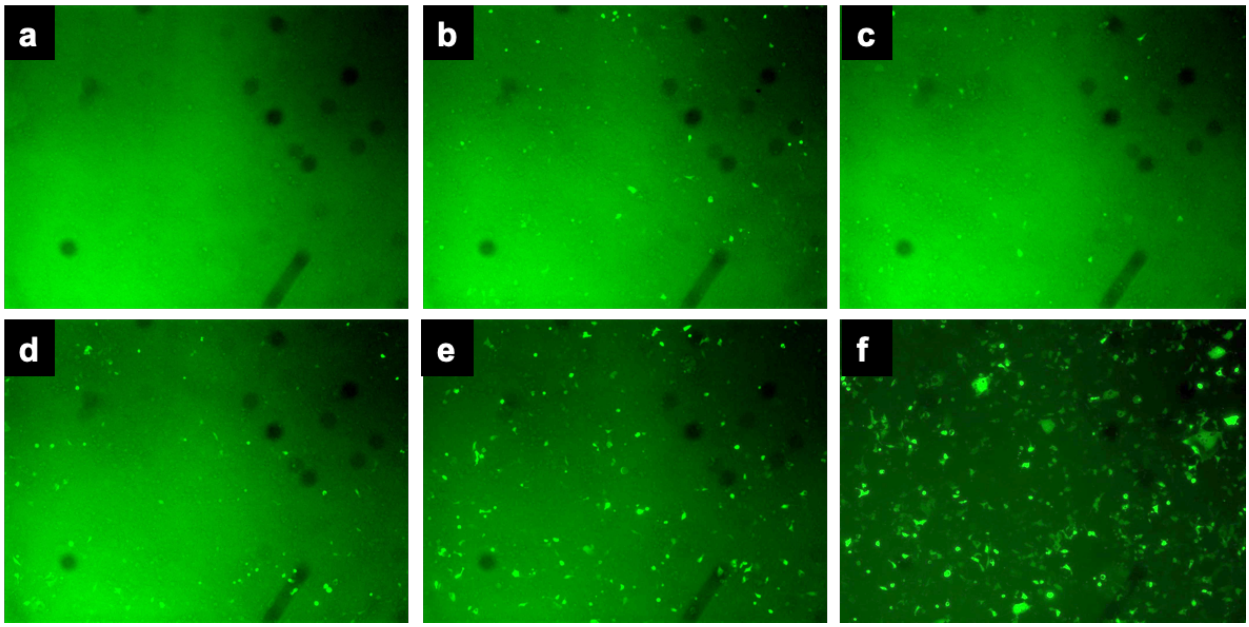


Figure S11. Effect of Tet-v3.0 amino acids on GFP-TAG suppression in HEK293 cells. HEK293 cells were transfected for 72 h with pAcBac1-R2-84-RS and pAcBac1-sfGFP-150TAG in the presence of 0.1% DMSO (a), 0.1 mM Tet-v3.0Me (b), 0.1 mM Tet-v3.0Et (c), 0.1 mM Tet-v3.0Ip (d), or 0.1 mM Tet-v3.0Bu (e) using Lipofectamine2000. sfGFP-wt was used as a positive control (f). The TAG suppression efficiency was estimated by examining the expression of GFP protein on epifluorescence microscope.

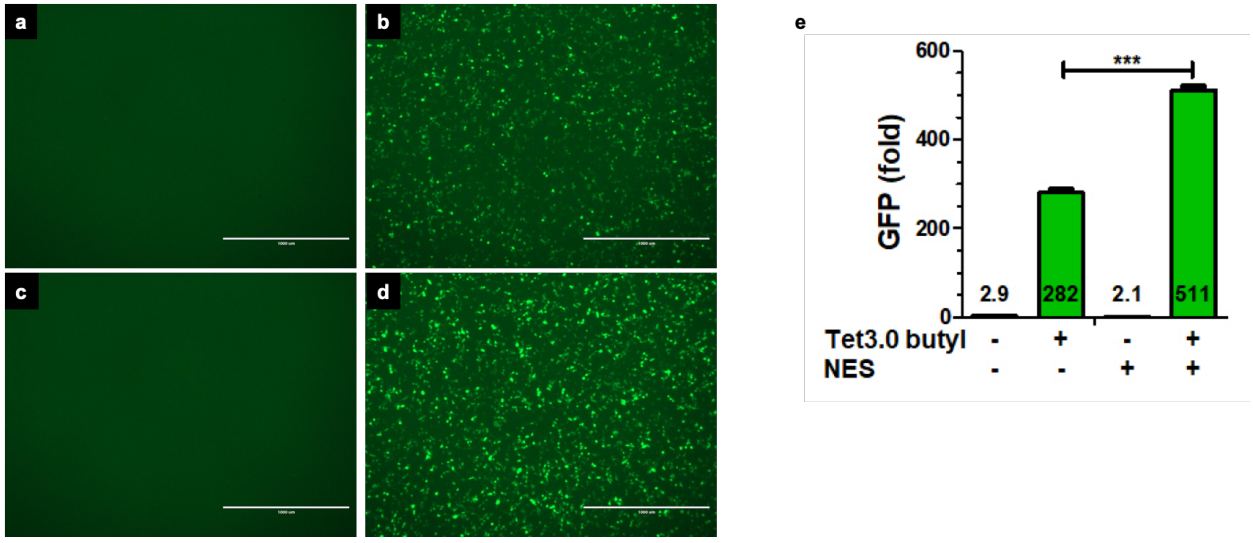


Figure S12. Effect of nuclear export signal on TAG suppression efficiency in HEK293T. HEK293T cells were transfected with pAcBac1-R2-84-RS (a, b) or pAcBac-NES-R2-84-RS (c, d) along with pAcBac1-sfGFP-150TAG in the presence of 30 μM Tet-v3.0Bu (b, d) or 0.1% DMSO (a, c). After 18 h, GFP fluorescence images were captured using EVOS FL imaging system. (e) Cells were analyzed by flow cytometry and mean fluorescence intensity was normalized to untransfected control. Scale bar, 1 mm

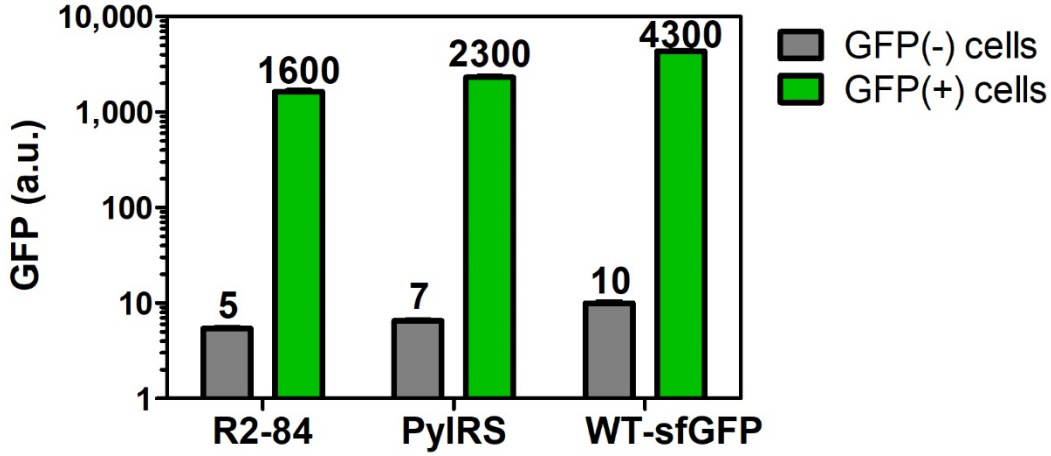


Figure S13. Eukaryotic comparison ncAA-RS suppression ability for R2-84/tRNA with Tet-v3.0Bu and Wt-Pyl-RS/tRNA with Boc-lysine. HEK293T cells were transfected with pAcBac1-NES-R2-84-RS or pAcBac1-NES-WT-pyl-RS along with pAcBac1-sfGFP-150TAG in the presence of 30 μM Tet-v3.0Bu or 1 mM Boc-lysine or 0.1% DMSO for minus ncAA cells. As a positive control WT-sfGFP was transfected. Cells were washed and incubated for 30 min with 30 μM sTCO to un-quench GFP fluorescence. Cells were analyzed by flow cytometry and mean fluorescence intensity of GFP-expressing or GFP-non-expressing cells was calculated.

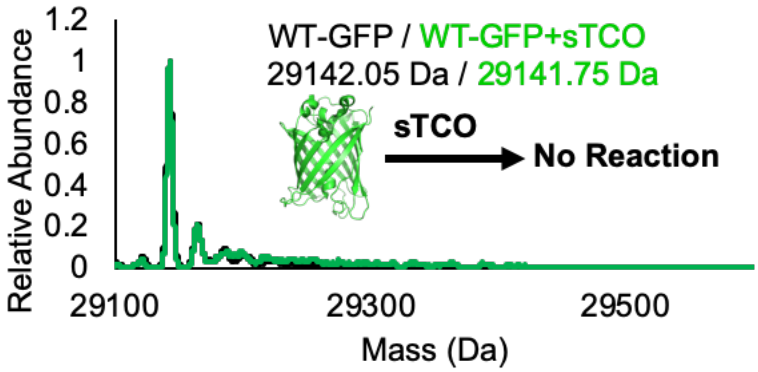


Figure S14. Assessment of GFP-wt mass and reactivity sTCO by ESI-Q-TOF mass spectrometry. GFP produced in HEK293T cells had an observed average mass of 29142.05 Da (predicted 29141.49 Da). GFP-wt exposed to sTCO and no mass change (observed 29141.75 Da was detected verifying sTCO is not reactive to GFP-wt.

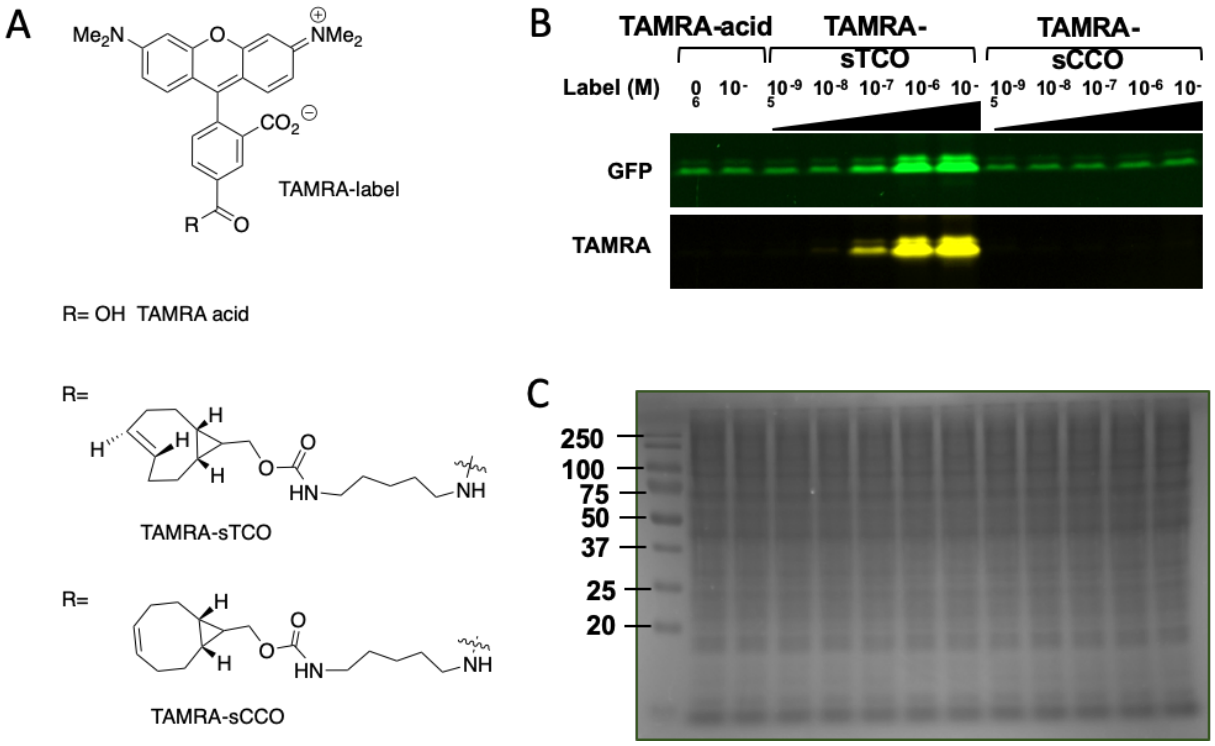
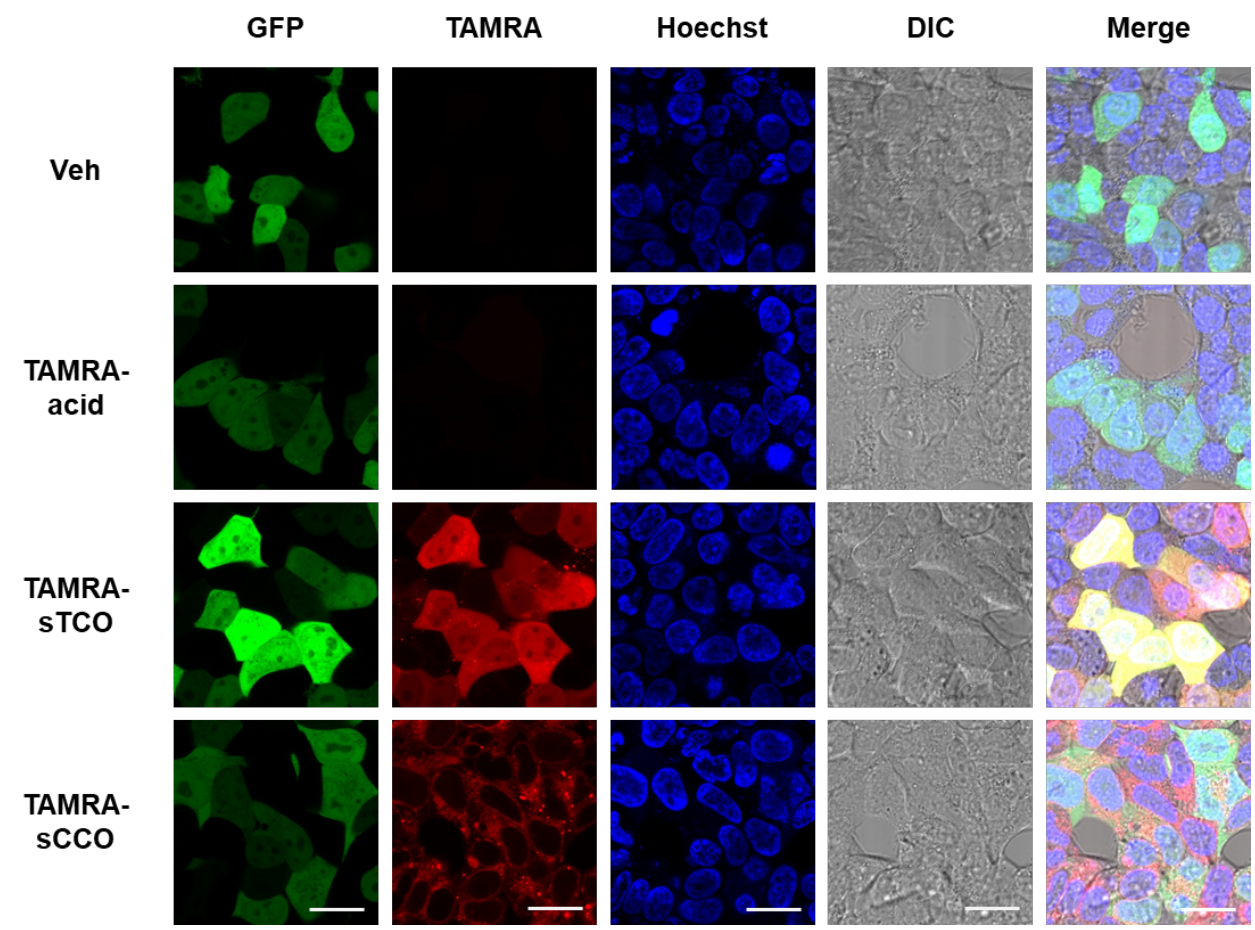
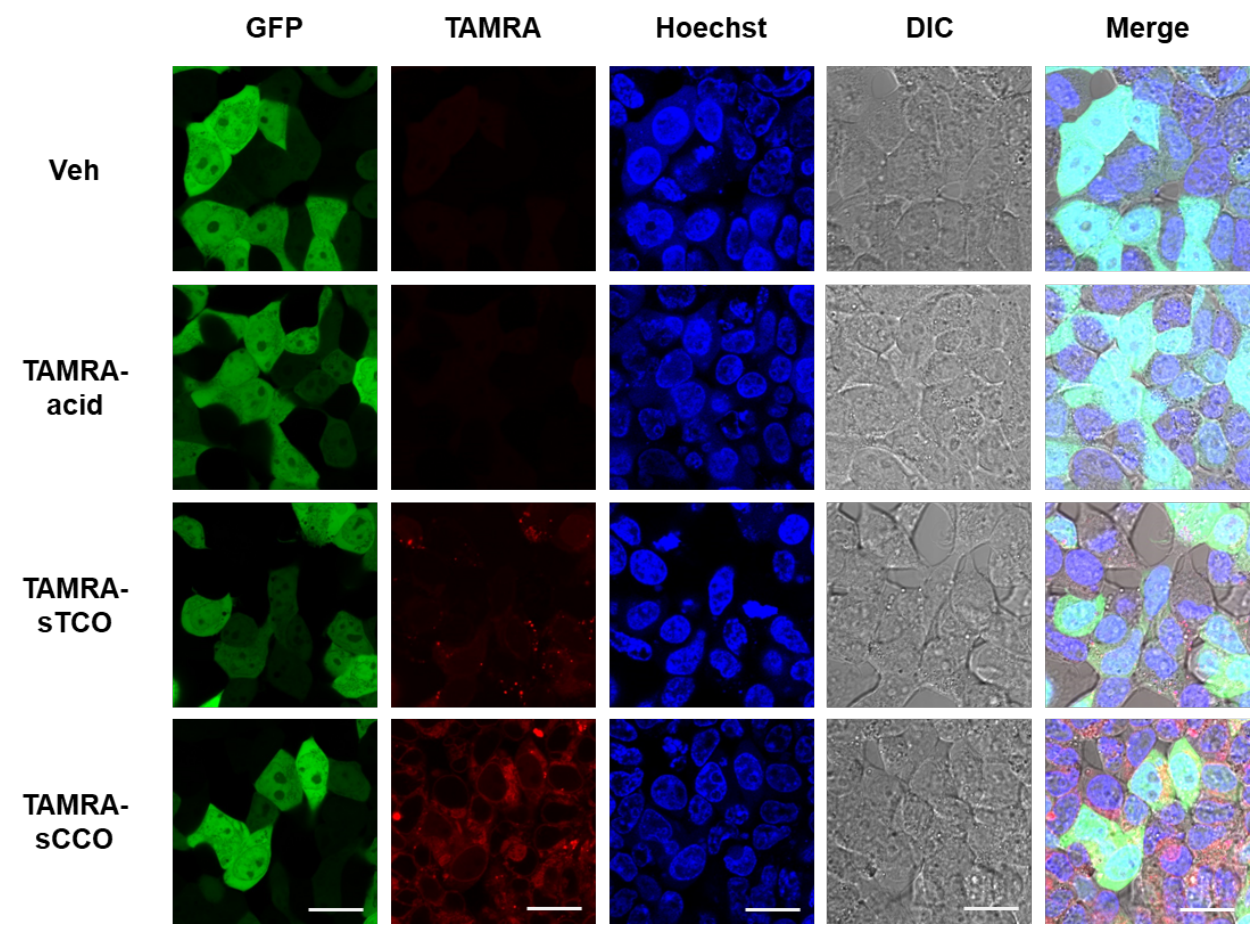


Figure S15. Labeling of GFP-Tet-v3.0Bu in eukaryotic cells with TAMRA-sTCO label. A) Dyes incubated with HEK293T cells, TAMRA acid, TAMRA-sTCO, and TAMRA-sCCO. B) HEK293T cells were incubated with TAMRA dyes for 30 min at 37 ºC in media. Subsequently, cells were incubated for 10 min at 37 ºC in media containing 30 µM Tet-v2.0Me to quench free label. Cell lysates were prepared in non-reducing Laemmli buffer and analyzed by 12% SDS-PAGE. Fluorescence gel images were captured on ChemiDoc imaging system. C) Coomassie staining of gel used for fluorescent imaging to verify equal sample load.

A


GFP
TAMRA
Hoechst
DIC
Merge
Veh
TAMRA-acid
TAMRA-sTCO
TAMRA-sCCO

B


GFP
TAMRA
Hoechst
DIC
Merge
Veh
TAMRA-acid
TAMRA-sTCO
TAMRA-sCCO

C

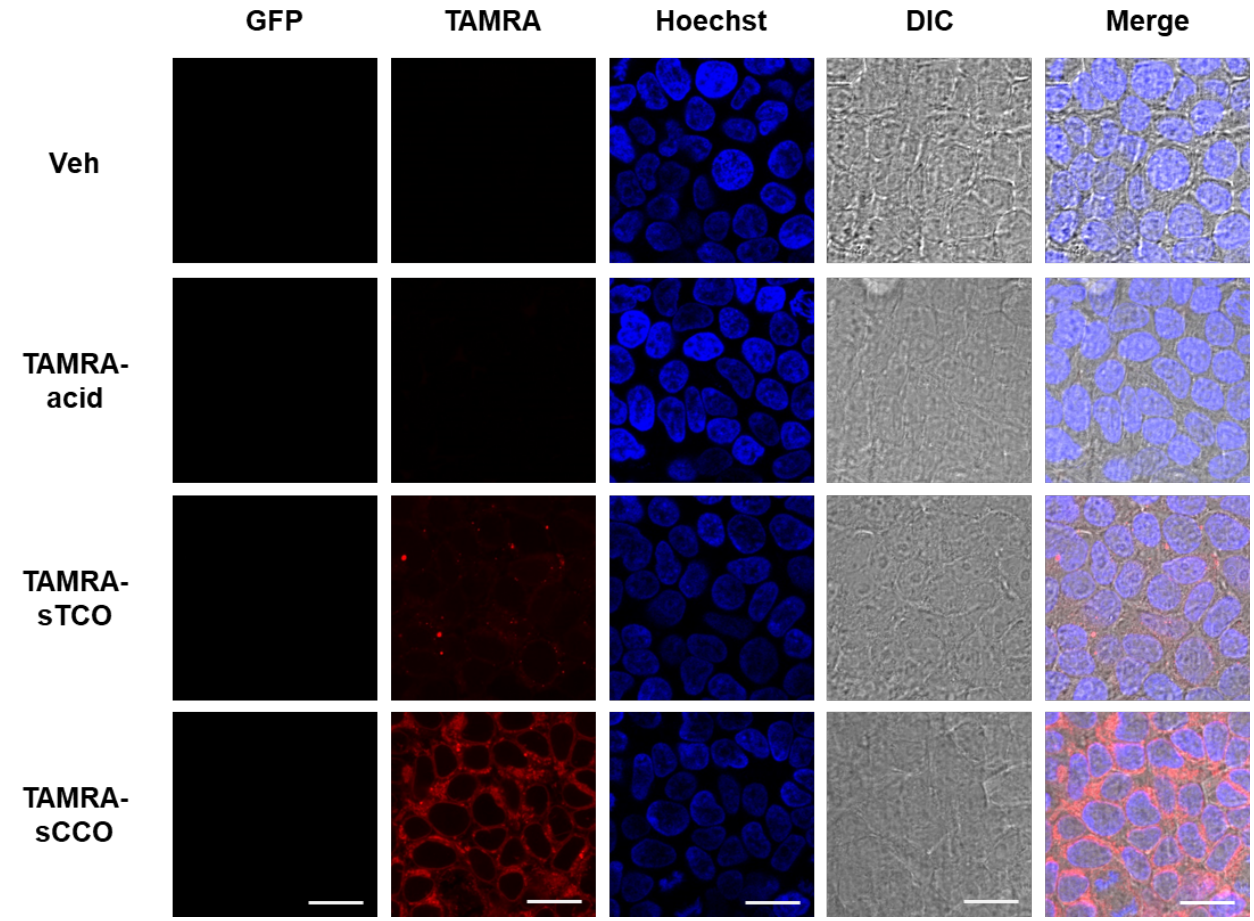


Figure S16. Bioorthogonal ligation of intracellular Tet-containing protein with fluorescent dye. HEK293T cells expressing (A) sfGFP-Tet-v3.0 Bu protein or (B) WT-sfGFP were labeled with 1 μM TAMRA-sTCO, TAMRA-sCCO, TAMRA-acid or vehicle for 30 min. (C) HEK293T cells with no DNA transfection and no Tet-v3.0 were negative control. The unreacted free label was quenched with 30 μM Tet-v2.0 Me for at least 30 min. Nuclei were stained by incubating live cells with Hoechst33342. Without washing away excess labels, fluorescence images were captured on the Zeiss LSM780 confocal microscope. TAMRA-sTCO specifically labeled GFP-Tet proteins in live cells. Scale bar, 20 μm

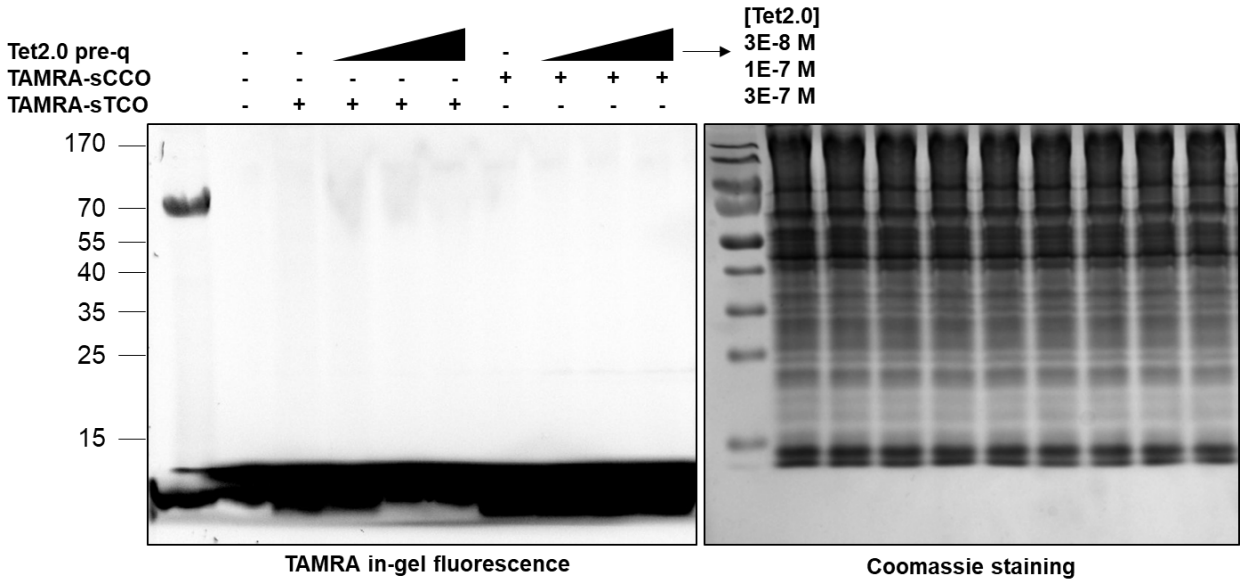


Figure S17. Assessment of non-specific protein labeling of proteins with TAMRA-sTCO and TAMRA-sCCO. HEK293T cells were incubated for 30 min with 100 nM TAMRA-sTCO, 100 nM TAMRA-sCCO, 100 nM TAMRA-sTCO-Tet-v2.0Me (pre-quenched with 30 nM, 100 nM, or 300 nM Tet-v2.0Me), 100 nM TAMRA-sCCO-Tet-v2.0Me (pre-quenched with 30 nM, 100 nM, or 300 nM Tet-v2.0Me) or vehicle as indicated. Subsequently, cells were incubated for 10 min at 37 °C in media containing 300 nM Tet-v2.0Me to quench free label. Cell lysates were prepared in non-reducing Laemmli buffer and analyzed by 12% SDS-PAGE. Fluorescence gel images were captured on ChemiDoc imaging system.

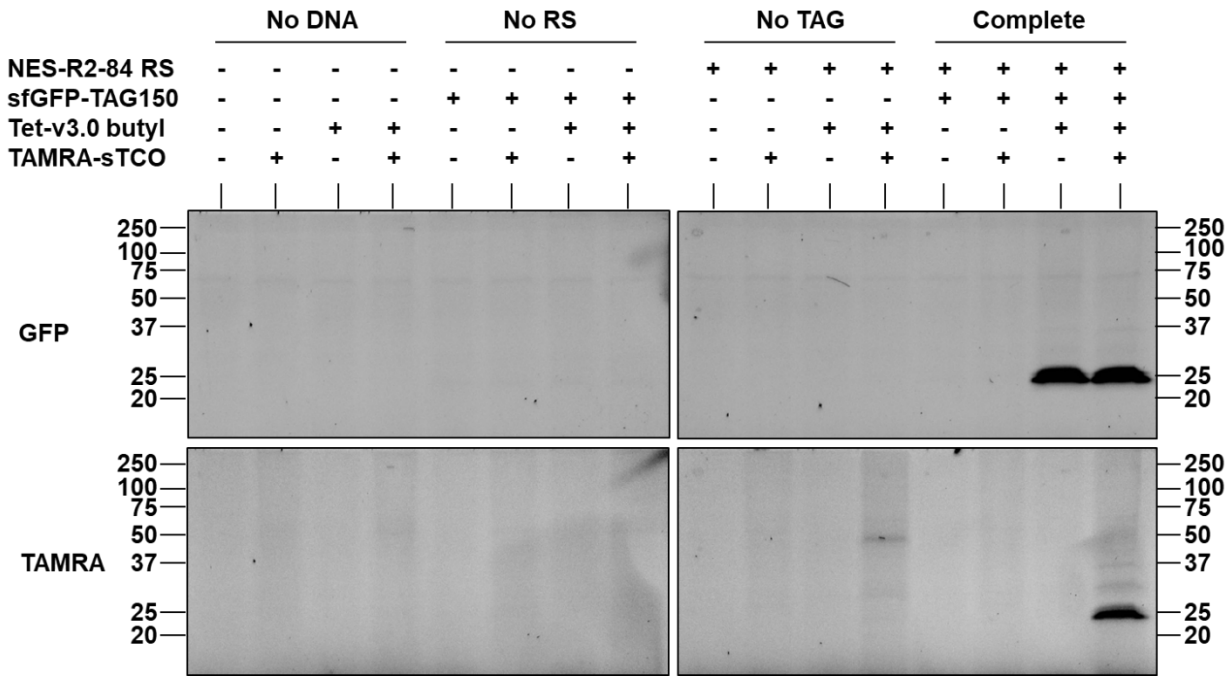


Figure S18. Assessment of non-specific labeling of proteins with TAMRA-sTCO. HEK293T cells were transfected for 24 h with pAcBac1-sfGFP-150TAG only (No RS), pAcBac1-NES-R2-84-RS only (No TAG), or pAcBac1-NES-R2-84-RS and pAcBac1-sfGFP-150TAG (complete) in the presence or absence of 30 μM Tet-v3.0Bu. Transfection without any DNA was included as a negative control (No DNA). Cells were washed three times with culture media for 2 h and incubated for 30 min with 100 nM TAMRA-sTCO or 0.1% DMSO. After quenching the unreacted label with 1 μM Tet-v2.0Me, crude cell lysates in non-reducing Laemmli buffer were analyzed by SDS-PAGE. Fluorescence gel images were captured on ChemiDoc imaging system. Excitation and emission for GFP: UV transillumination and 580/120 nm filter. Ex and Em for TAMRA: Green epi LED (520 nm-545 nm) and 605/50 nm. Minor non-specific labeling of proteome by sTCO-TAMRA was detected only if cells are exposed to tRNA/R2-84-RS and Tet-v3.0Bu.

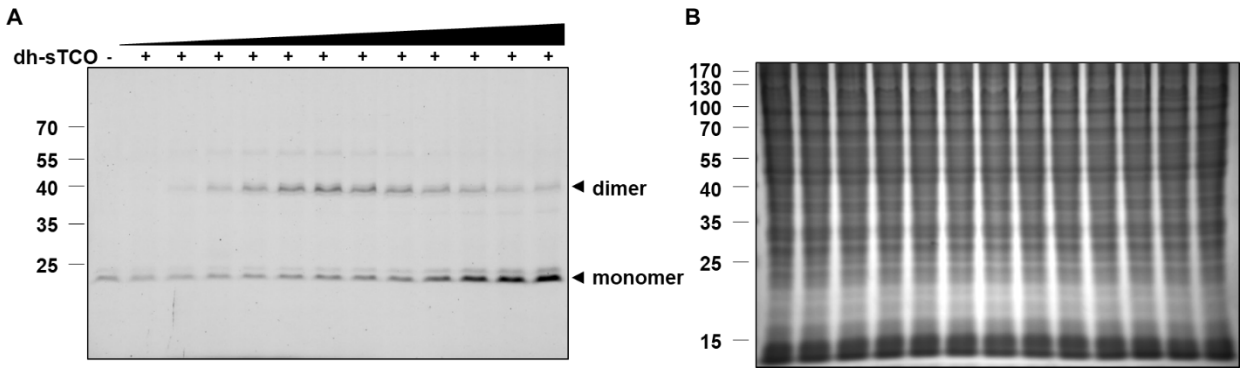


Figure S19. Concentration-dependent protein dimerization of GFP-Tet-v3.0Bu in HEK293T cells. A) HEK293T cells were transfected with pAcBac1-NES-R2-84-RS and pAcBac1-sfGFP-150TAG in the presence of 30 μM Tet-v3.0Bu for 24 h. After replacing the culture media three times, cells were incubated for 30 min with increasing concentrations (1 nM, 10 nM, 20 nM, 50 nM, 100 nM, 200 nM, 500 nM, 1 μM, 10 μM, 50 μM, 100 μM, and 200 μM) of dh-sTCO or 0.1% DMSO followed by 10 min quenching with Tet-v2.0Me. The whole cell lysate in non-reducing Laemmli buffer was separated by SDS-PAGE, and GFP fluorescence images were captured to estimate the dimerization. Densitometry analysis was performed to quantify the monomer and the dimer band intensity using ImageJ software (Figure 5B). B) The gel was stained with Coomassie Brilliant Blue R to visualize the total loading proteins.

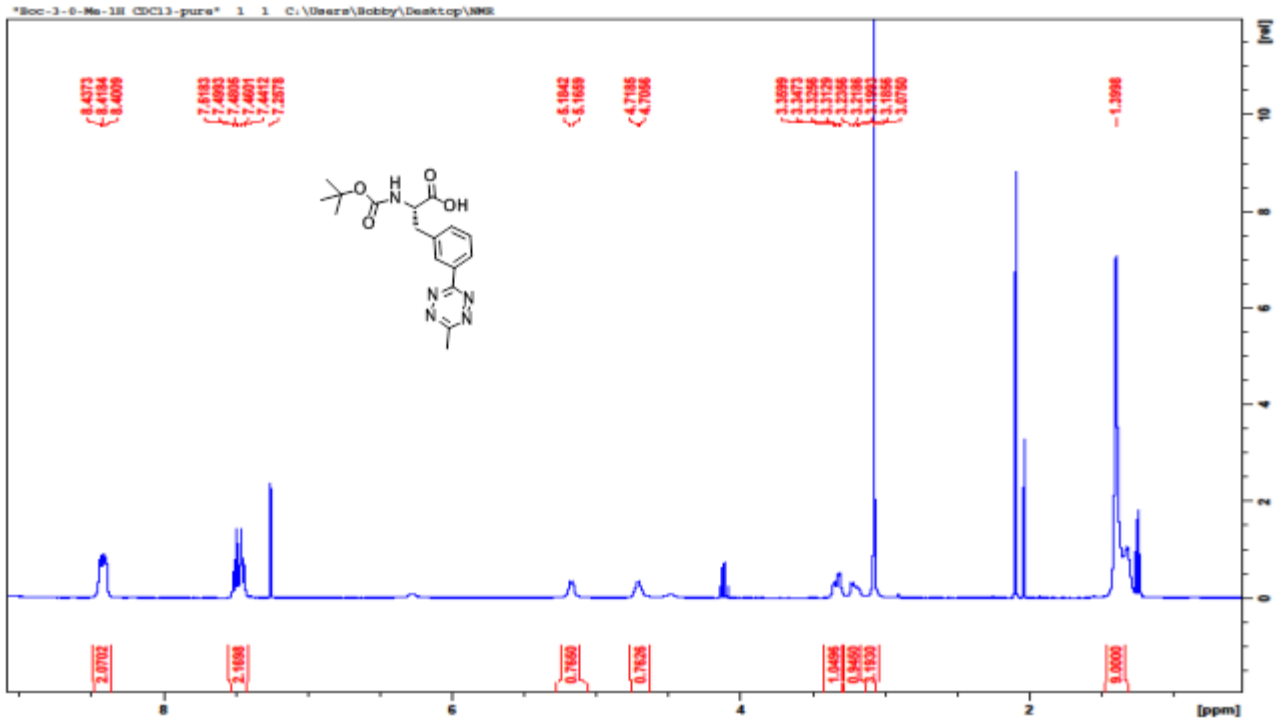


Figure S20. [1]H NMR spectra of Boc-Tet3.0Me. **2a**

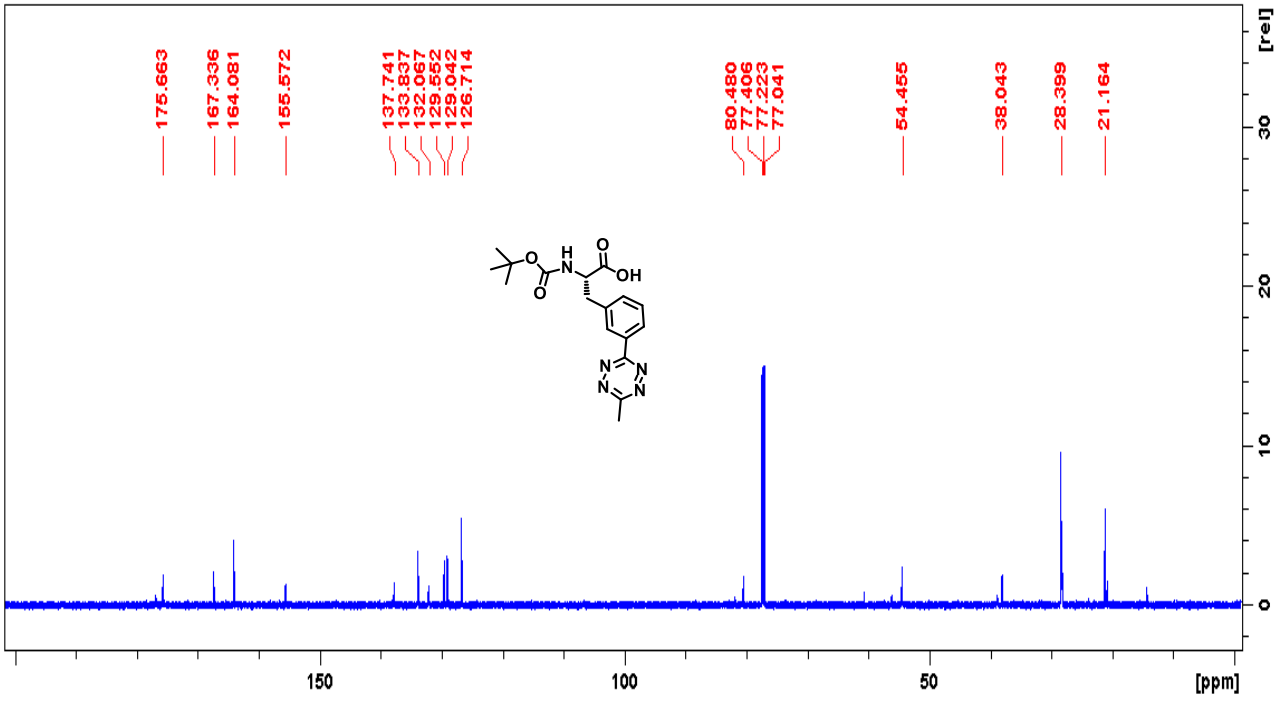


Figure S21. [13]C NMR spectra of Boc-Tet3.0Me. **2a**

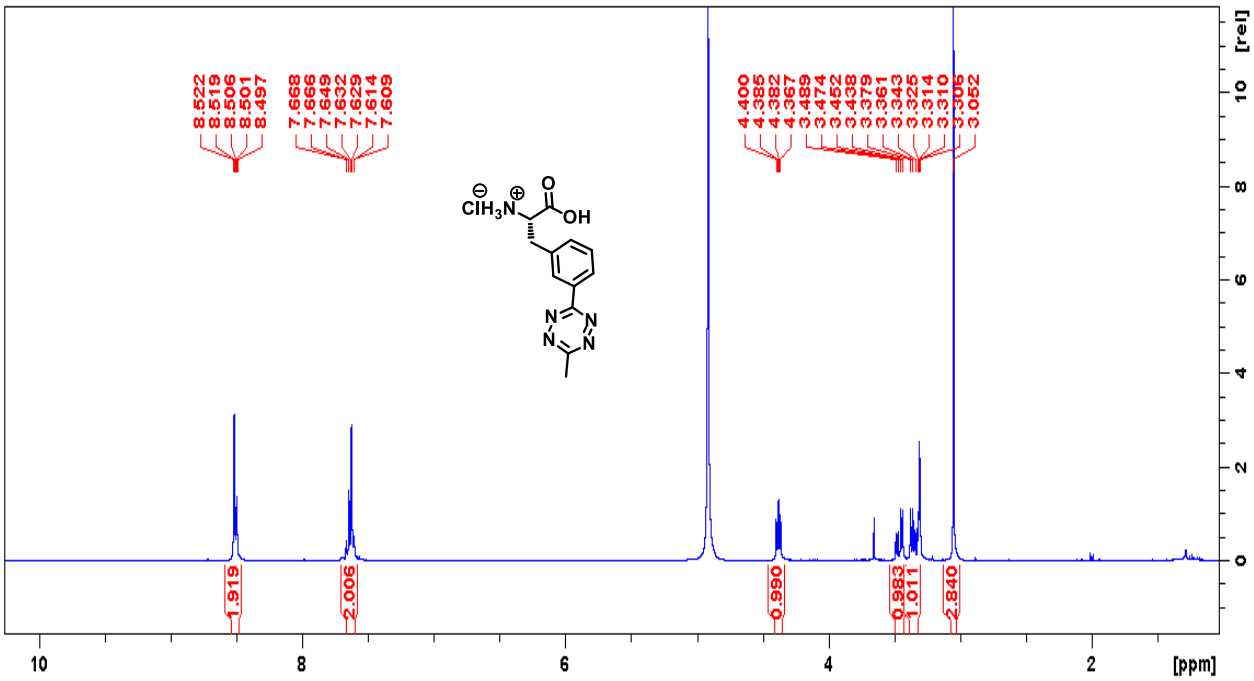


Figure S22. ^{1}H NMR spectra of Tet3.0Me. **1a**

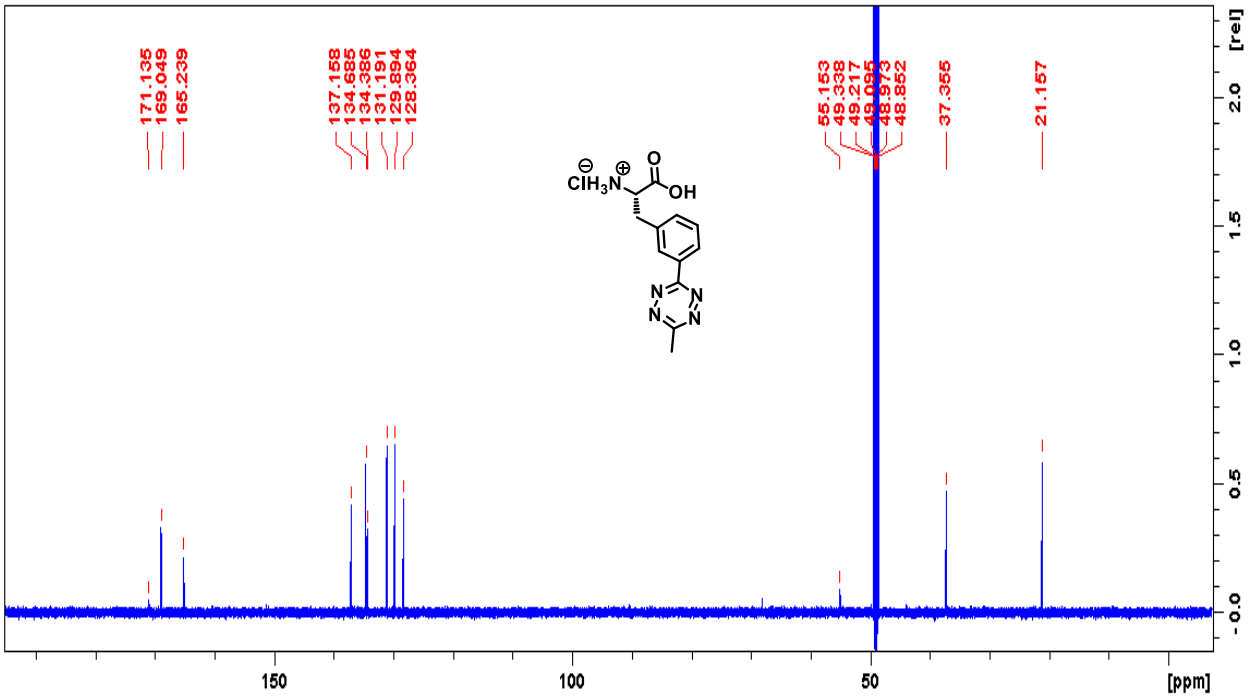


Figure S23. ^{13}C NMR spectra of Tet3.0Me. **1a**

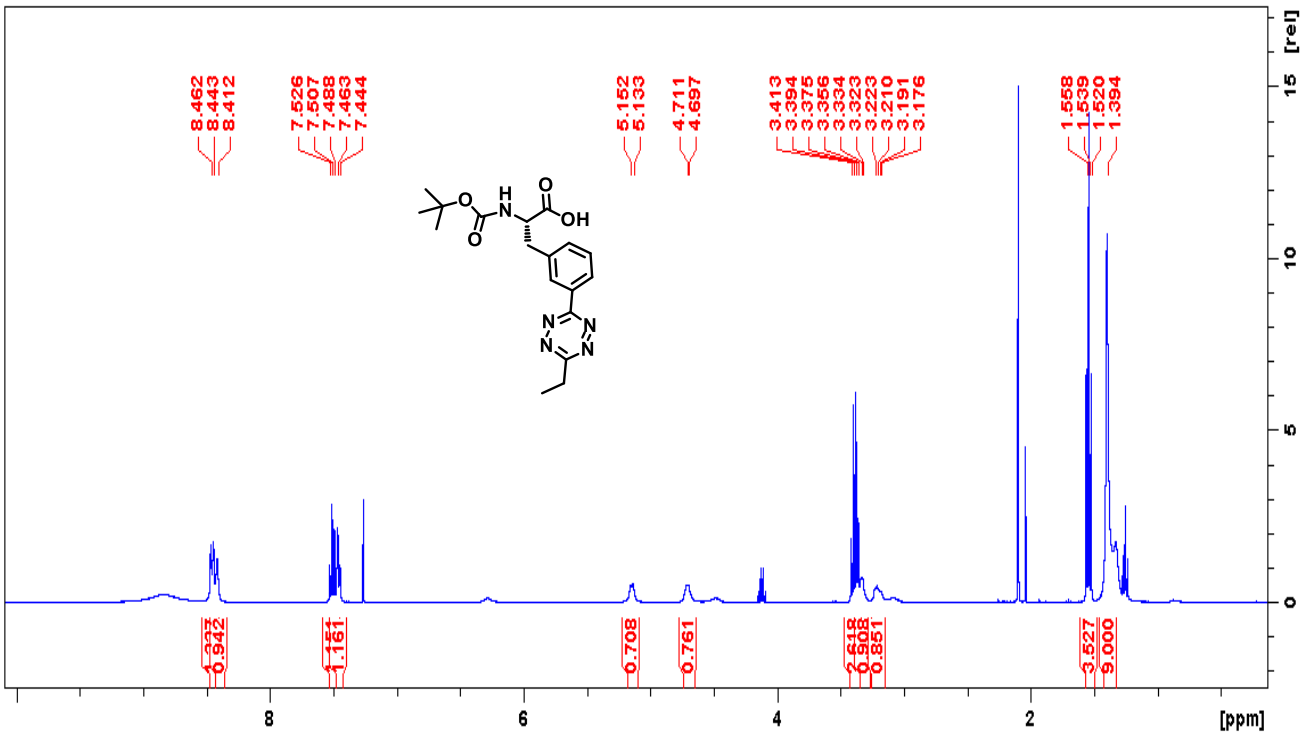


Figure S24. ^{1}H NMR spectra of Boc-Tet3.0Et. **2b.**

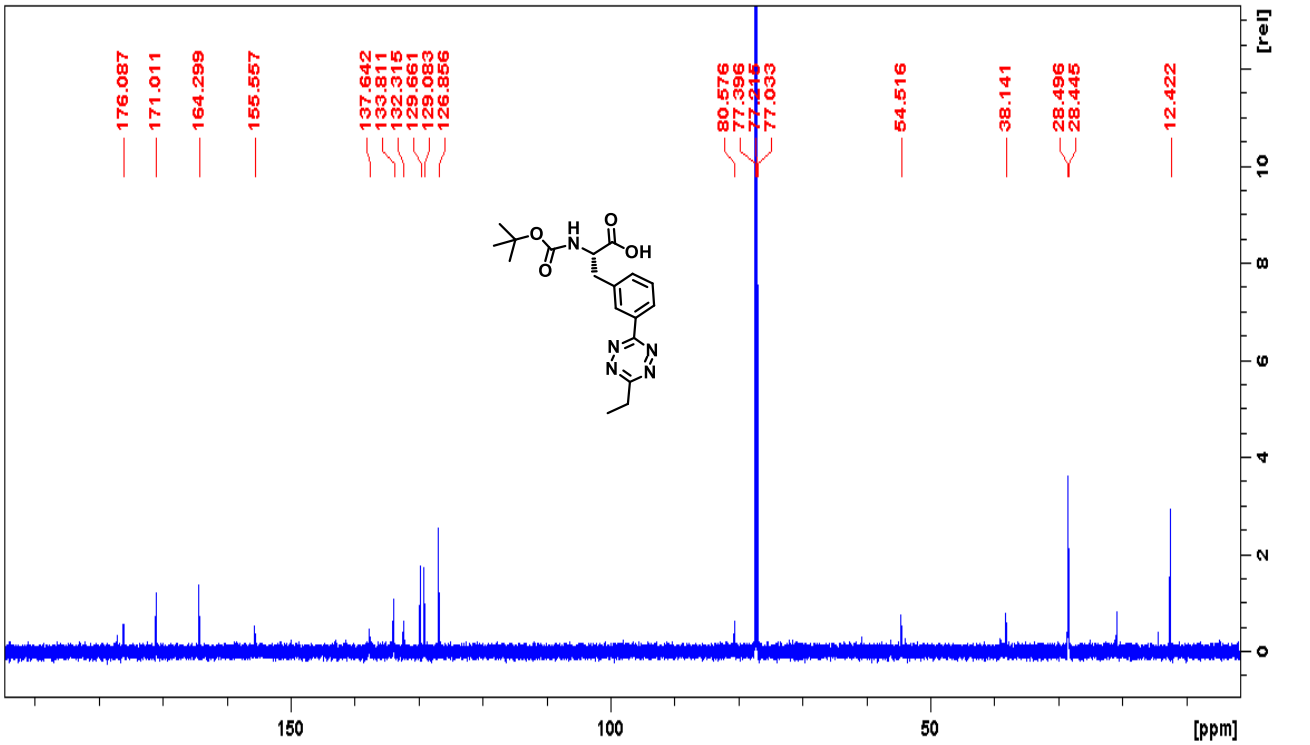


Figure S25. ^{13}C NMR spectra of Boc-Tet3.0Et. **2b**

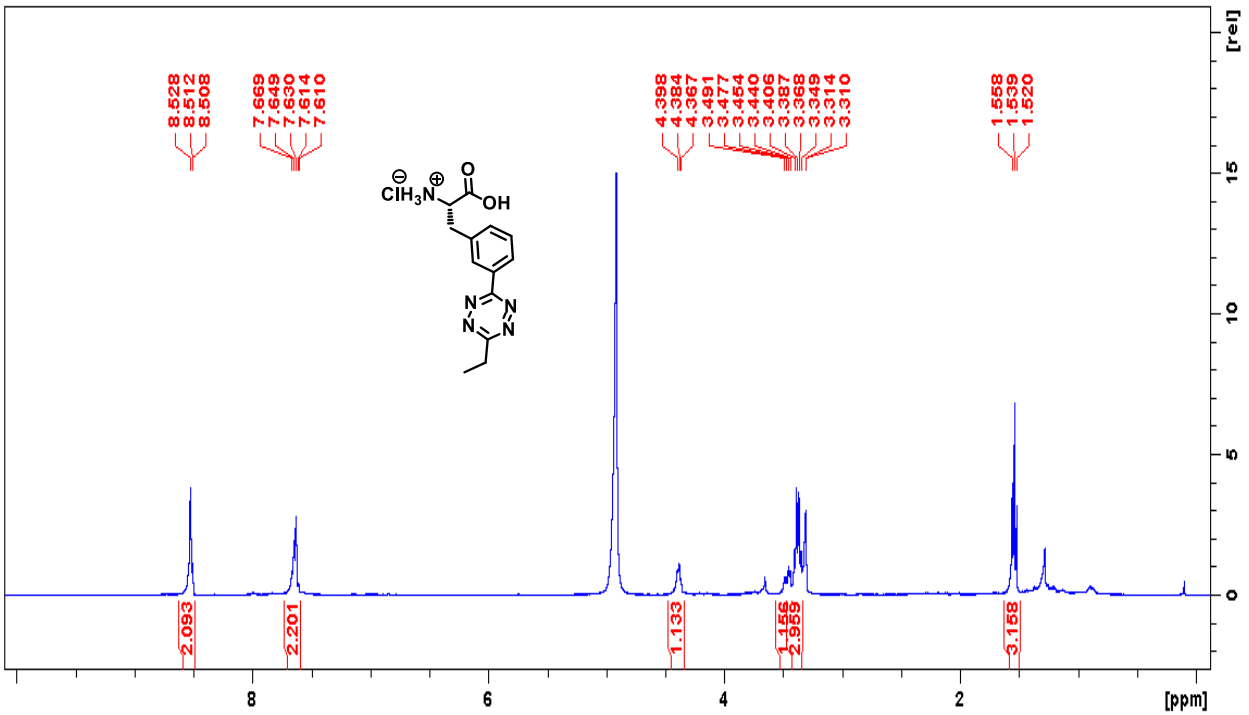


Figure S26. ^{1}H NMR spectra of Tet3.0Et. **1b**

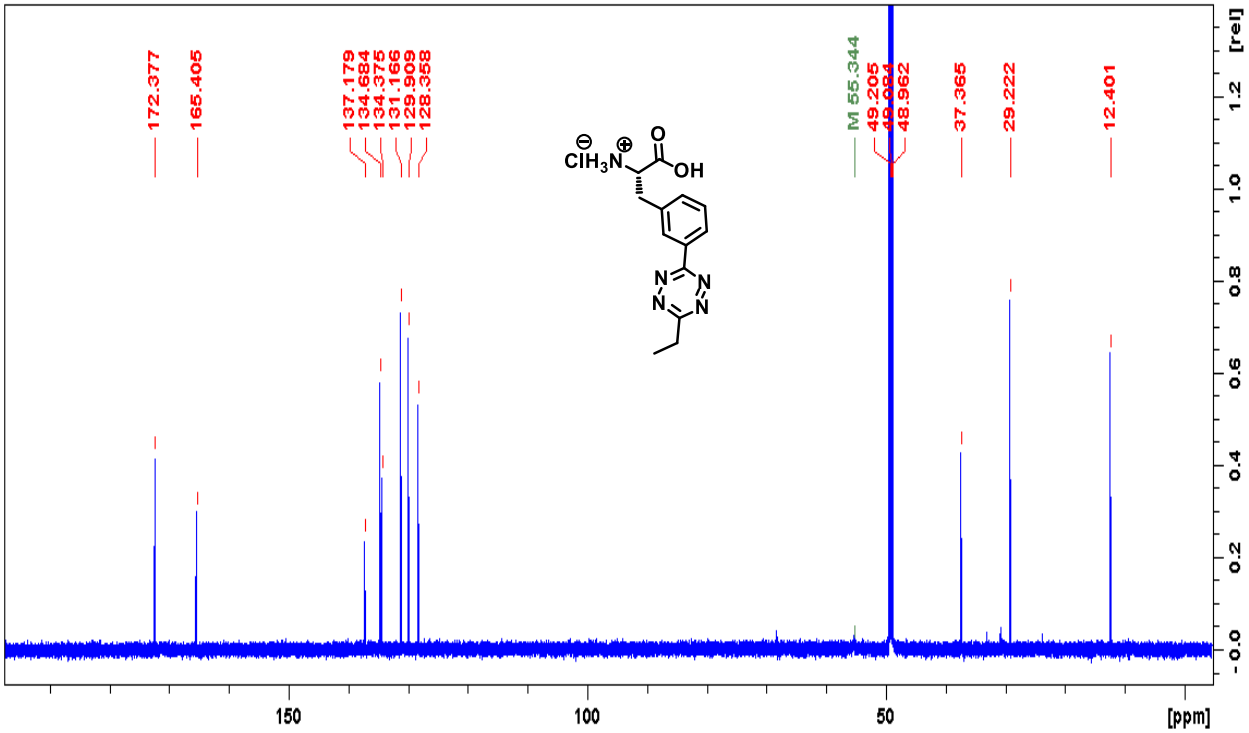


Figure S27. ^{13}C NMR spectra of Tet3.0Et. **1b.**

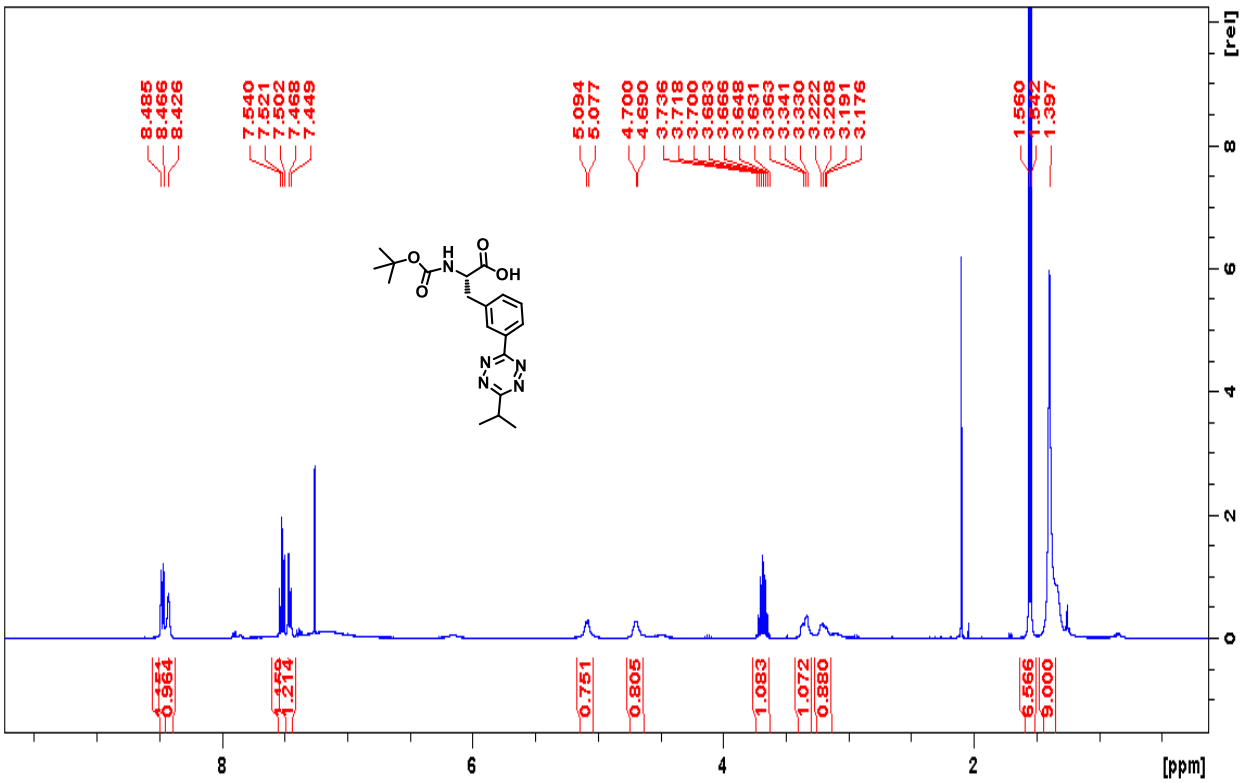


Figure S28. ^{1}H NMR spectra of Boc-Tet3.0Ip. **2c**

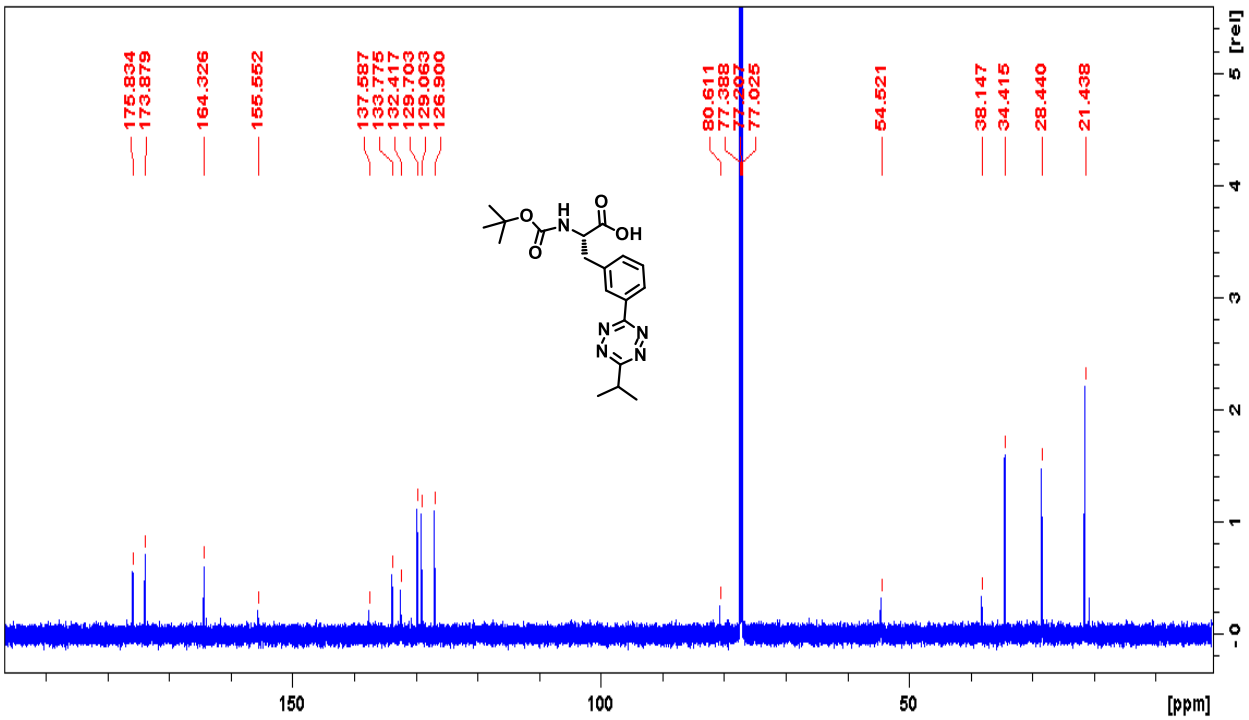


Figure S29. ^{13}C NMR spectra of Boc-Tet3.0Ip. **2c**

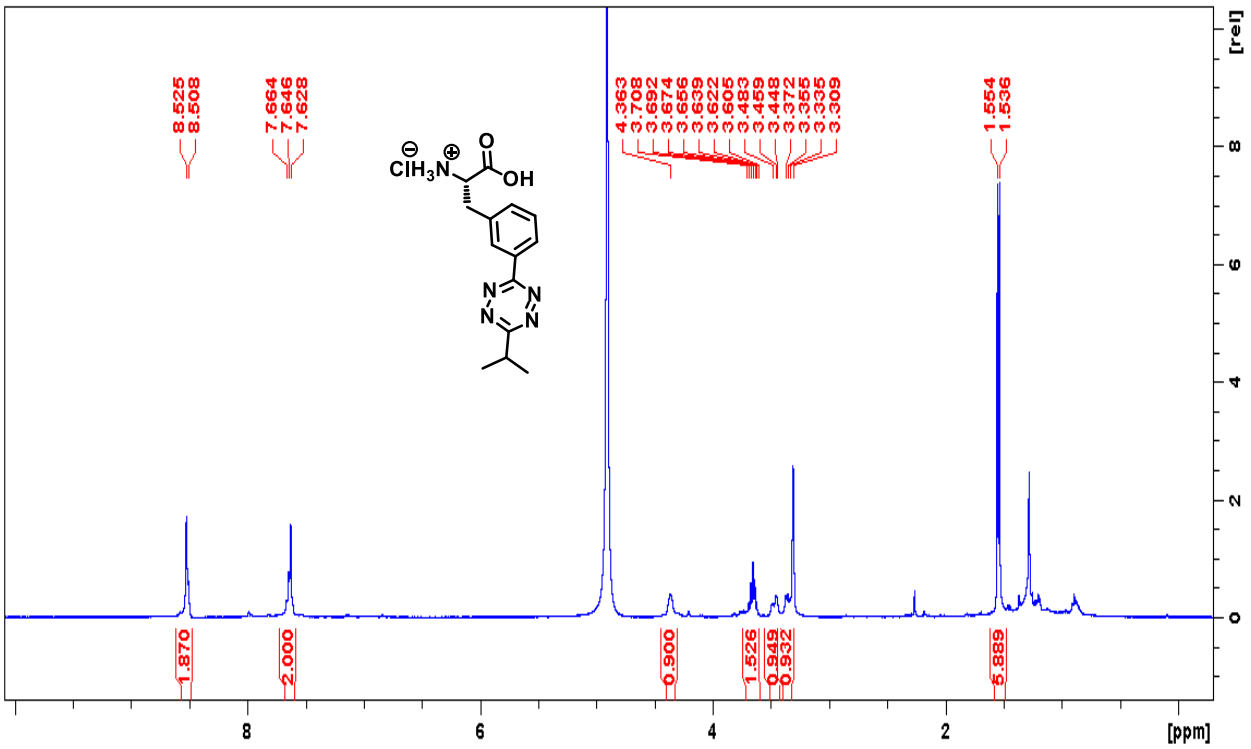


Figure S30. ^{1}H NMR spectra of Tet3.0Ip. **1c**

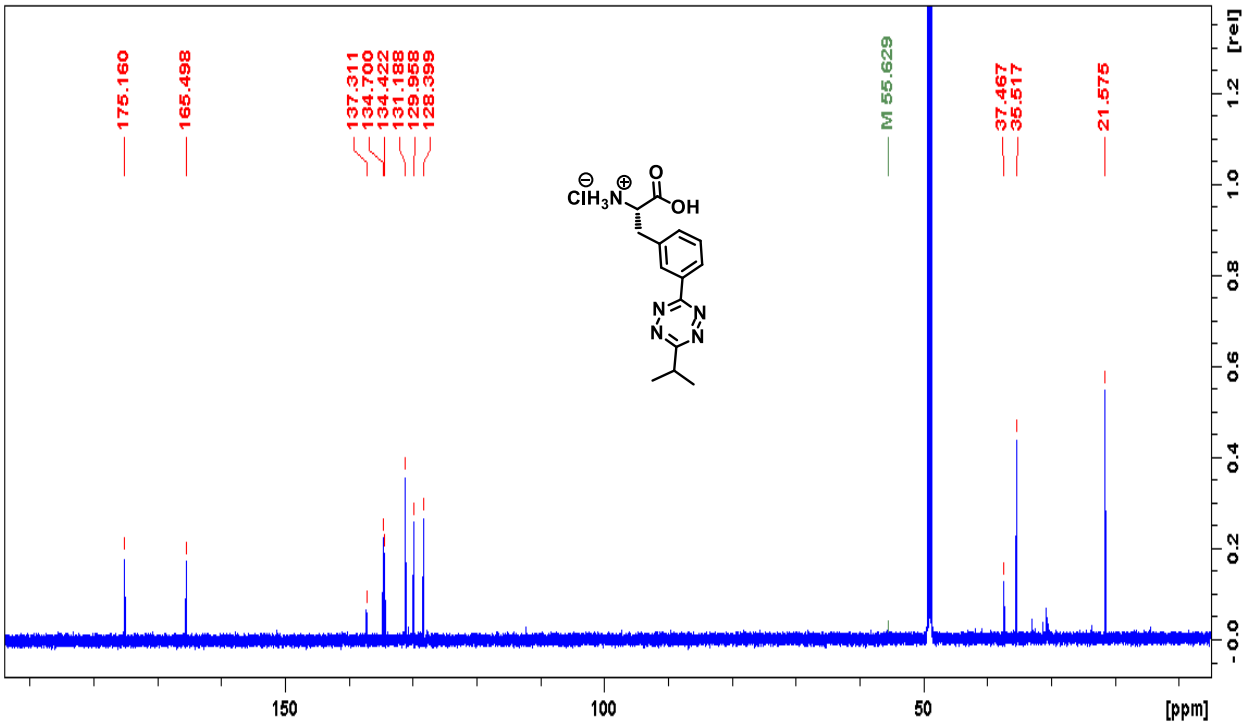


Figure S31. ^{13}C NMR spectra of Tet3.0Ip. **1c**

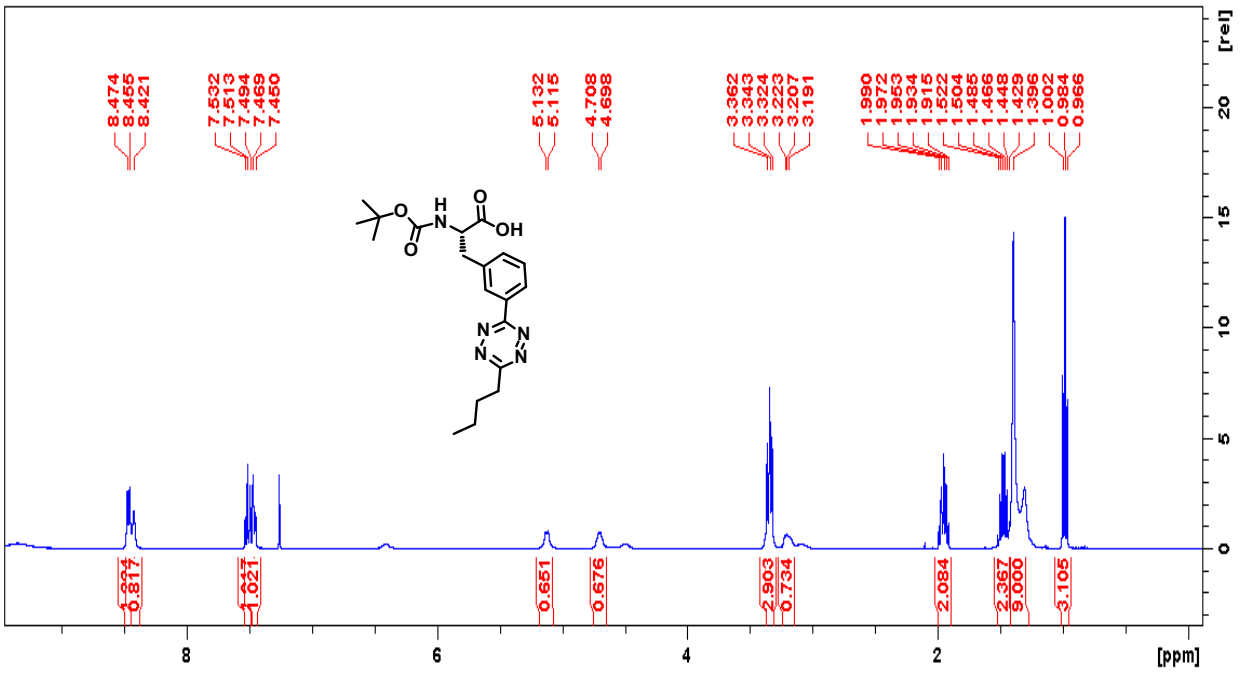


Figure S32. ^{1}H NMR spectra of Boc-Tet3.0Bu. **2d**

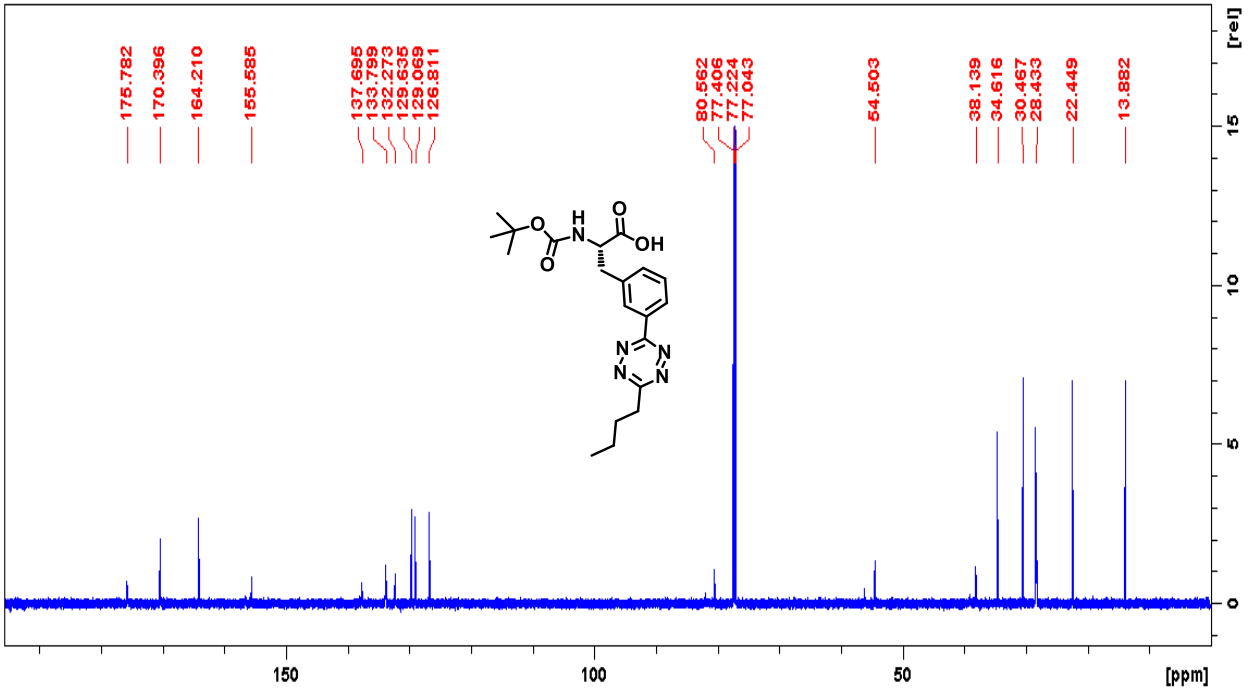


Figure S33. ^{13}C NMR spectra of Boc-Tet3.0Bu. **2d**

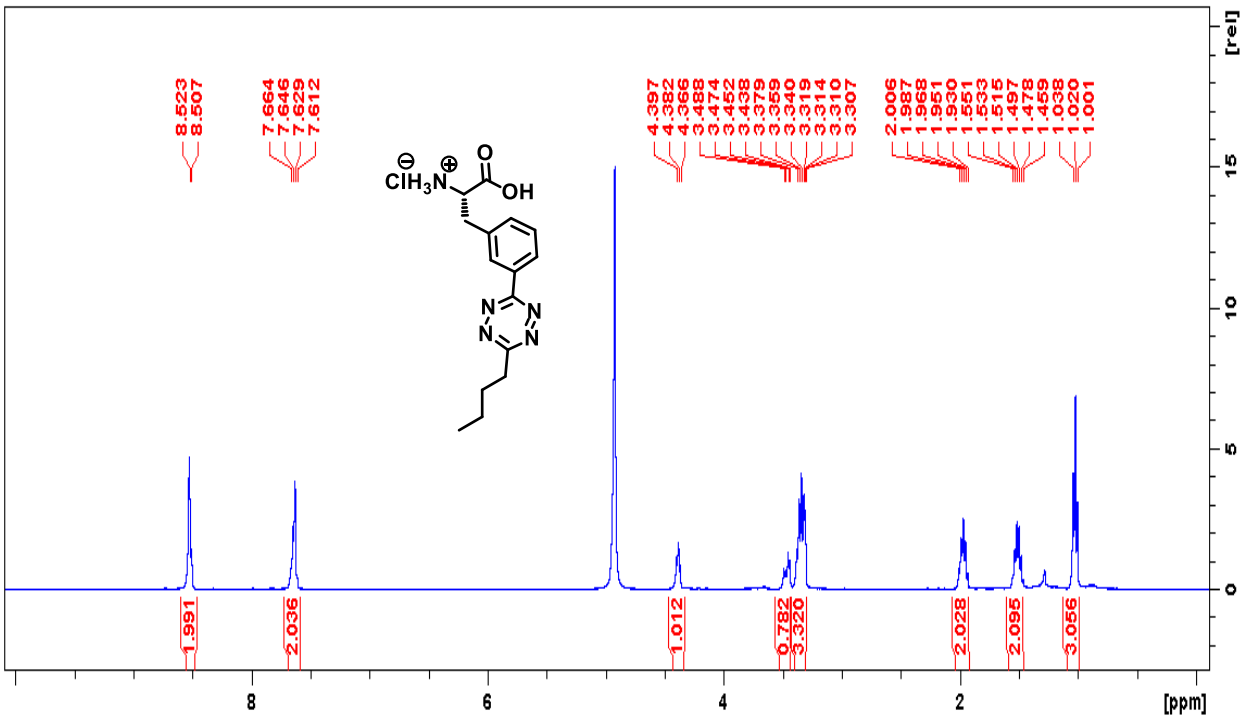


Figure S34. ^{1}H NMR spectra of Tet3.0Bu. **1d**

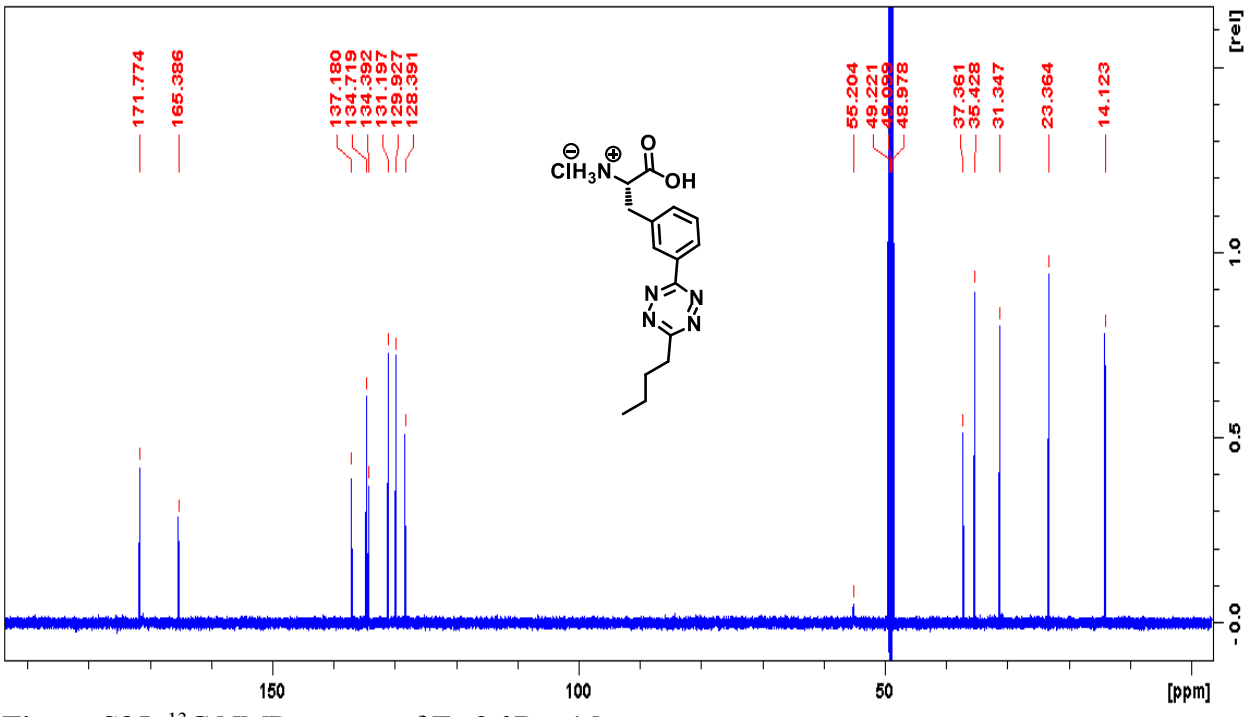


Figure S35. ^{13}C NMR spectra of Tet3.0Bu. **1d**

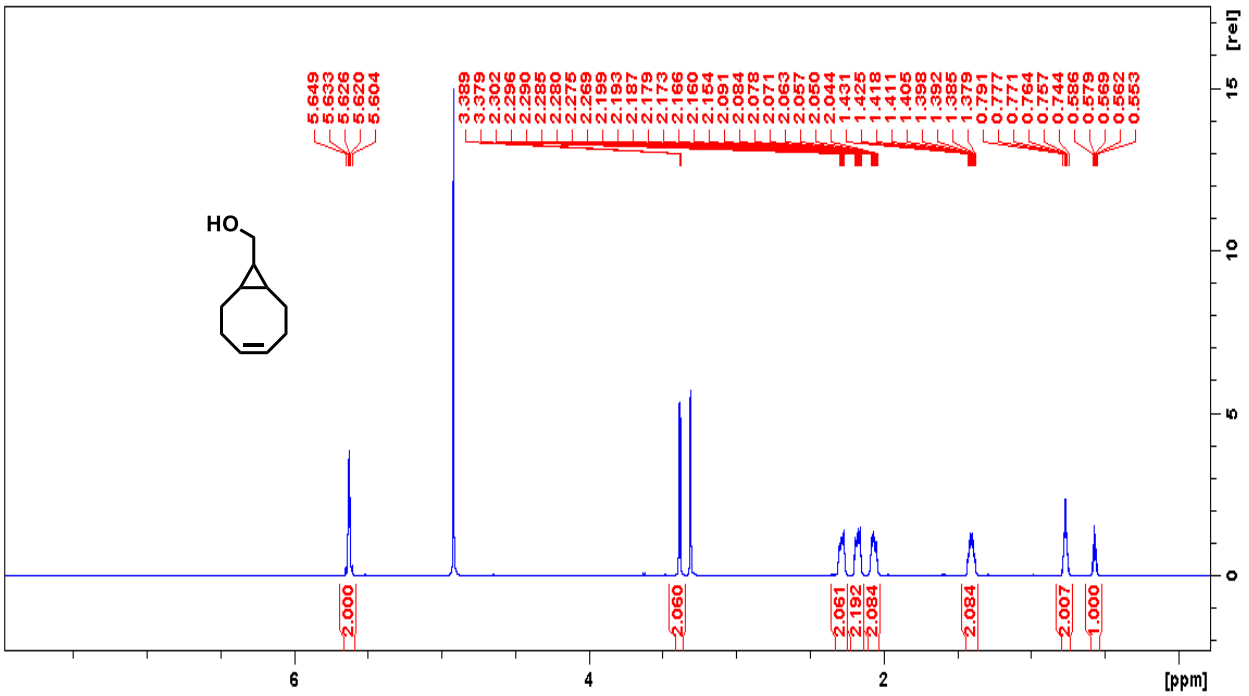


Figure S36. ^{1}H NMR spectra of sCCO **7.**

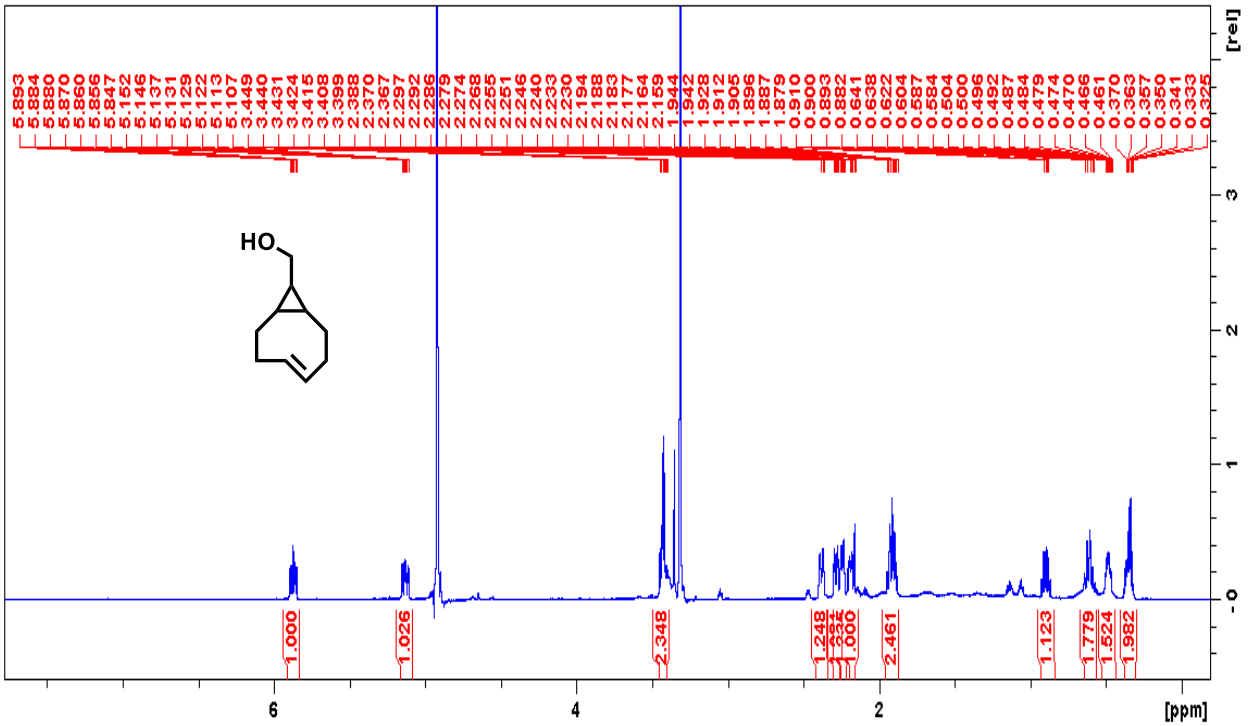


Figure S37. ^{1}H NMR spectra of sTCO **8**.

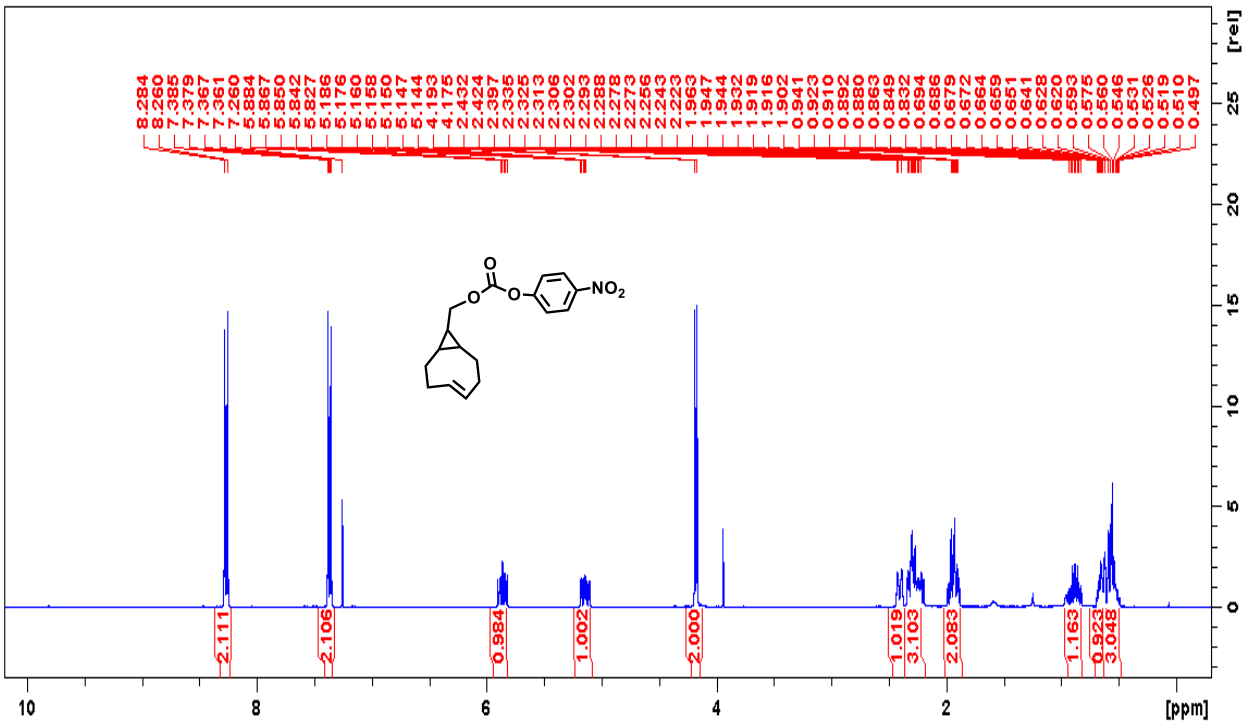


Figure S38. ^{1}H NMR spectra of activated ester of sTCO **9**.

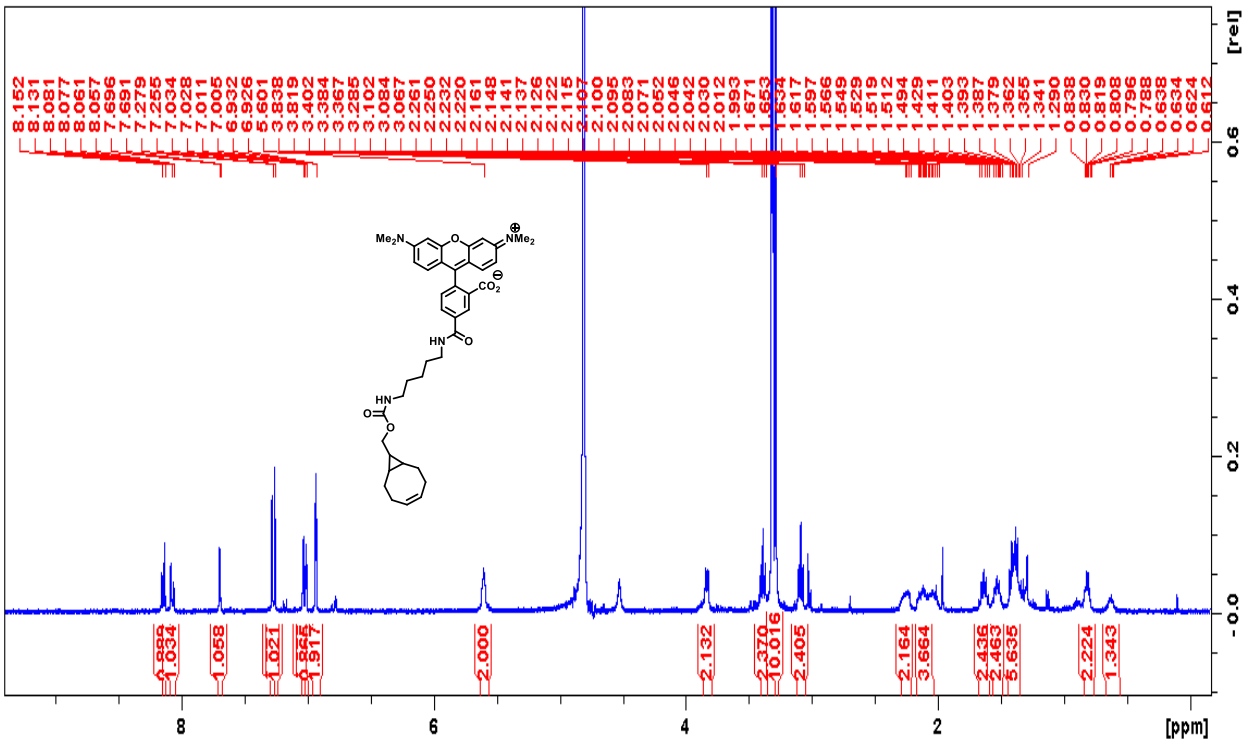


Figure S39. ^{1}H NMR spectra of sCCO linked TAMRA **5.**

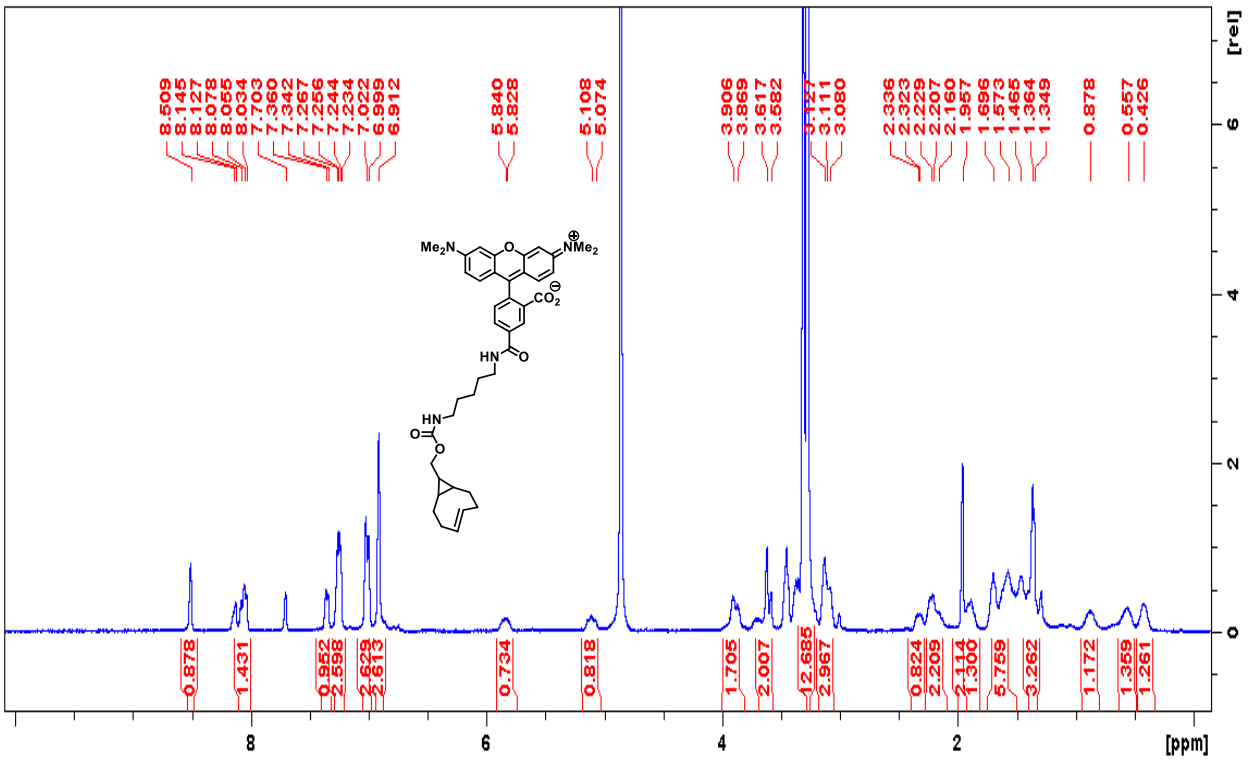


Figure S40. ^{1}H NMR spectra of sTCO linked TAMRA **6.**

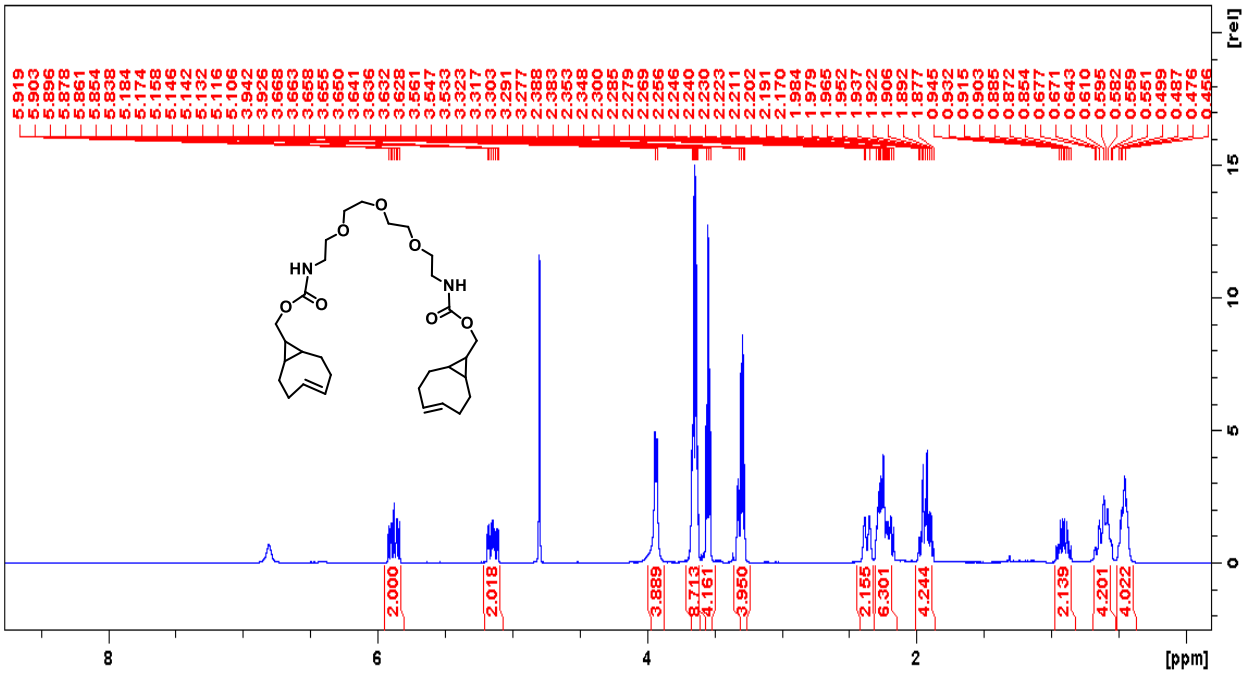


Figure S41. ^{1}H NMR spectra of PEG192 linked dh-sTCO **11.**

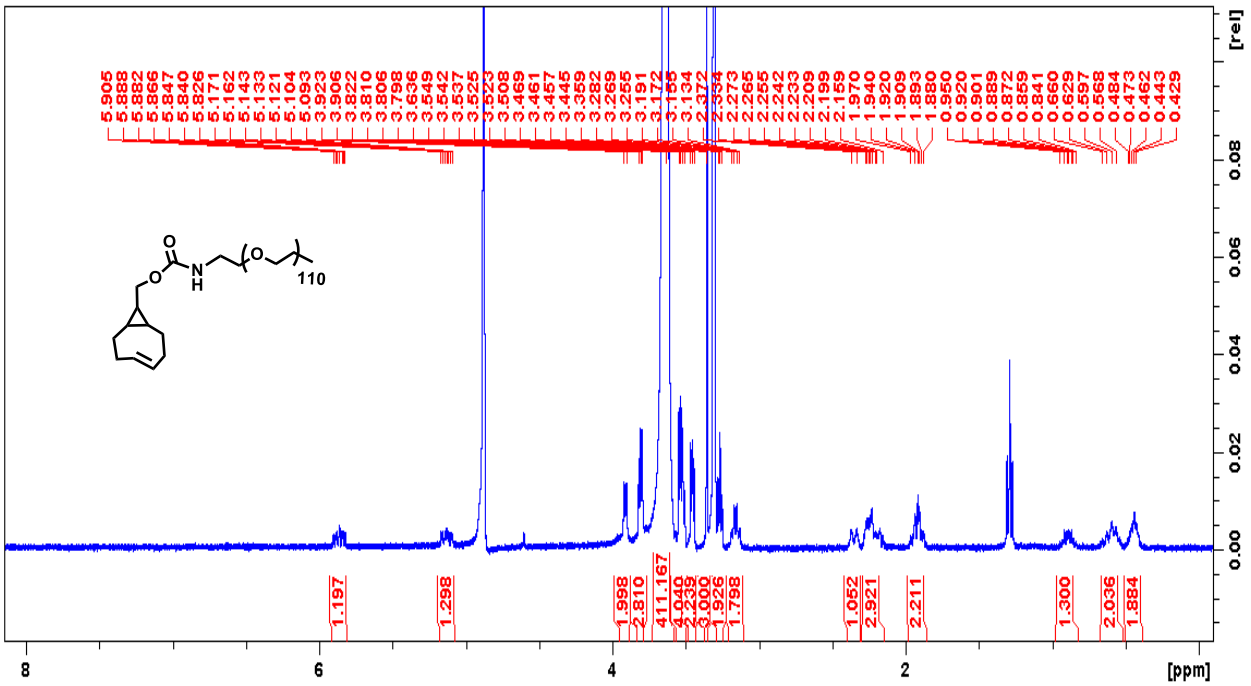


Figure S42. ^{1}H NMR spectra of mPEG5000 linked sTCO **12**.

Supplemental References:

1. O'Brien, J. G. K.; Chintala, S. R.; Fox, J. M., Stereoselective Synthesis of Bicyclo[6.1.0]nonene Precursors of the Bioorthogonal Reagents s-TCO and BCN. *J Org Chem* **2018,** *83* (14), 7500-7503.
2. Royzen, M.; Yap, G. P.; Fox, J. M., A photochemical synthesis of functionalized trans-cyclooctenes driven by metal complexation. *J Am Chem Soc* **2008,** *130* (12), 3760-1.
3. Porter, J. J.; Jang, H. S.; Van Fossen, E. M.; Nguyen, D. P.; Willi, T. S.; Cooley, R. B.; Mehl, R. A., Genetically Encoded Protein Tyrosine Nitration in Mammalian Cells. *ACS Chem Biol* **2019,** *14* (6), 1328-1336.
4. Hammill, J. T.; Miyake-Stoner, S.; Hazen, J. L.; Jackson, J. C.; Mehl, R. A., Preparation of site-specifically labeled fluorinated proteins for 19F-NMR structural characterization. *Nat Protoc* **2007,** *2* (10), 2601-7.
5. Chatterjee, A.; Xiao, H.; Bollong, M.; Ai, H. W.; Schultz, P. G., Efficient viral delivery system for unnatural amino acid mutagenesis in mammalian cells. *Proc Natl Acad Sci U S A* **2013,** *110* (29), 11803-8.