Supporting Information

A Fluorogenic TMP-tag for High Signal-to-Background Intracellular Live Cell Imaging

Chaoran Jing and Virginia W. Cornish*

Department of Chemistry, Columbia University
550 West 120th Street, MC 4854, NWC Building, New York, NY 10027

E-mail: vc114@columbia.edu

Table of Contents

Molecular design

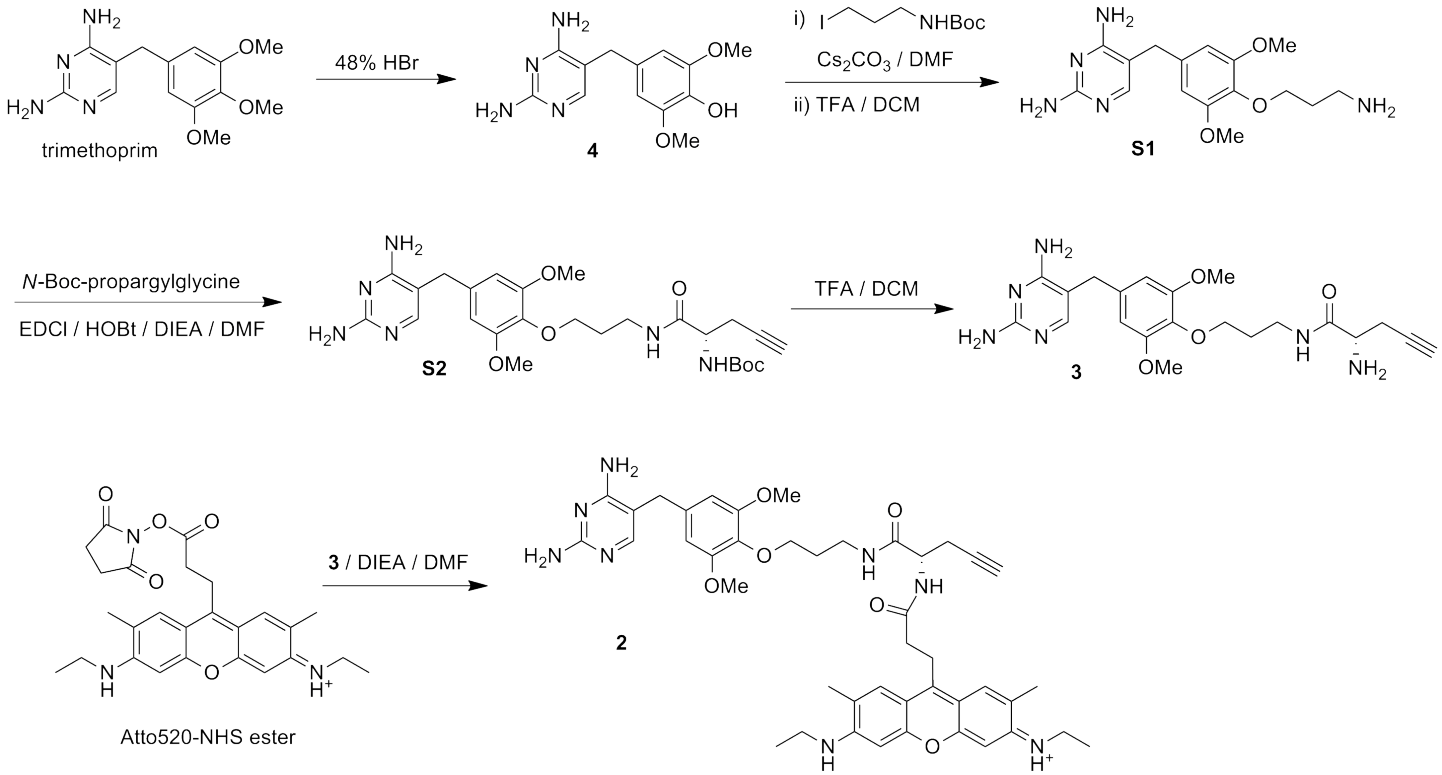
Using the model described in Figure 3, an approximation was made of the minimum linker length between TMP and the electrophile that would allow proximity-induced S_N2 reaction to occur upon binding to a mutant eDHFR containing each of these Cys mutants (Table S1). The minimum length was approximated as the arc, $L = \pi R\vartheta/180°$, where R is the radius of gyration and ϑ is the angle between the 4'-OH and the β-atom of each selected residue, using the center of mass of eDHFR as the vertex.

Table S1: Estimation of the Minimum Linker Length for Proximity-Induced Covalent Labeling

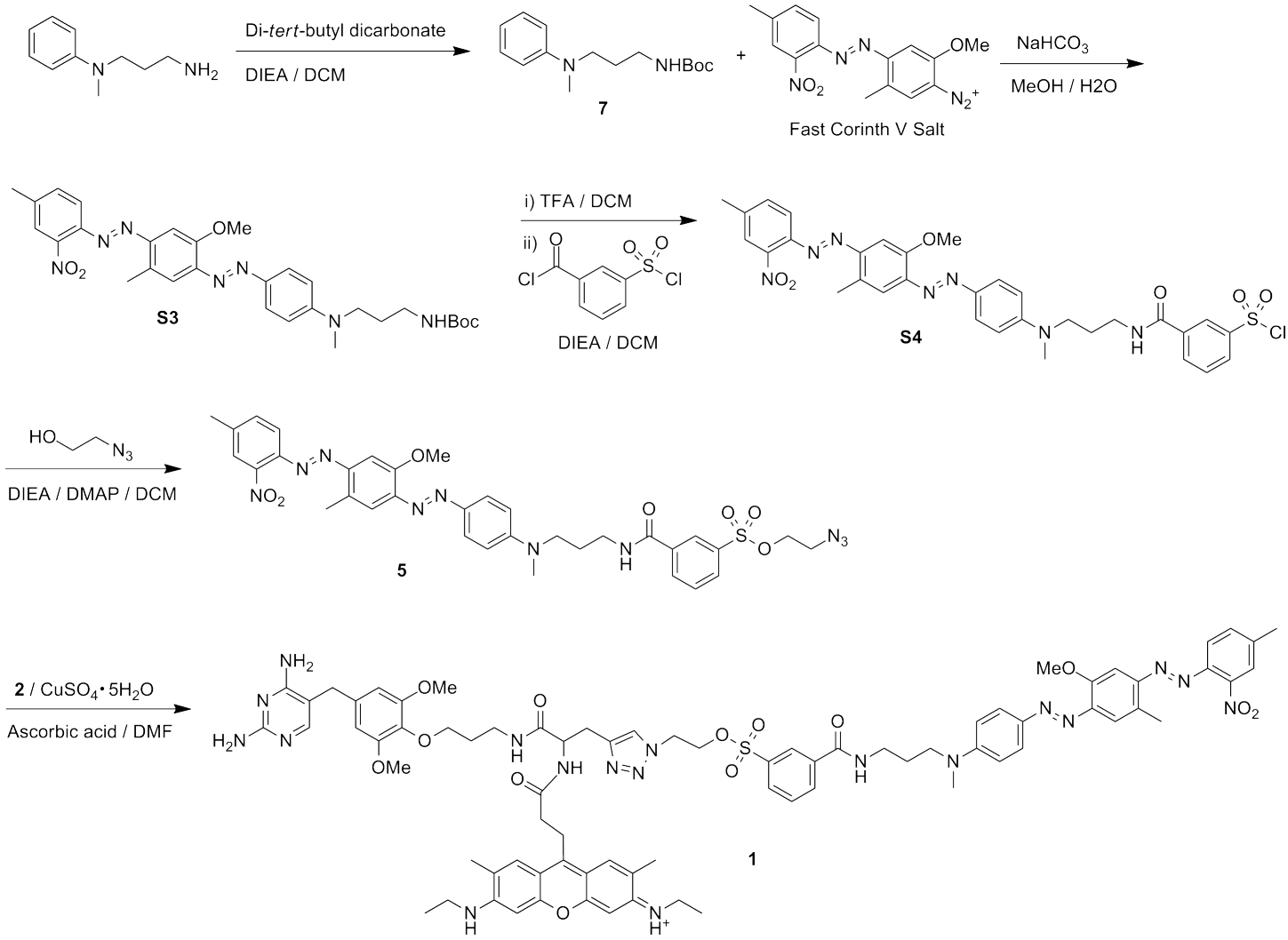
Residue	Angle ϑ	Radius R (Å)	Length L (Å)
Glu17	50.8°	15.2	13.47
Asn18	47.8°	15.2	12.68
Ala19	38.7°	15.2	10.27
Met20	33.7°	15.2	8.94
Pro21	46.7°	17.2	14.01
Asn23	30.8°	15.2	8.18
Pro25	28.6°	15.2	7.59
Leu28	35.6°	15.2	9.45
Ala29	55.9°	15.2	14.82
Lys32	54.8°	15.2	14.53
Ser49	33.0°	15.2	8.75
Ile50	29.3°	15.2	7.77
Gly51	29.0°	17.3	8.75
Arg52	35.9°	15.2	9.51
Leu54	45.6°	15.2	12.10
Pro55	48.7°	18.0	15.30

Synthetic Chemistry.

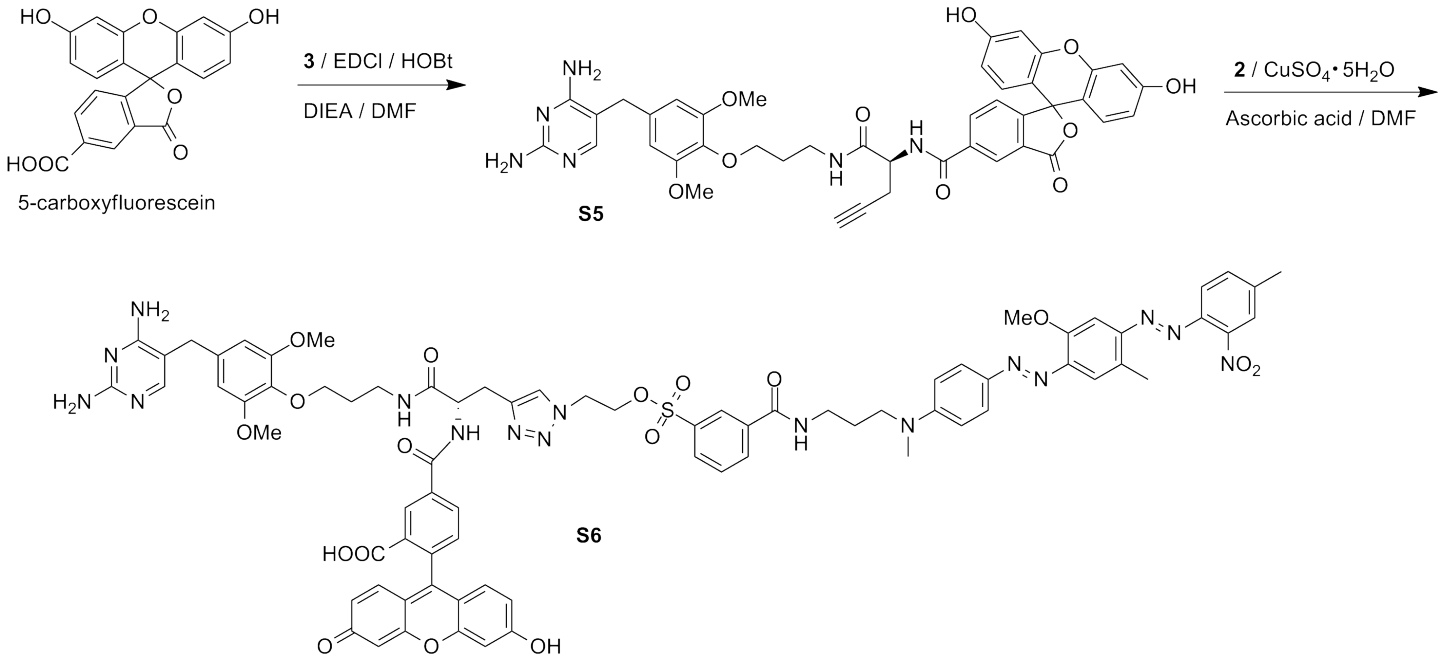
General Methods. Unless otherwise noted, reagents and solvents were obtained from Aldrich and were used without further purification. Anhydrous *N*,*N*-dimethylformamide (DMF) was from Sure Seal™ bottles purchased from Aldrich. *N*-(3-aminopropyl)-*N*-methylaniline and 3-(chlorosulfonyl)benzoyl chloride are purchased form Tokyo Chemical Industry (TCI) Co., Ltd. *N*-Boc-propargylglycine is purchased from Chem-Impex International, Inc. Nuclear magnetic resonance (NMR) spectra were recorded on a Bruker 400 (400 MHz) or Bruker 300 (300 MHz) Fourier Transform (FT) NMR spectrometers at Columbia University, Chemistry Department. ^{1}H NMR spectra are tabulated in the following order: multiplicity (s, singlet; d, doublet; t, triplet; q, quartet; p, pentet; m, multiplet; br, broad), number of protons. Fast Atom Bombardment (FAB) high resolution mass spectra (HRMS) were recorded on a JMS-HX110A mass spectrometer, and low resolution electron spray ionization (ESI) MS were recorded on a JMS-LC mate mass spectrometer.



Scheme S1: Synthesis of compound **2**.



Scheme S2: Synthesis of compound **1**.



Scheme S3: Synthesis of compound **S6**.

Synthesis of compound 4 TMP-OH. A round bottom flask containing 48% aqueous hydrobromic acid (250 mL) was heated to 100 ˚C. Trimethoprim (25.0 g, 86 mmol) was suspended and stirred for 20 minutes. The flask was removed from the oil bath and 50% (v/w) sodium hydroxide in water (75 mL) was added. The reaction mixture was then cooled to room temperature and the resulting crystals were filtered and rinsed with cold water. The crystals were then dissolved in hot water (350 mL), neutralized with aqueous NH_4OH and allowed to crystallize. The crystals were filtered, rinsed with cold water, and dried under vacuum to produce 15.9 g (67% yield). ^{1}H NMR (400 MHz, CD_3OD) δ ppm: 7.18 (s, 1H), 6.53 (s, 2H), 3.81 (s, 6H), 3.62 (s, 2H). ^{13}C NMR (400 MHz, CD_3OD) δ ppm: 163.45, 156.58, 148.88, 135.25, 130.78, 107.16, 106.90, 56.50, 33.98. HRMS (FAB+) m/z Calcd. for $C_{13}H_{16}O_3N_4$ $[M+H]^+$:277.1222. Found:277.1312.

Synthesis of compound S1 TMP-C3-NH_2. Compound **4** (600 mg, 2.16 mmol), *tert*-butyl *N*-(3-iodopropyl)carbamate (618 mg, 2.16 mmol) and cesium carbonate (1.41 g, 4.32 mmol) were dissolved in anhydrous DMF (30 mL). The reaction was allowed to stand at 70 °C for 7 h, followed by the removal of solvent *in vacuo*. The crude mixture was purified by column chromatography with silica gel (1:9 methanol:DCM) to yield TMP-C3-NHBoc (Rf = 0.3 in 1:9 methanol:DCM), which was subsequently deprotected by directly dissolving in 10 mL trifluoroacetic acid (TFA) while stirring. After 4 h, the TFA was removed *in vacuo* to yield compound **S2** in the form of its TFA salt (715 mg, 1.66 mmol, 77% yield in two steps) as colorless solid. Rf = 0.02 in 1:9 methonal:DCM. ^{1}H NMR (400 MHz, CD_3OD) δ ppm: 7.28 (s, 1 H); 6.64 (s, 2 H); 4.12 (t, J = 5.4 Hz, 2 H); 3.87 (s, 6 H); 3.70 (s, 2 H); 3.30 (t, J = 6.0 Hz, 2 H); 2.08 (m, 2 H). ^{13}C NMR (400 MHz, CD_3OD) δ ppm: 164.45, 163.25, 155.96, 154.71, 136.84, 136.04, 108.03, 106.63, 74.31, 56.52, 41.95, 34.50, 30.95. HRMS (FAB+) m/z Calcd. for $C_{16}H_{24}O_3N_5$ $[M+H]^+$: 334.1879. Found: 334.1883.

Synthesis of compound 3 TMP-alkyne-NH_2. *N*,*N*-Diisopropylethylamine (DIEA, 871 µL, 4 mmol) was added dropwise to a solution of compound **S1** (192 mg, 0.4 mmol, in the form of TFA salt), *N*-Boc-propargylglycine (85.3 mg, 0.4 mmol), hydroxybenzotriazole (HOBt, 16.2 mg, 0.12 mmol) and 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDCl, 230 mg, 1.2 mmol) in 3 mL anhydrous DMF at RT. The reaction was stirred under Ar for 12 h at RT and then concentrated. The mixture was then purified by silica gel flash chromatography (1:6 methanol:DCM) yielding the coupling product **S2** (Rf = 0.45 in 1:6 methanol:DCM), which was subsequently dissolved in 20 mL TFA-DCM (1:1, v/v) and stirred for 4 h at RT. The solvent was removed *in vacuo* to yield compound **3** (TFA salt, 140 mg, 257 µmol, 64% yield in two steps) as a colorless solid. Rf = 0.05 in 1:6 methonal:DCM. Compound **S2** ^{1}H NMR (400 MHz, $CDCl_3$) δ ppm: 7.66 (s, 1H), 6.39 (s, 2H), 5.43 (d, J = 8 Hz, 1H), 5.06 (s, 1H), 4.29 (q, J = 8, 6.5 Hz, 1H), 4.03 (m, 2H), 3.80 (s, 6H), 3.64 (S, 2H), 3.53 (m, 2H), 2.74 (d, J = 17.1 Hz, 1H), 2.59 (m, 1H), 1.99 (m, 2H), 1.91 (m, 2H), 1.37 (s, 9H). ^{13}C NMR (400 MHz, CD_3OD) δ ppm: 168.60, 166.25, 162.98, 155.08, 153.89, 140.44, 133.94, 111.02, 107.29, 77.33, 74.97, 72.22, 56.68, 53.05, 38.49, 30.79, 25.46, 22.43. Compound **3** HRMS (FAB+) m/z Calcd. for $C_{21}H_{28}O_4N_6$ $[M+H]^+$:429.2172. Found: 429.2249.

Synthesis of compound 2 TMP-Atto520-alkyne. DIEA (15 µL, 89 µmol) was added to a solution of compound **3** (7 mg, 13 µmol) and Atto520-NHS ester (5.0 mg, 8.9 µmol) in 0.5 mL anhydrous DMF. The mixture was stirred under Ar at RT for 12 h before concentrated. Reaction residue was purified by reverse phase HPLC to yield compound **2** (TFA salt, 6 mg, 6.7 µmol, 76% yield). HPLC condition: start with 10:90 acetonitrile:water, gradient elution for 80 min, end with 60:40 acetonitrile:water. Retention time: 40–41 min. ^{1}H NMR (400 MHz, CD_3OD) δ ppm: 7.91 (s, 2H), 7.23 (s, 1H), 6.85 (s, 2H), 6.56 (s, 2H), 4.40 (dd, J = 7.2, 4.8 Hz, 1H), 3.98 (t, J = 5.9 Hz, 2H), 3.81 (s, 6H), 3.73 (m, 2H), 3.65 (s, 2H), 3.52 (q, J = 7.2 Hz, 6H), 3.22 (m, 2H), 2.87 (s, 1H), 2.67 (m, 2H),

2.52 (m, 2H), 2.35 (s, 6H), 2.30 (m, 1H), 1.87 (p, J = 6.3 Hz, 2H), 1.33 (m, 6H). HRMS (FAB+) m/z Calcd. for $C_{43}H_{53}O_8N_8$ M^+:777.41. Found: 777.41.

Synthesis of compound 7. *N*-(3-aminopropyl)-*N*-methylaniline (1 mL, 6.09 mmol) was dissolved in 10 mL DCM. Then DIEA (2.1 mL, 12.18 mmol) and Di-*tert*-butyl dicarbonate (7 mL, 30.45 mmol) were added to the solution and the reaction was stirred under Ar for 20 minutes. Then the solvent was removed *in vacuo* and the product was purified by column chromatography on silica gel (1:5 EtOAc:hexane) to yield compound **7** (1.6 g, 6.05 mmol, 99% yield) as colorless oil. Rf = 0.35 in 1:5 EtOAc:hexane. ^{1}H NMR (400 MHz, $CDCl_3$) δ ppm: 7.23 (m, 2H), 6.70 (m, 3H), 4.56 (s, 1H), 3.40 – 3.32 (t, J = 7.2 Hz, 2H), 3.17 (q, J = 6.8 Hz, 2H), 2.92 (s, 3H), 1.77 (p, J = 7.0 Hz, 2H), 1.45 (s, 9H). ^{13}C NMR (400 MHz, $CDCl_3$) δ ppm: 156.12, 149.29, 129.07, 116.53, 112.62, 79.34, 50.52, 39.82, 38.69, 28.53, 27.41. HRMS (FAB+) m/z Calcd. for $C_{15}H_{24}O_2N_2$ $[M]^+$:264.1838. Found: 264.1826.

Synthesis of compound 6 BHQ1-NH_2. Fast Corinth V Salt (472 mg, dye content 90%, 1.02 mmol) was disolved in 8 mL water and chilled to 0 °C in ice bath. Compound 7 (600 mg, 2.27 mmol) was dissolved in 8 mL methanol and 6 mL 5% $NaHCO_3$ aqueous solution, and then added dropwise to the Fast Corinth V Salt solution. The reaction mixture was stirred in ice bath for 2 h during which the color changed from orange to dark red. The crude product was filtered and washed with 3 X 10 mL of cold water and then dried *in vacuo*. The product was purified by column chromatography on silica gel (4:6 EtOAc:hexane) to yield compound **S3** as dark red solid. Compound **S3** was subsequently deprotected by dissolving in 10 ml 30% TFA in DCM (v/v) and stirring for 1 h; the color changes from red to dark blue. Then the product was obtained by removing solvent *in vacuo*, yielding compound **6** (TFA salt, 360 mg, 0.61 mmol, 60% yield in 2 steps) as dark purple solid. Rf = 0.3 in 1:2 EtOAc:hexane. Compound **S3** ^{1}H NMR (400 MHz, $CDCl_3$) δ ppm: 7.91 (m, 2H), 7.77 (s, 1H), 7.66 (d, J = 8.0 Hz, 2H), 7.58 (s, 1H), 7.47 (m, 2H), 7.40 (s, 1H), 6.4 (m, 2H), 4.57 (s, 1H), 4.02 (s, 3H), 3.49 (t, J = 7.3 Hz, 2H), 3.20 (q, J = 6.7 Hz, 2H), 3.08 (s, 3H), 2.70 (s, 3H), 2.51 (s, 3H), 1.84 (p, J = 7.1 Hz, 2H), 1.46 (s, 9H). Compound **6** HRMS (FAB+) m/z Calcd. for $C_{25}H_{29}O_3N_7$ $[M+H]^+$:476.2332. Found: 476.2408.

Synthesis of compound S4 BHQ1-SO_2Cl. Compound **6** (TFA salt, 100 mg, 0.173 mmol) was dissolved in anhydrous DCM and chilled to 0 °C in ice bath. Then 3-(chlorosulfonyl)benzoyl chloride (55 µL, 0.346 mmol) was added to the solution dropwise, followed by DIEA (0.3 mL, 1.73 mmol). The reaction mixture was stirred in ice bath under Ar for 15 min, and then warmed to RT and stirred for 1 h. The crude product was collected by evaporating the solvent, and was subsequently purified by column chromatography on silica gel (7:3 EtOAc:hexane), yielding compound **S4** (99 mg, 0.145 mmol, 84% yield) as dark red solid. Rf = 0.42 in 2:1 EtOAc:hexane. ^{1}H NMR (400 MHz, $CDCl_3$) δ ppm: 8.38 (t, J = 1.8 Hz, 1H), 8.16 (ddd, J = 8.0, 2.0, 1.1 Hz, 1H), 8.10 (dt, J = 8.0, 1.4 Hz, 1H), 7.91 (m, 2H), 7.70 (m, 2H), 7.66 (d, J = 8.2 Hz, 1H), 7.6 (s, 1H), 7.47 (ddd, J = 8.1, 1.8, 0.9 Hz, 1H), 7.39 (s, 1H), 6.78 (d, J = 8.8 Hz, 2H), 4.01 (s, 3H), 3.59 (m, 4H), 3.10 (s, 3H), 2.70 (s, 3H), 2.51 (s, 3H), 2.03 (p, J = 7.2 Hz, 3H). MS (FAB+) m/z Calcd. for $C_{32}H_{32}ClO_6N_7S$ $[M+H]^+$:678.18. Found: 660.27 (hydrolysis product).

Synthesis of compound 5 BHQ1-azide. Compound S4 (99 mg, 0.146 mmol) and DMAP (18 mg, 0.164 mmol) was dissolved in 2 mL anhydrous DCM. Then 2-azidoethanol (25 µL, 0.292 mmol) and DIEA (128 µL, 0.73 mmol) were added and the reaction mixture was stirred under Ar at RT for 4 h. After removing the solvent *in vacuo*, the product was purified by column chromatography on silica gel (100% EtOAc), yielding compound 5 (65 mg, 89 µmol, 61% yield) as dark red solid. Rf = 0.58 in 100% EtOAc. ^{1}H NMR (400 MHz, $CDCl_3$) δ ppm: 8.30 (s, 1H), 8.11 (m, 1H), 8.07 (m, 1H), 7.93 (m, 3H), 7.71 (m, 2H), 7.66 (m, 2H), 7.58 (s, 1H), 7.48 (m, 1H), 7.39 (s, 1H), 4.24

(t, J = 4.8 Hz, 2H), 4.02 (s, 3H), 3.78 (t, J = 5.0 Hz, 1H), 3.59 (m, 4H), 3.50 (t, J = 4.8 Hz, 2H), 3.12 (s, 3H), 2.70 (s, 3H), 2.52 (s, 3H), 2.02 (m, 2H). ^{13}C NMR (400 MHz, $CDCl_3$) δ ppm: 165.58, 158.31, 156.70, 147.80, 143.44, 142.41, 136.54, 133.74, 132.95, 132.21, 130.62, 130.04, 125.22, 125.03, 134.61, 119.28, 119.12, 100.24, 100.13, 68.99, 59.79, 58.13, 56.56, 49.81, 29.85, 21.50, 16.84. HRMS (FAB+) m/z Calcd. for $C_{34}H_{36}O_7N_{10}S$ $[M+H]^+$:729.2489. Found: 729.2531.

Synthesis of compound 1 TMP-BHQ1-Atto520. Compound **2** (TFA salt, 6 mg, 6.7 µmol) and compound **5** (9 mg, 12.3 µmol) were transferred to the same vial and dissolved in 0.2 mL DMF. Then ascorbic acid (0.68 mg, 3.86 µmol) and $CuSO_4{\cdot}5H_2O$ (0.19 mg, 0.771 µmol) were added and the reaction mixture was stirred under Ar at RT for 24 h. After removing the solvent *in vacuo*, the product was purified by HPLC to yield compound **1** (1.3 mg, 0.86 µmol, 11% yield) as dark purple solid. HPLC condition: start with 30:70 acetonitrile:water, gradient elution for 80 min, end with 80:20 acetonitrile:water. Retention time: 49-51 min. Rf = 0.72 in 1:4 methanol:DCM with 0.5% acetic acid (v/v). ^{1}H NMR (400 MHz, CD_3OD) δ ppm: 8.86 (t, J = 5.7 Hz, 1H), 8.24 – 8.14 (m, 3H), 8.05 (t, J = 5.6 Hz, 1H), 7.93 (m, 1H), 7.79 (dd, J = 1.8, 0.9 Hz, 1H), 7.73 – 7.58 (m, 7H), 7.28 (d, J = 0.8 Hz, 1H), 7.22 (d, J = 1.1 Hz, 1H), 7.17 (s, 1H), 6.61 (s, 2H), 6.56 (s, 2H), 4.66 – 4.56 (m, 3H), 4.50 (m, 2H), 3.92 (m, 2H), 3.86 (s, 3H), 3.80 (s, 6H), 3.63 (s, 2H), 3.62 – 3.56 (m, 3H), 3.55 – 3.50 (m, 2H), 3.38 (q, J = 7.2 Hz, 4H), 3.33 (d, J = 1.6 Hz, 2H), 3.11 (s, 3H), 3.06 (dd, J = 14.9, 5.7 Hz, 1H), 2.90 (dd, J = 14.9, 8.3 Hz, 1H), 2.57 (s, 3H), 2.54 (s, 3H), 2.18 (s, 6H), 2.01 (m, 2H), 1.82 (m, 2H), 1.29 (t, J = 7.2 Hz, 6H). MS (ESI) m/z Calcd. for $C_{77}H_{89}O_{13}N_{18}S$ $[M]^+$:1505.66. Found: 1506.5.

The TMP-Q-Atto520 is dissolved in anhydrous DMF and the stock solution can be stored at -80 °C for long-term (> 18 months) or -20 °C for short-term (up to 6 months) storage. It is recommended to use amber glass vials sealed by parafilm. To use the TMP-Q-Atto520 stock solution, warm up the container to room temperature before opening the vial. Do not store the compound in aqueous solution.

Synthesis of compound S5 TMP-fluorescein-alkyne. DIEA (188 µL, 1.08 mmol) was added dropwise to a solution of compound **3** (59 mg, 0.108 mmol, in the form of TFA salt), 5-carboxyfluorescein (40 mg, 0.108 mmol), HOBt (4.4 mg, 32.4 µmol) and EDCl (62.1 mg, 0.324 mmol) in 2 mL anhydrous DMF at RT. The reaction was stirred under Ar for 12 h at RT and then concentrated. The mixture was first purified by silica gel flash chromatography (1:4 methanol:DCM with 1% acetic acid) and then refined by HPLC, yielding the coupling product **S5** (16 mg, 20.3 µmol, 19%) as orange solid. HPLC condition: start with 27:73 acetonitrile:water, gradient elution for 50 min, end with 60:40 acetonitrile:water. Retention time: 10–11 min. Rf = 0.25 in 1:4 methanol:DCM with 1% acetic acid. ^{1}H NMR (400 MHz, $CDCl_3$) δ ppm: 8.34 (d, J = 1.7 Hz, 1H), 8.14 (m, 1H), 7.23 (m, 2H), 6.71 (s, 2H), 6.55 (m, 4H), 6.51 (m, 2H), 4.77 (m, 1H), 4.02 (m, 2H), 3.77 (s, 6H), 3.60 (s, 2H), 3.53 (m, 2), 2.91 (m, 1H), 2.75 (m, 1H), 2.39 (t, J = 2.6 Hz, 1H), 1.91 (p, J = 6.3 Hz, 2H). Compound **3** MS (FAB+) m/z Calcd. for $C_{42}H_{38}O_{10}N_6$ $[M+H]^+$:787.26. Found: 787.23.

Synthesis of compound S6 TMP-BHQ1-fluorescein. Compound **S5** (10 mg, 12.7 µmol) and compound **5** (9.3 mg, 12.7 µmol) were transferred to the same vial and dissolved in 0.2 mL DMF. Then ascorbic acid (1.1 mg, 6.35 µmol) and $CuSO_4{\cdot}5H_2O$ (0.32 mg, 1.27 µmol) were added and the reaction mixture was stirred under Ar at RT for 24 h. After removing the solvent *in vacuo*, the product was purified by HPLC to yield compound **S6** (4.7 mg, 3.1 µmol, 24% yield) as dark purple solid. HPLC condition: start with 30:70 acetonitrile:water, gradient elution for 80 min, end with 60:40 acetonitrile:water. Retention time: 67-68 min. Rf = 0.53 in 1:3 methanol:DCM with 0.5% acetic acid (v/v). ^{1}H NMR (400 MHz, CD_3OD) δ ppm: 8.32 (s, 1H), 8.12 (m, 3H), 7.87 (dt, J = 8.0, 1.4 Hz, 1H), 7.76 (m, 4H), 7.70 (d, J = 8.2 Hz, 1H), 7.67 (t, J = 7.8 Hz, 1H), 7.58 (dd, J = 8.0, 1.6 Hz, 1H), 7.50 (s, 1H), 7.33 (s, 1H), 7.17 (m, 2H), 6.85 (d, J = 9.0 Hz, 2H), 6.69 (d, J = 2.1 Hz, 2H), 6.57 – 6.49 (m, 4H),

6.47 (s, 2H), 4.93 (dd, J = 9.5, 4.9 Hz, 1H), 4.63 (t, J = 4.7 Hz, 2H), 4.46 (m, 2H), 4.01 – 3.89 (m, 5H), 3.74 (s, 6H), 3.63 (t, J = 7.5 Hz, 2H), 3.56 (s, 2H), 3.50 (m, 4H), 3.38 – 3.18 (m, 2H), 3.14 (s, 3H), 2.65 (s, 3H), 2.51 (s, 3H), 2.01 (m, 2H), 1.867 (m, 2H). MS (ESI) m/z Calcd. for $C_{76}H_{74}O_{17}N_{16}S$ $[M+H]^+$:1514.51. Found: 1515.04.

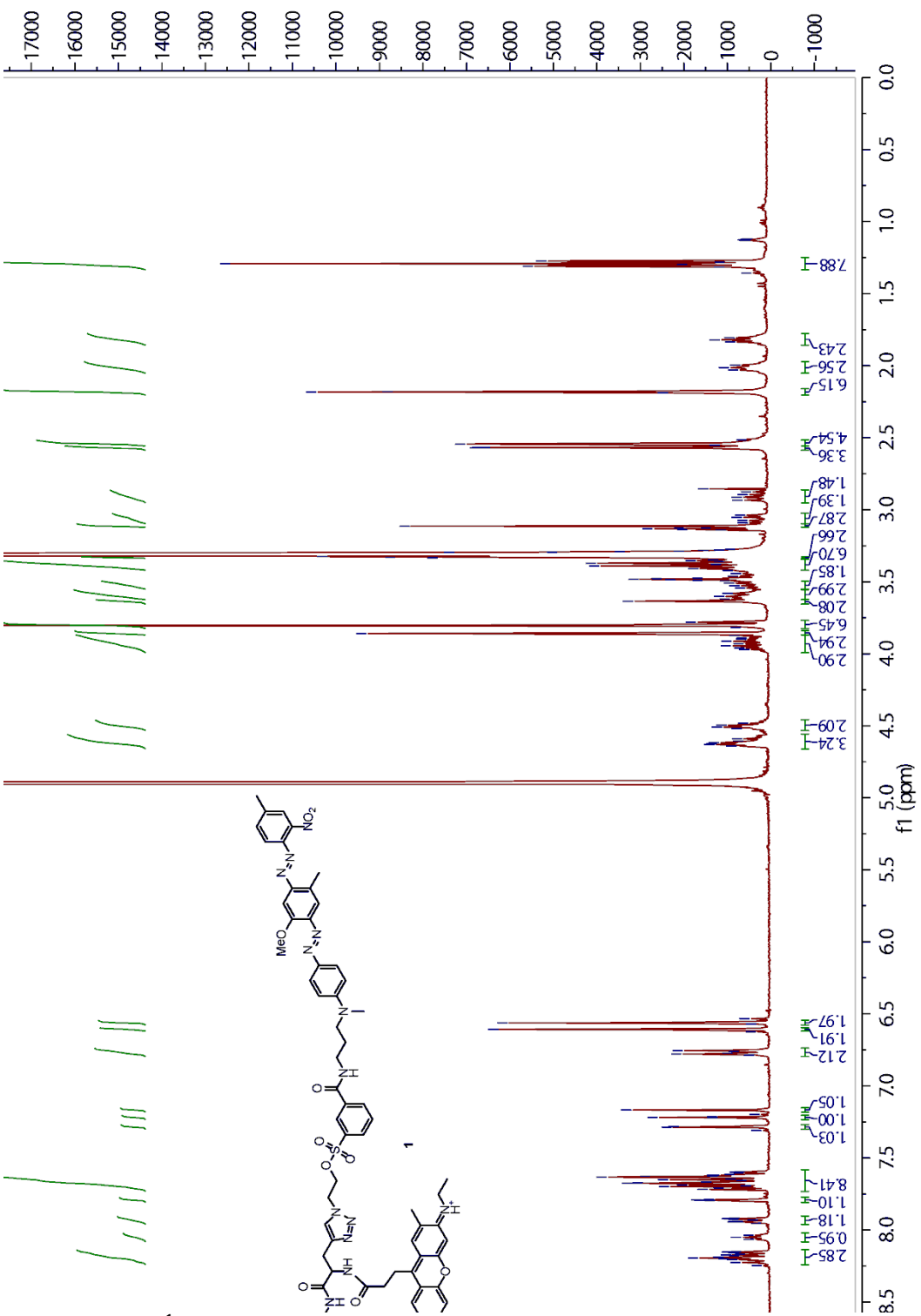


Figure S1: ^{1}H-NMR of compound **1** TMP-Q-Atto520 (400 MHz, CD_3OD)

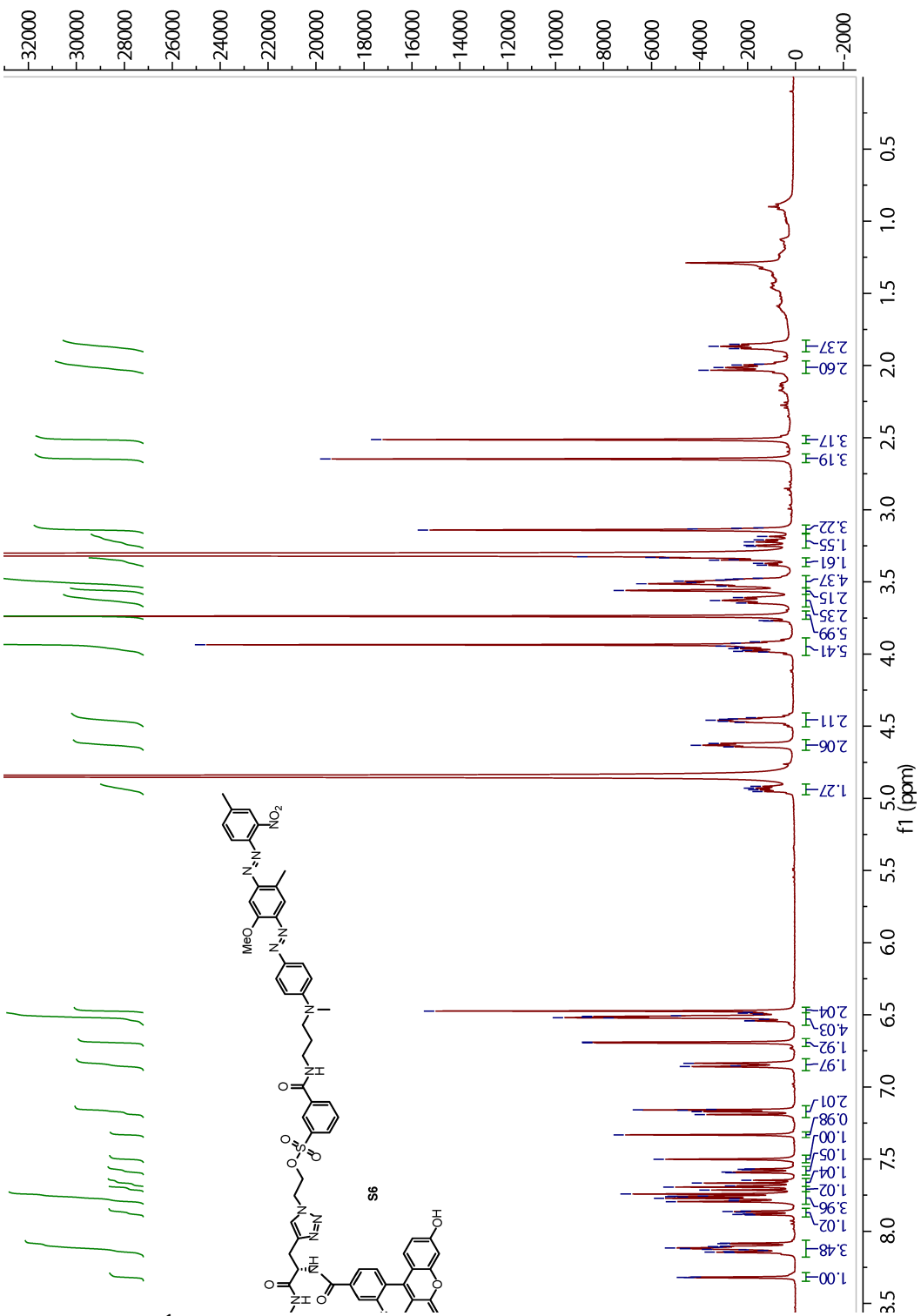


Figure S2: ^{1}H-NMR of compound **S6** TMP-Q-fluorescein (400 MHz, CD_3OD).

Site-Directed Mutagenesis

Genes encoding eDHFR variants were produced using Strategene's QuikChange Mutagenesis Kit. The parent vector was the previously published p2247 (eDHFR C85S and C152S in pAED4), encoding eDHFR-2C. 16 individual rounds of mutagenesis yielded 16 eDHFR variants with additional Cys mutations: E17C, N18C, A19C, M20C, P21C, N23C, P25C, L28C, A29C, K32C, S49C, I50C, P51C, R52C, L54C, P55C, respectively. The forward primers for site-directed mutagenesis to generate 16 eDHFR:Cys mutants are listed in Table S2. The reverser primers are anti-parallel (reverse complementary) to the forward primers.

Table S2. Primer Sequences for Site-Directed Mutagenesis

Mutation	Forward Primer Sequence (5' → 3')
E17C	GAT CGC GTT ATC GGC ATG TGC AAC GCC ATG CCG TGG AAC
N18C	CGG CAT GGA ATG CGC CAT GCC ATG GAA CCT GCC TG
A19C	GTT ATC GGC ATGGAA AAC TGC ATG CCG TGG AAC CTG CC
M20C	GGC ATG GAA AAC GCC TGT CCG TGG AAC CTG CCT G
P21C	GGC ATG GAA AAC GCC ATG TGT TGG AAC CTG CCT GCC
N23C	CGC CAT GCC GTG GTG CCT GCC TGC CGA T
P25C	CAT GCC GTG GAA CCT GTG TGC AGA TCT CGC CTG GTT TAA ACG
L28C	GGA ACC TGC CTG CCG ATT GCG CAT GGT TTA AAC GCA AC
A29C	GGA ACC TGC CTG CAG ATC TCT GCT GGT TTA AAC GCA ACA C
K32C	CTG CCG ATC TCG CCT GGT TTT GCC GCA ACA CCT TAA ATA AAC C
S49C	GCC ATA CCT GGG AAT GCA TCG GTC GTC CGT TGC
I50C	CAT ACC TGG GAA TCA TGC GGC CGT CCG TTG CCA GG
G51C	CCT GGG AAT CAA TCT GCC GTC CGT TGC CAG G
R52C	CTG GGA ATC AAT CGG GTG CCC GTT GCC AGG ACG C
L54C	CAA TCG GTC GTC CGT GCC CAG GAC GCA AAA ATA TTA TCC
P55C	CAA TCG GTC GTC CGT TGT GTG GAC GCA AAA ATA TTA TCC TCA G

Biochemistry and Cell Biology Methods

Protein purification. The resulting eDHFR:Cys plasmid was expressed in BL21(DE3)pLysS cells (Invitrogen). Cells were grown at 37 ˚C to an OD_{600} of 0.6 – 0.8 and then induced with 0.4 mM IPTG for three hours and purified using nickel-NTA spin columns (Qiagen). All 16 recombinant eDHFR:Cys variants were judged to be >95% pure by Coomassie staining of a SDS-PAGE gel. The protein was dialyzed three times in phosphate buffered saline (PBS) at 4 ˚C, snap frozen and stored at –80 ˚C.

***In vitro* fluorometric measurement.** Purified eDHFR:Cys variants at a concentration of 1 μM were incubated with 1 μM TMP-Q-Atto520 in PBS with 100 μM NADPH and 1 mM glutathione at 37 ˚C. Fluorescence intensity was measured at various time points in a 96-well plate using a Tecan Infinite® M200 fluorometer, excited at 520 nm (bandwidth 9 nm) and emission collected at 550 nm (bandwidth 20 nm). To take fluorescence spectra, eDHFR:L28C (1 μM) was incubated with TMP-Q-Atto520 (1 μM) at 37 ˚C for 3 h. For emission spectrum, the excitation wavelength was 488 nm. For excitation spectrum (Supporting Information Figure S4), the emission wavelength was 580 nm.

***In vitro* alkylation reactions.** Purified eDHFR:L28C (2 μM) was incubated with TMP-Q-Atto520 (10 μM) in PBS with reduced 100 μM NADPH and glutathione (1 mM) at 37 ˚C. At selected time points, aliquots (30 μL) were removed from the reaction mixture, quenched with 6X SDS. Samples from the *in vitro* alkylation experiments

were run on Criterion 15% Tris-HCl gels (BioRad) for 55 min at 200 V. Bands were quantified by densitometry analysis of Coomassie stained gels by ImageJ.[1] The labeled proteins were also detected by in-gel fluorescence scanning using a Typhoon™ Trio scanner excited by the 488 nm laser and an emission filter of 520 (BP 40) nm.

Construction of eDHFR:L28C fusion proteins. The plasmid encoding H2B-mCherry was purchased from Addgene (plasmid 20972). The construction of H2B-eDHFR, H2B-eDHFR:L28C, TOMM20-eDHFR:L28C, MLC-eDHFR:L28C and α-actinin-eDHFR:L28C has been previously published.[2, 3]

Cell culture and transfection. HEK 293T cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) w/ glutamine (Gibco #11995) with 10% v/v fetal bovine serum (FBS) and 1% v/v Pen/Strep. All cells were maintained under 5% CO_2 at 37 °C. For live cell protein labeling, cells were plated in 8-well chambered #1 borosilicate coverglass (Thermo, Nunc 155411) 24 h before transfected with expression plasmids for eDHFR:L28C fused target protein (0.4 μg DNA for one well) using Fugene HD (Roche). MEF cells were cultured in DMEM (Gibco #11965) with 10% v/v FBS, 1.5% v/v HEPES and 1% v/v Pen/Strep. All cells were maintained under 5% CO_2 at 37 °C. Transient transfection of plasmids was performed 1 day before labeling experiment through electroporation using a Nucleofector (Lonza VPD-1004) kit according to the manufacturer's protocol (MEF 1 kit and program T-20).

Intracellular protein labeling. 24 h after transfection, cells were stained with TMP-Q-Atto520 in 200 μL supplemented DMEM media in 8-well chambered coverglass (5 μM of TMP-Q-Atto520 for H2B labeling; 10 μM for other target proteins). Cells were incubated with this staining solution for 3 h at 37 °C, followed by washing with fresh media for 2 times before imaging or cellular analysis.

Live cell imaging. Confocal images were obtained using LEICA TCS SP5 confocal microscope with a HCX PL APO CS 20x 0.70 dry UV objective or a HCX PL APO CS 100x 1.46 oil immersion objective. Green channel was excited with a 488 nm laser and emission collected between 520 – 580 nm; red channel was excited with a 594 nm laser and emission collected between 620 – 680 nm. Images were processed by LAS AF software.

In-gel fluorescence scanning. Cells labeled with TMP-Q-Atto520 were trypsinized and the cell density was determined. The cells were then pelleted by centrifugation at 1000 rpm for 4 minutes. The supernatant was removed, and the pellet was washed with 1 mL of PBS/1x10^6 cells and repelleted. Again, the supernatant was removed and the pellet was resuspended in lysis buffer to a concentration of 1x10^6 cells/12.5 μL of lysis buffer and lysed on ice for 20 minutes. The lysate was then centrifuged at 13,000 rpm for 10 minutes and the supernatant was mixed with loading buffer and heated to 95 ˚C for five minutes. Samples from the live cell staining were run on Criterion 15% Tris-HCl gels (BioRad) for 1.5 hours at 150 V. The BenchMark™ Fluorescent Protein Standard from Invitrogen (Catalog No: LC5928) was used to determine the protein molecular weights. The labeled proteins were also detected by in-gel fluorescence scanning using a Typhoon™ Trio scanner excited by the 488 nm laser and an emission filter of 520 (BP 40) nm.

In Vitro Fluorometry Characterization of TMP-Q-Atto520

Figure S3 shows the *in vitro* screening of the eDHFR:Cys variant library with TMP-Q-Atto520.

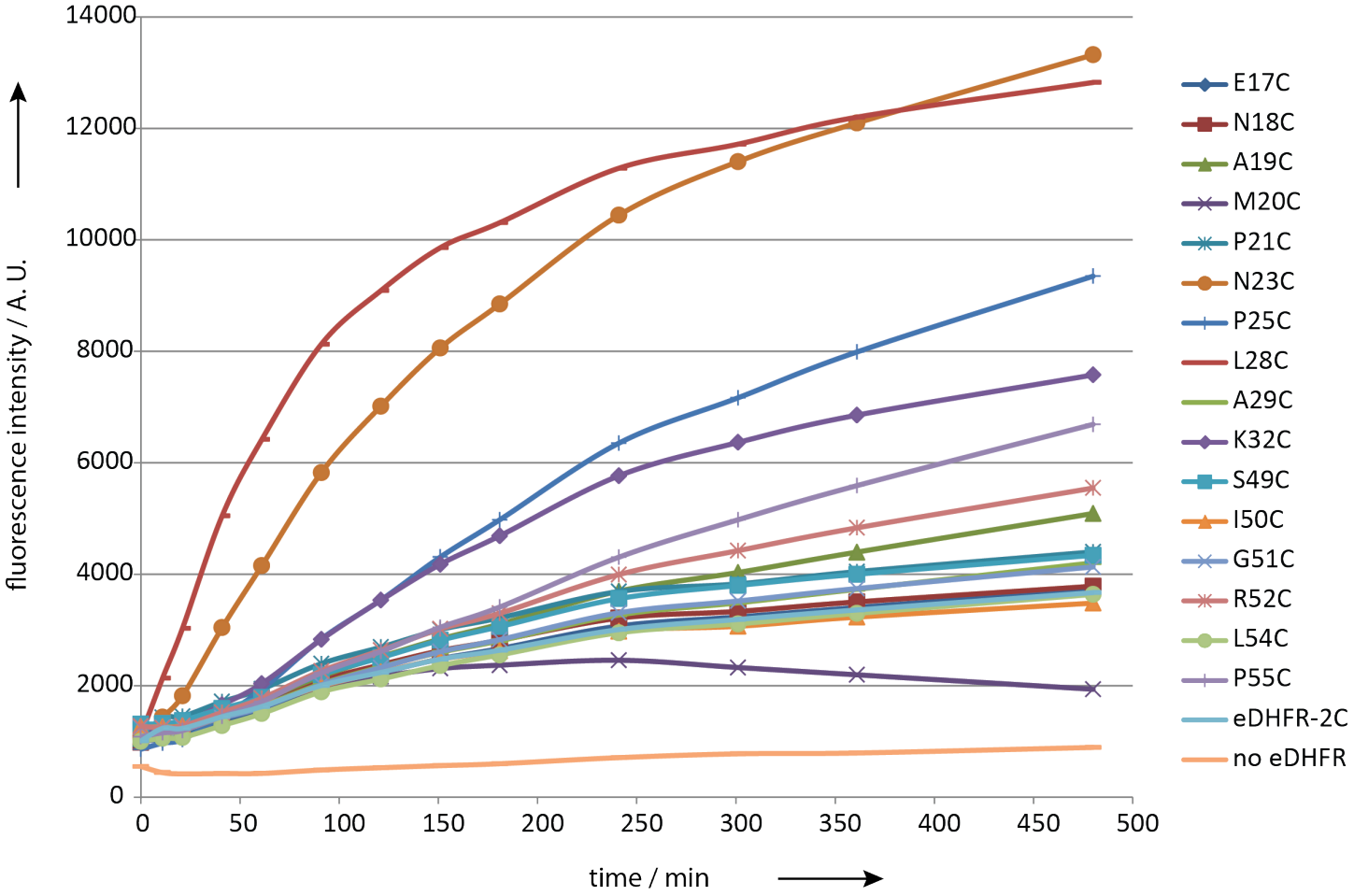


Figure S3. Screening of the eDHFR:Cys variant library with TMP-Q-Atto520. Purified eDHFR:Cys variants at a concentration of 1 µM were incubated with 1 µM TMP-Q-Atto520 in PBS with 100 µM NADPH and 1 mM glutathione at 37 ˚C. Fluorescence intensity was measured at various time points. As controls, the eDHFR-2C variant that lacks Cys and the buffer without eDHFR induced little increase in fluorescence intensity.

Figure S4 shows the fluorescence excitation spectra of TMP-Q-Atto520 labeling eDHFR:L28C.

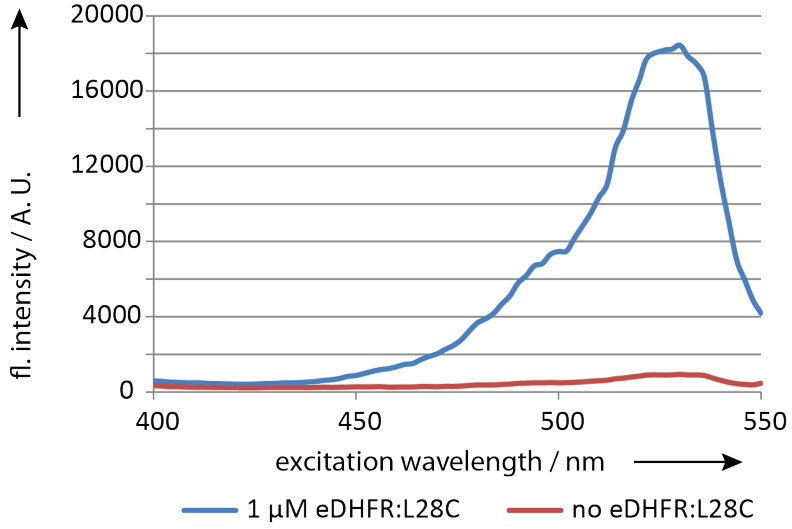


Figure S4. The fluorescence excitation spectra of TMP-Q-Atto520. Purified eDHFR:L28C at 1 µM concentration was labeled with 1 µM TMP-Q-Atto520 in PBS buffer supplemented with 1 mM GSH and 100 µM NADPH, incubated at 37 °C for 3 h. The fluorescence spectrum showed 20 fold enhancement compared to the control group in which 1 µM TMP-Q-Atto520 was incubated with the buffer without eDHFR:L28C.

Figure S5 shows the results of TMP competition assay. 1–10 equivalents of TMP could efficiently compete off labeling of eDHFR:L28C by TMP-Q-Atto520.

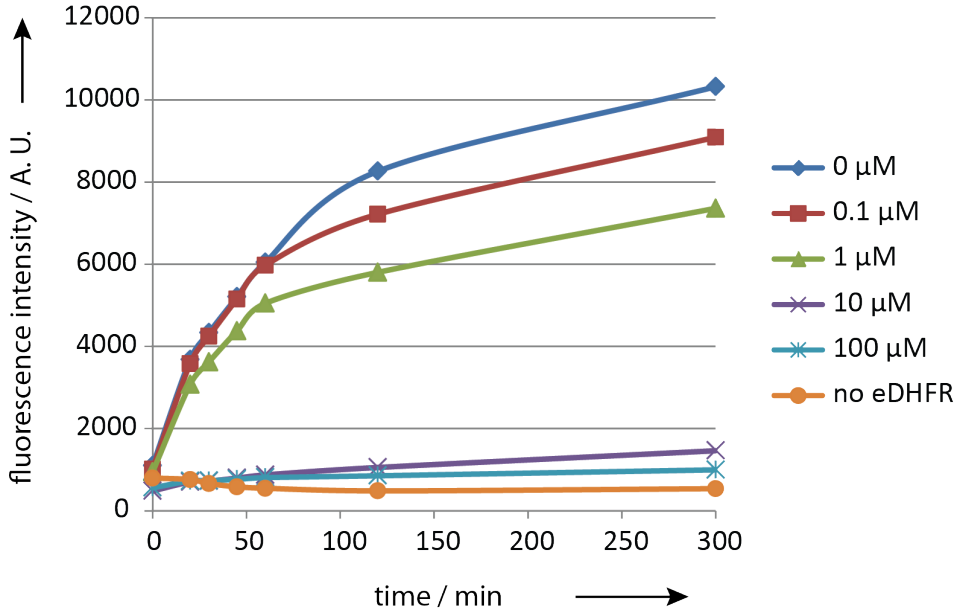


Figure S5. TMP competition assay. Purified eDHFR:L28C at 1 µM concentration was labeled with 1 µM TMP-Q-Atto520 in presence of TMP at various concentration (see legend), under the same condition as in Figure S4. The TMP-Q-Atto520 incubated with buffer without eDHFR:L28C was used as a negative control.

MALDI-TOF and LC-MS/MS Analysis

MALDI-TOF analysis was performed to further confirm that eDHFR:L28C was covalently labeled by fluorogenic TMP-tag. MALDI-TOF analysis was carried out on an AB Sciex Voyager DE Pro MALDI mass spectrometer with the following instrument parameters: accelerating voltage, 25,000 V; grid voltage, 93%; guide wire 0.15%; extraction delay time, 200 ns; laser intensity 2100 - 2200. The MH^+ and $M2H^{2+}$ ions of myoglobin were used for internal two-point calibration. Matrix solution for intact protein analysis was prepared by dissolving sinapinic acid in 50% acetonitrile/0.1% TFA (v/v) at a concentration of 10 mg/mL. A 1 µL aliquot of unlabeled protein (0.21 mM in PBS) was mixed with 5 µL of matrix and 0.6 µL was spotted onto a sample plate and allowed to dry. When the spot was dry, it was washed with 2 µL of water and dried again. A calibration standard of 130 fmol myoglobin in sinapinic acid matrix was applied on top of the protein spot and allowed to dry completely. The same procedure was followed for the TMP-Q-fluorescein labeled protein: 1 µL (0.07 mM in PBS) was mixed with 5 µL of matrix solution and spotted on the sample plate. The spot was dried and washed, dried again and spotted on top with the myoglobin standard solution.

The labeling residue was identified by chymotrypsin digestion followed by LC-MS/MS. For chymotrypsin digestion, eDHFR:L28C labeled with TMP-Q-fluorescein was first purified by SDS-PAGE and gel bands were transferred to a clean tube and 100 µL 0.01 M DTT/0.1 M Tris, pH 8.5 was added. The tube was placed in a

heating block at 55° for 1 – 2 h. After cooling the tube to room temperature, the liquid was removed and replaced with 100 µL 0.015 M iodoacetamide/0.1 M Tris, pH 8.5. This was allowed to react for 30 min. in the dark after which the liquid was removed and the gel was washed as described below. The gel was prepared for digestion by washing once with 200 µL 0.05 M Tris, pH 8.5/25% acetonitrile (v/v) and twice with 200 µL 0.05 M Tris, pH 8.5/50% acetonitrile (v/v) for 20 min. with shaking. After removing the washes, the gel piece was dried for 30 min in a Speed-Vac concentrator.

In-gel digestion was carried out by adding 0.1 µg chymotrypsin (sequencing grade, Roche Molecular Biochemicals) in 15 µL 0.025 M Tris, pH 8.5. The tube was placed in a heating block at 32° and left overnight. Peptides were extracted with 2X 50 µL 50% acetonitrile/2% TFA (v/v) and the combined extracts were dried in a Speed-vac concentrator then re-dissolved in 20 µL 0.2% formic acid (v/v).

LC-MS/MS analysis was carried out on a Waters Q-Tof Ultima hybrid quadrupole/time-of-flight mass spectrometer with a nanoelectrospray source. Capillary voltage was set at 1.8 kV and cone voltage 32 V; collision energy was set according to mass and charge of the ion, from 14 eV to 50 eV. Chromatography was performed on an LC Packings HPLC with a C18 Inertsil column (75 µm x 15 cm) using a linear acetonitrile gradient with flow rate of 200 nL/min. The molecular mass spectrum of the labeled peptide was created from the raw MS/MS spectrum using the MaxEnt 3 program in the Waters MassLynx software suite.

Sequence of recombinant eDHFR:L28C:

^{1}MISLIAALAV ^{11}DRVIGMENAM ^{21}PWNLPAD**C**AW ^{31}FKRNTLNKPV ^{41}IMGRHTWESI ^{51}GRPLPGRKNI
^{61}ILSSQPGTDD ^{71}RVTWVKSVDE ^{81}AIAACGDVPE ^{91}IMVIGGGRVY ^{101}EQFLPKAQKL ^{111}YLTHIDAEVE
^{121}GDTHFPDYEP ^{131}DDWESVFSEF ^{141}HDADAQNSHS ^{151}YCFEILERRH ^{161}HHHHH

Chymotrypsin cleaves the C-terminus of large hydrophobic residues (Tyr/Phe/Trp-Xaa), indicated by red letters. L28C mutation is highlighted as bold and underlined letter.

Figure S6 shows MS graphs of intact eDHFR:L28C; eDHFR:L28C conjugated to TMP-Q-fluorescein; and peptide with TMP-Q-fluorescein label generated by chymotrypsin digestion.

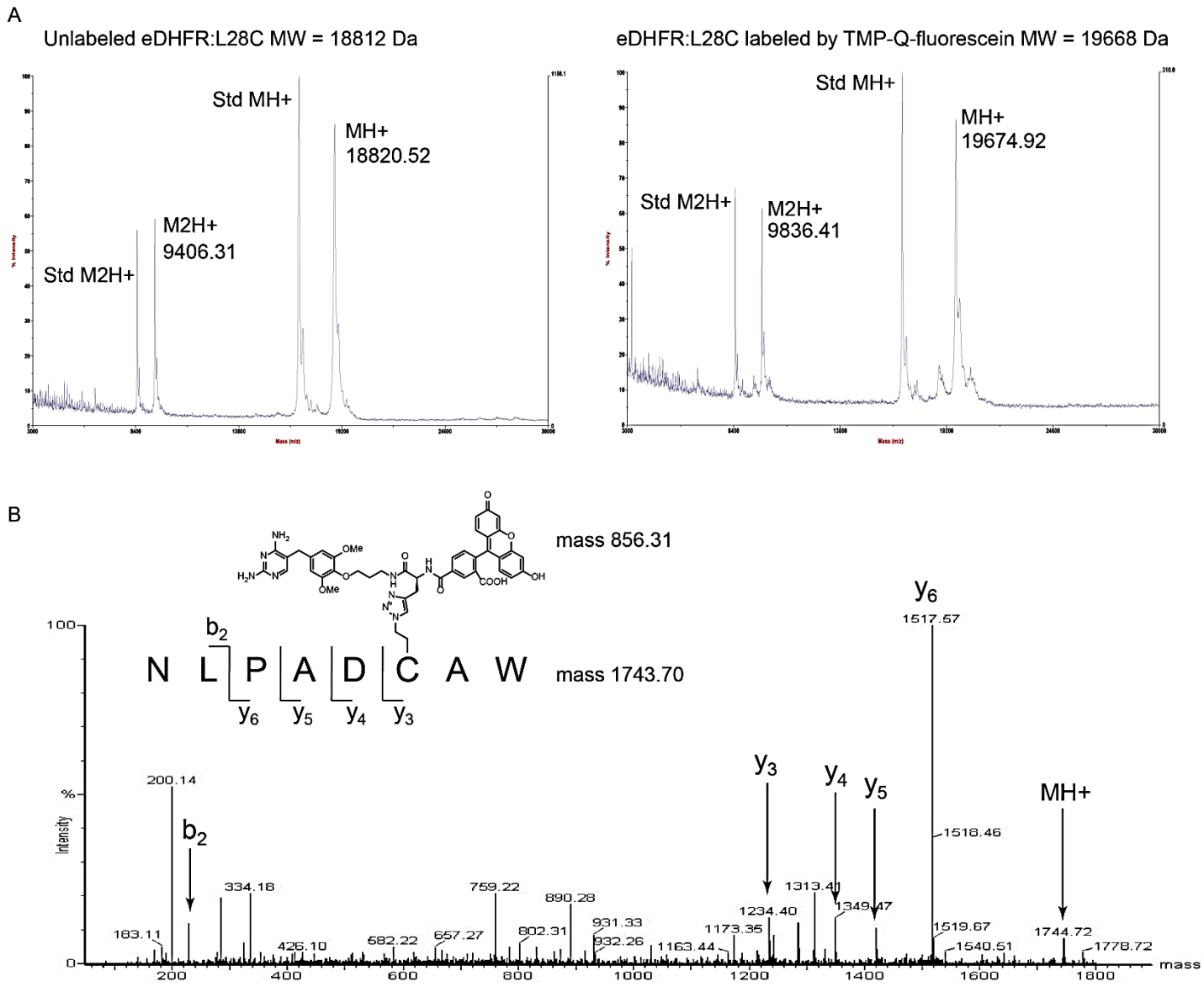


Figure S6. MS analysis of eDHFR:L28C labeled with TMP-Q-fluorescein. (A) MALDI-TOF graphs of unlabeled eDHFR:L28C (left) and eDHFR:L28C labeled with TMP-Q-fluorescein (right), respectively. The increase in MW confirmed the covalent labeling with fluorogenic TMP-tag. (B) MS/MS graph of labeled peptide generated by chymotrypsin digestion. The fragmentation pattern was consistent with covalent labeling on L28C.

In Vitro Labeling Kinetics

Hypothesized mechanism for the proximity-induced S_N2 reaction:

$$R + L \underset{k_3}{\overset{k_1}{\rightleftharpoons}} RL \xrightarrow{k_2} P$$

R: eDHFR:L28C (receptor)
L: TMP-Q-Atto520 (ligand)
RL: TMP-Q-Atto520 non-covalently binding to eDHFR:L28C
P: eDHFR:L28C covalently labeled with TMP-Q-Atto520

Derive the rate equation using steady-state hypothesis:

$$\frac{d[RL]}{dt} = k_1[R][L] - k_2[RL] - k_3[RL] = 0 \qquad (1)$$

$$[RL] = \frac{k_1[R][L]}{k_2+k_3} \quad (2)$$

The reaction rate $r = -\frac{d[R]}{dt} = \frac{d[P]}{dt} = k_2[RL] = \frac{k_1k_2[R][L]}{k_2+k_3} = k_{app}[R][L]$ (3)

In which $k_{app} = \frac{k_1k_2}{k_2+k_3}$.

Initial concentration: $[R]_0 = [L]_0 = 1.0\times10^{-5}$ M;

Therefore during the reaction at any given time: $[R] = [L]$;

Which is plugged into equation (3) to get: $-\frac{d[R]}{dt} = k_{app}[R]^2$ (4)

Integrated rate equation: $\frac{1}{[R]} = \frac{1}{[R]_0} + k_{app}t$ (5)

In gel-shift assay, [R] is determined as: $[R] = [R]_0 \frac{I_{upper}}{I_{upper}+I_{lower}}$ (6)

In which I_{upper} and I_{lower} represents the intensity of the upper and lower bands, respectively.

Linear regression of $\frac{1}{[eDHFR{:}L28C]}$ versus time should result in k_{app} as the slope.

Figure S7 shows the gel-shift assay and linear regression to determine k_{app}.

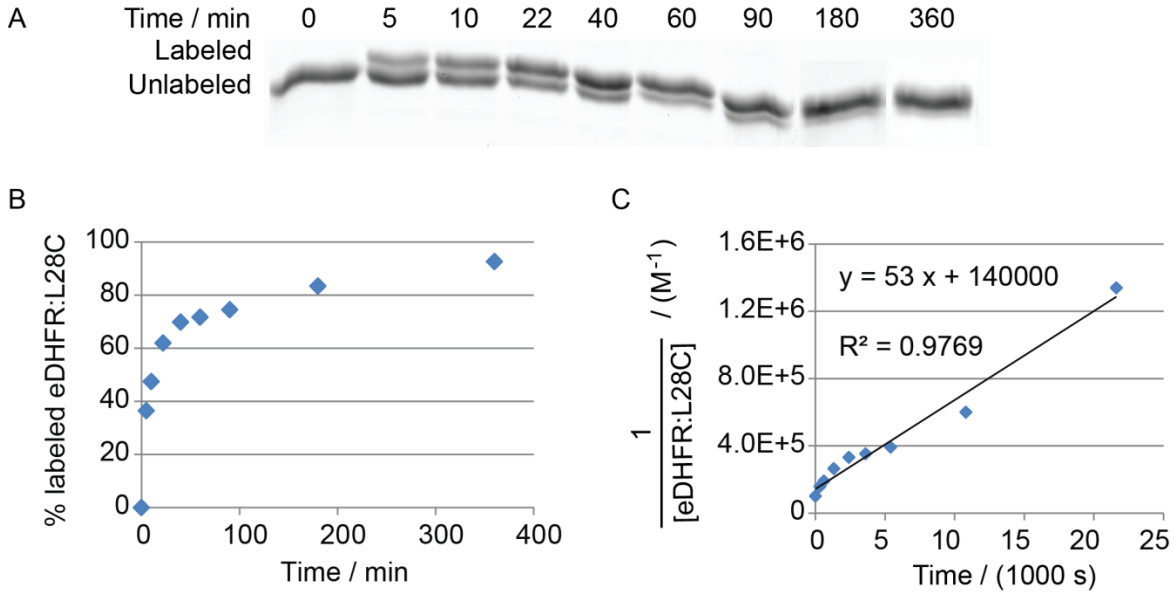


Figure S7. *In vitro* labeling kinetics of fluorogenic TMP-tag. (A) The gel-shift assay was performed using 10 µM TMP-Q-Atto520 and 10 µM eDHFR:L28C in the same buffer as in figure S3. (B) (C) Density of bands were quantified by ImageJ and fitted to equation (5). The k_{app} was determined to be 53 $M^{-1}s^{-1}$ as the slope of the trend line in (C).

In Vitro Characterization of TMP-Q-fluorescein

Figure S8 shows the *in vitro* screening of the eDHFR:Cys variant library with TMP-Q-fluorescein.

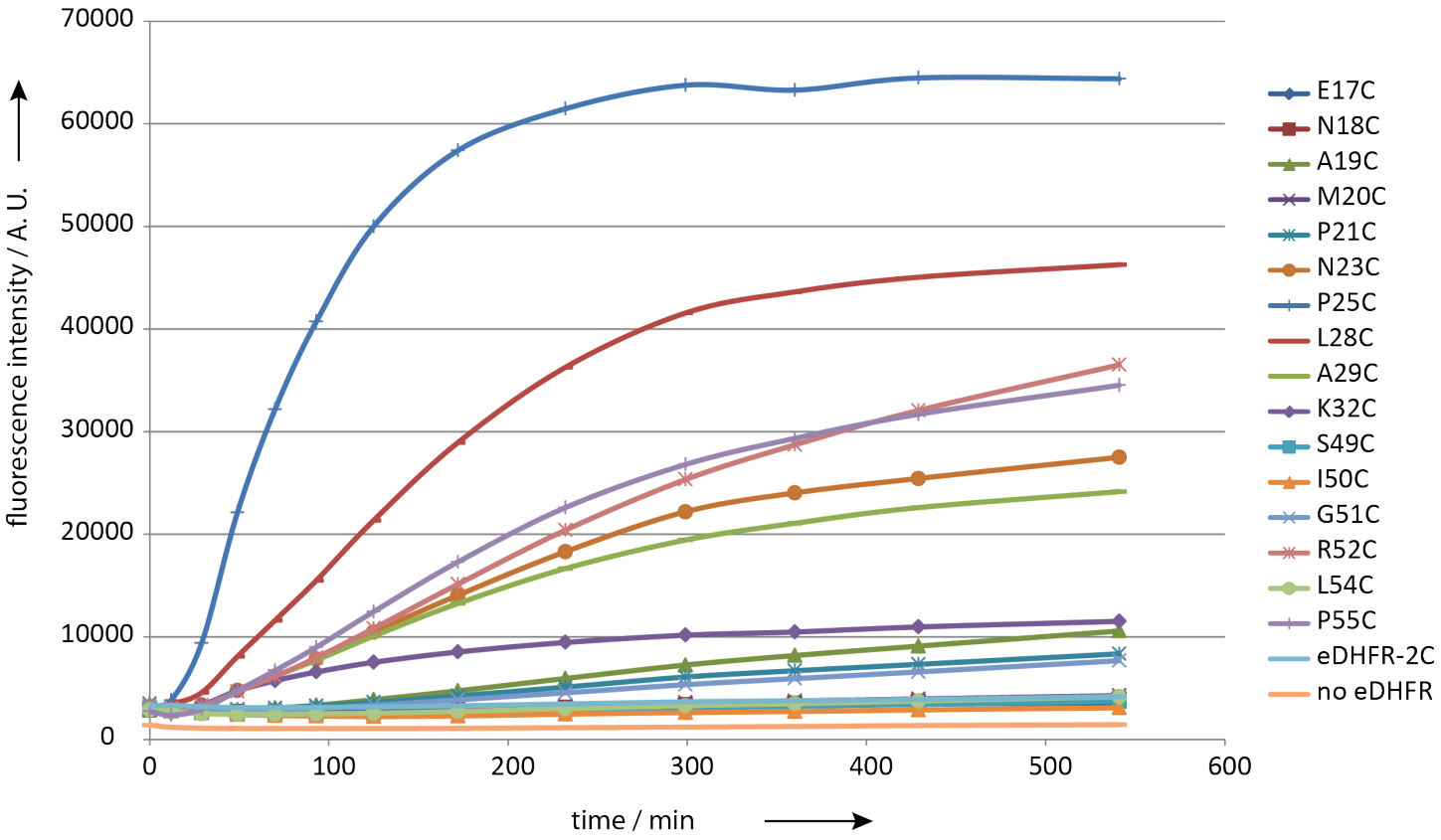


Figure S8. Screening of the eDHFR:Cys variant library with TMP-Q-fluorescein. Purified eDHFR:Cys variants at a concentration of 1 µM were incubated with 1 µM TMP-Q-fluorescein under the same condition as in Figure S3. Fluorescence intensity was measured at various time points. As controls, the eDHFR-2C variant that lacks Cys and the buffer without eDHFR induced little increase in fluorescence intensity.

Figure S9 shows the fluorescence excitation and emission spectra of TMP-Q-fluorescein labeling eDHFR:P25C.

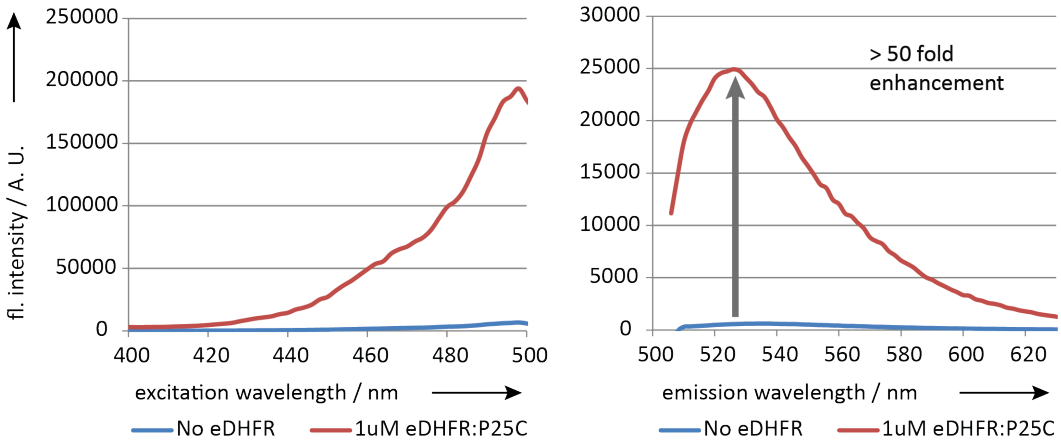


Figure S9. The fluorescence excitation and emission spectra of TMP-Q-fluorescein. Purified eDHFR:P25C at 1 µM concentration was labeled with 1 µM TMP-Q-fluorescein under the same condition as in Figure S4. The fluorescence spectrum showed > 50 fold enhancement compared to the control group in which 1 µM TMP-Q-fluorescein was incubated with the buffer without eDHFR:P25C.

Figure S10 shows the results of gel shift assay of TMP-Q-fluorescein labeling eDHFR:P25C and eDHFR:L28C.

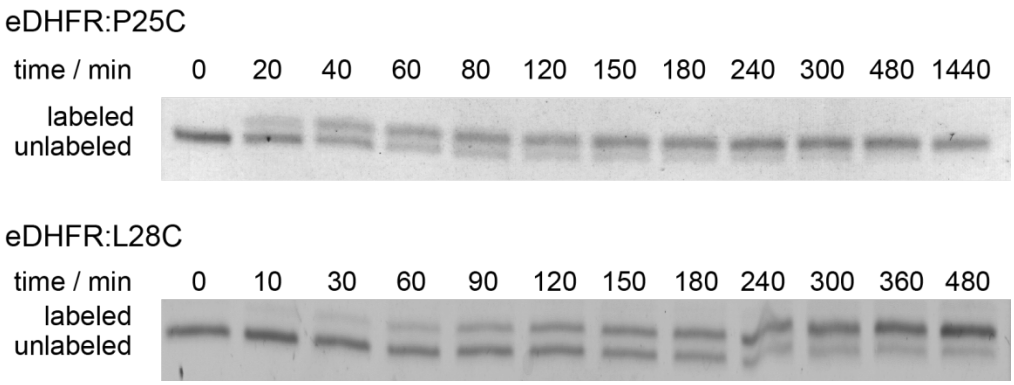


Figure S10. Gel shift assay of TMP-Q-fluorescein. In the same buffer as in Figure S3, 2 µM of eDHFR:P25C or eDHFR:L28C was incubated with 10 µM TMP-Q-fluorescein at 37 °C. At various time points, aliquots were quenched with 6X SDS. The reaction was analyzed by SDS-PAGE followed by Coomassie staining, and it was determined that 50% labeling occurs in approximately 50 minutes with eDHFR:P25C and 150 min with eDHFR:L28C, respectively.

Live Cell Imaging Using TMP-Q-fluorescein

Figure S11 shows live cell imaging of TMP-Q-fluorescein labeling eDHFR:L28C fused to H2B and plasma membrane (PM) localization sequence, respectively.

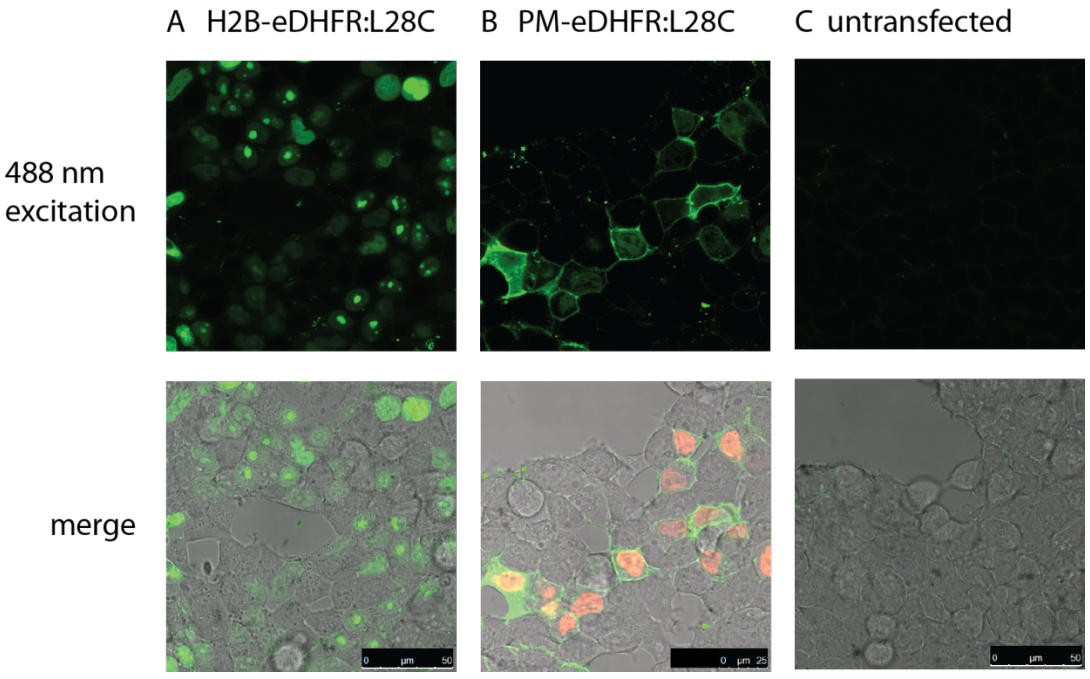


Figure S11. Live cell imaging using TMP-Q-fluorescein. HEK 293T cells were transiently transfected with plasmid encoding H2B-eDHFR:L28C (A) and PM-eDHFR:L28C (B), respectively, and labeled with 10 µM TMP-Q-fluorescein for 3 h. Live cell imaging was achieved using confocal microscope, displaying the specific labeling of cell nucleus (A) and plasma membrane (B), respectively. For PM labeling, the cells were co-transfected with H2B-mCherry as transfection indicator. (C) As control, untransfected cells were treated with TMP-Q-

fluorescein under the same condition as (A), showing very weak background fluorescence. For comparison, images in A and C were obtained using the same microscope setup with excitation at 488 nm. Black bars: 50 µm.

Variation on Live Cell Imaging Protocol

Figure S12 shows fluorescent images of live HEK 293T cells labeled with TMP-Q-Atto520 using alternative staining protocols.

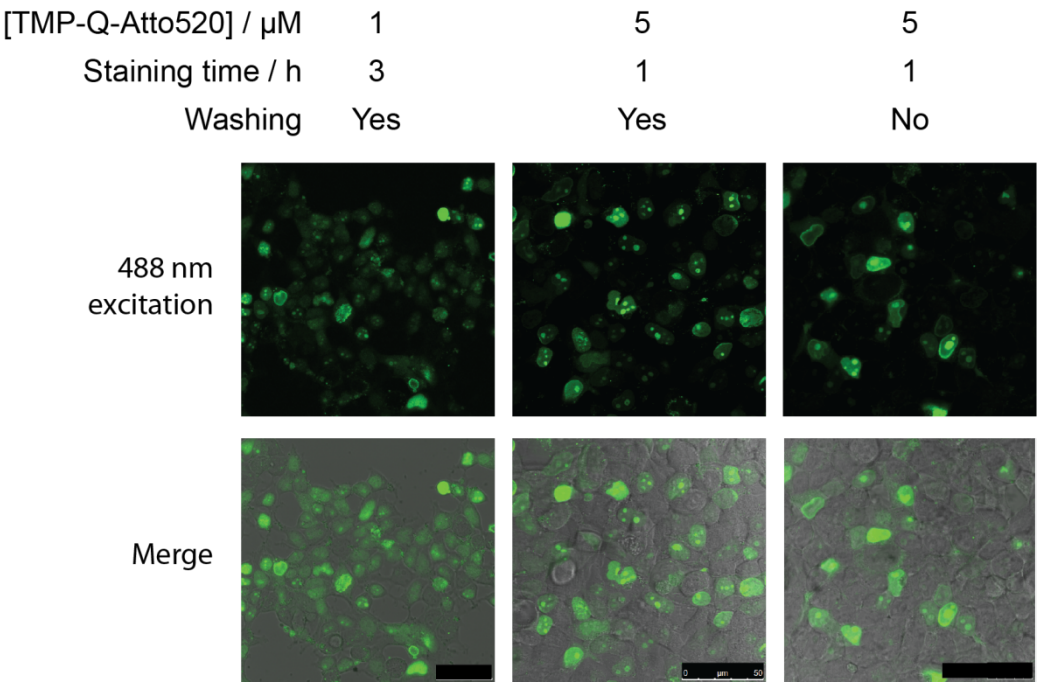


Figure S12. Live cell imaging of H2B-eDHFR:L28C labeled with TMP-Q-Atto520 using different staining protocols. HEK 293T cells were transiently transfected with plasmid encoding H2B-eDHFR:L28C and subsequently labeled with TMP-Q-Atto520 with different protocol as indicated above each column.

Labeling Kinetics in Live Cells

Figure S13 shows the result of gel-shift assay in live HEK 293T cells using western blot.

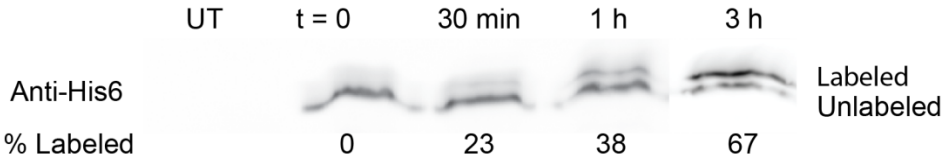


Figure S13. Gel-shift of eDHFR:L28C in live HEK 293T cells labeled with 10 µM TMP-Q-Atto520. HEK 293T cells were transiently transfected with plasmids encoding eDHFR:L28C-His_6 for mammalian expression. Transfected cells were labeled with 10 µM TMP-Q-Atto520 and lysed at different time points as indicated in the figure. Cell lysates were analyzed with SDS-PAGE gel and eDHFR:L28C was detected by western blot using anti-His_6 monoclonal antibody conjugated with horseradish peroxidase for chemiluminescence imaging. Because of change in molecular weight, covalent labeling on eDHFR:L28C by TMP-Q-Atto520 causes upward gel shift so that the labeling efficiency can be estimated by densitometric measurement.

Cytotoxicity Evaluation of TMP-Q-Atto520

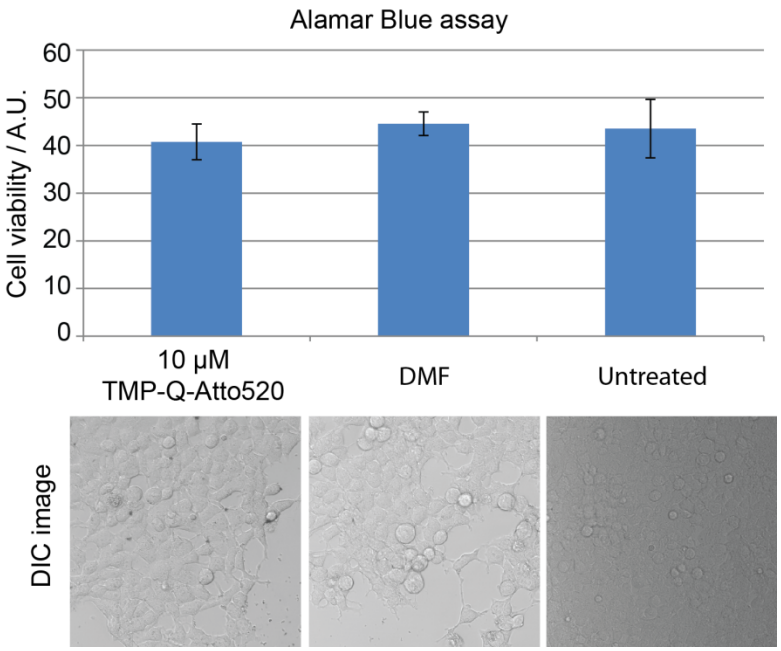


Figure S14. Evaluation of cytotoxicity of TMP-Q-Atto520 using Alamar Blue assay and DIC imaging, respectively. HEK 293T cells were incubated for 3 h at 37 °C with 10 μM TMP-Q-Atto520 and DMF, respectively, with the untreated cells as control. Then the cells were treated with 10% Alamar Blue for 4 h at 37 °C and fluorescence intensity (Ex: 580 nm; Em: 620 nm) was measured to indicate cell viability. Error bar represents standard deviation from three trials. With Alamar Blue assay, 10 μM TMP-Q-Atto520 was not found to significantly impact the cell viability of HEK 293T cells, in comparison to cells treated with DMF and untreated cells. Under DIC microscope the HEK 293T cells treated with 10 μM TMP-Q-Atto520 also showed normal morphology and viability compared with DMF treated and untreated cells.

References

1. Rasband, W. S. (1997-2007) ImageJ, U. S. National Institutes of Health: Bethesda.
2. Wombacher, R., Heidbreder, M., van de Linde, S., Sheetz, M. P., Heilemann, M., Cornish, V. W., and Sauer, M. (2010) Live-cell super-resolution imaging with trimethoprim conjugates, *Nat Methods 7*, 717-719.
3. Chen, Z. X., Jing, C. R., Gallagher, S. S., Sheetz, M. P., and Cornish, V. W. (2012) Second-Generation Covalent TMP-Tag for Live Cell Imaging, *Journal of the American Chemical Society 134*, 13692-13699.