Supporting Information

"Cell-permeable small molecule probes for site-specific labeling of proteins"

Dawn S. Y. Yeo, Rajavel Srinivasan, Mahesh Uttamchandani, Grace Y. J. Chen, Qing Zhu, and Shao Q. Yao,*

Departments of Biological Sciences & Chemistry, National University of Singapore, 3 Science drive 3, Singapore 117543, Republic of Singapore

1. General

All chemicals were purchased from commercial sources, unless indicated otherwise. All chemical reactions were run under N_2, unless otherwise indicated. The ^1H and ^{13}C NMR spectra were taken on a Bruker 300 MHz spectrometer. Chemical shifts are reported in parts per million referenced to internal standard ($(CH3)_4Si$ = 0.00 ppm). The mass spectra were taken on a Finnigan LCQ spectrometer.

2. Live cell labeling strategy with our cell-permeable probes.

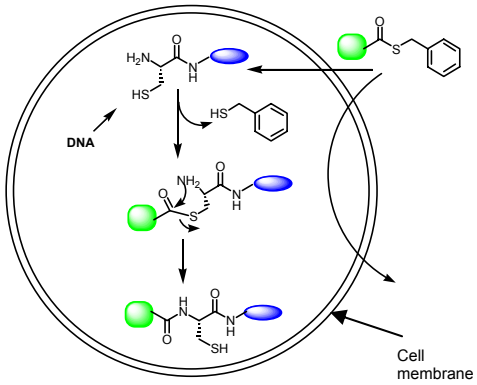


Fig 3. Chemoselective reaction between a thioester-containing probe and an N-terminal Cys protein in a living cell.

3. Chemical Synthesis

As shown in Scheme 2, **1** was readily synthesized from CM **9**, which was prepared as previously reported.[1] Acetylation of **9** gave **10** in 30% yield, which was then converted to **1** (56% yield) by coupling to thioester **8** under mild conditions.[2] The other three dyes, FL **12**, TMR **15** and CF **17**, are all commercially available but were conveniently synthesized in house using optimized procedures. **12** was prepared from resorcinol **11** and trimellitic anhydride in the presence of methanesulfonic acid, following published protocols.[3] Following acetylation, the resulting product, **13** upon purification,

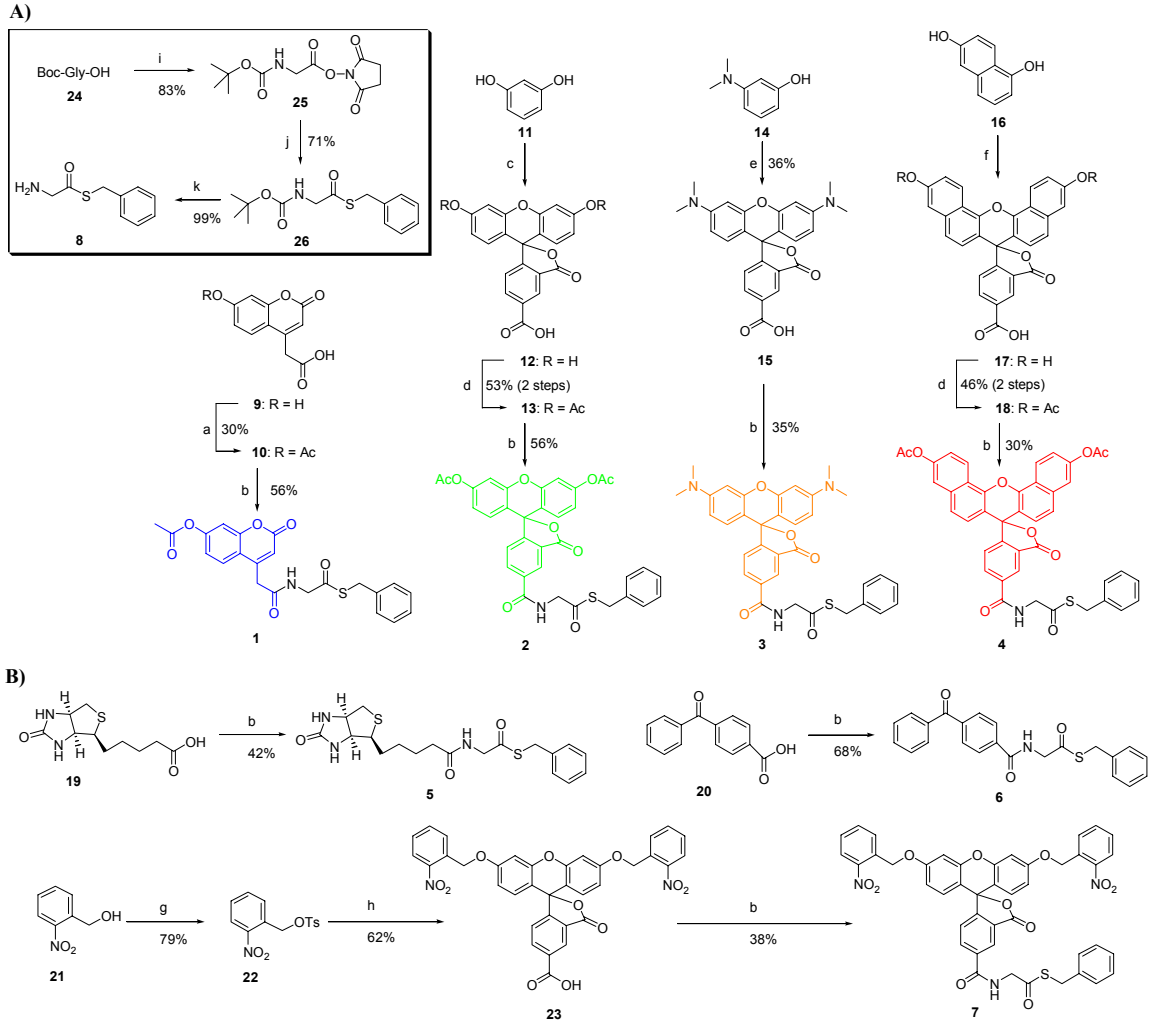
---

([1]) (a) Zhu, Q.; Uttamchandani, M.; Li, D.B.; Lesaicherre, M. L.; Yao, S. Q. *Org. Lett.* **2003**, *5*, 1257-1260. (b) Zhu, Q.; Li, D. B.; Uttamchandani, M.; Yao, S.Q. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 1033-1036.

([2]) Neutralizing bases (i.e. DIEA) commonly used in the coupling reaction should be avoided, as the diacetates in **1**, **2** & **4** are extremely base-labile. For the same reason, base extraction (e.g. $NaHCO_3$) in the workup should also be avoided.

([3]) Sun, W. C.; Gee, K. R.; Klaubert, D. H.; Haugland, R. P. *J. Org. Chem.* **1997**, *62*, 6469-6475.

was coupled with **8** to give probe **2** in 56% yield (in 2 steps).[11] **15** and **17** were similarly prepared. We found that, by adjusting the reaction conditions (e.g. temperature, time, solvent, CH_3SO_3H) used in the **FL** synthesis, both **TMR** and **CF** dyes could also be conveniently obtained in high yields from their respective starting materials (i.e. **14** and **16)**. Subsequently, **15**, and the corresponding acetylated product of the **CF** dye, **18**, were similarly coupled with **8** to give probes **3** and **4**, respectively.[11] Probes **5** and **6** were prepared in one step by coupling commercially available starting materials, **19** and **20**, with the thioester **8**, respectively. The "caged" probe **7** was generated similarly from **23**, which was readily prepared from **12** by masking the two phenolic alcohols with *o*-nitrobenzyl alcohol **21**.



Scheme 2. Reagents and Conditions: (a) Ac_2O, pyridine, 30 min, RT; (b) **8**, EDC, HOBt, THF/DMF, RT; (c) trimellitic anhydride, CH_3SO_3H, 80 °C, 12 h; (d) Ac_2O, pyridine, 15 min, 85 °C; (e) trimellitic anhydride, toluene, 110 °C, 12 h; (f) trimellitic anhydride, CH_3SO_3H, 100 °C, 12 h; (g) *p*-toluenesulfonyl chloride, DCM, RT, 2 h; (h) **12**, K_2CO_3, DMF, 60 °C, 6 h; (i) N-hydroxysuccinimide, DCC, THF, RT, 1 h; (j) benzyl mercaptan, THF, RT, 12 h; (k) 4 N HCl, dioxane, RT, 2 h.

tert-Butoxycarbonylamino-acetic acid 2,5-dioxo-pyrrolidin-1-yl ester (**25**):

Commercially available Boc-glycine-OH, **24** (7.0 g, 40 mmol) was dissolved in 60 mL of dry THF at 0 ºC. To this solution was added DCC (4.6 g, 40 mmol) and the mixture was stirred at RT for 1 hour. Upon filtration to remove the precipitate, the

resulting product was concentrated under reduced pressure. Pure crystals of **25** (9.03 g, 83%) was obtained by recrystallization with hot isopropanol. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.2 (broad, 1H), 4.28 (d, *J* = 5.22 Hz, 2H), 2.84 (s, 4H), 1.46 (s, 9H). MS (ESI): m/z 294.9 $[(M+Na)]^+$.

tert-Butoxycarbonlyamino-thioacetic acid S-benzyl ester (**26**):

25 (10 g, 36.73 mmol) was dissolved in dry THF. DIEA (6.8 ml, 38.7 mmol) was added followed by benzyl mercaptan (4.67 ml, 38.7 mmol). The resulting mixture was stirred at RT for overnight. Upon evaporation to dryness, the product was redissolved in ethyl acetate, washed with 2 x 1 N HCl, 2 x Sat. $NaHCO_3$, water and brine. The Organic layer was dried with anhydrous $MgSO_4$. The resulting white solid was dissolved in a minimal amount of DCM and recrystallized with hexane to afford **26** (7.3 g, 71%)**.** ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.29-7.26 (m, 5H), 4.14 (s, 2H), 4.06 (d, *J* = 5.61 Hz, 2H), 1.45 (s, 9H). MS (ESI): m/z 304.1 $[(M+Na)]^+$.

Amino thioacetic acid S-benzyl ester (**8**):

26 (7.3 g, 26mmol) was dissolved in 4 N HCl (20 ml) and dioxane (20 ml), followed by stirring at RT for 2 hours. The solution was then concentrated *in vacuo* and chased with ether (3 x 50 ml) to afford pure white crystals of **8** (4.67 g, 99%). ^{1}H-NMR (300 MHz, DMSO) δ 8.5 (broad, 3H), 7.29-7.33 (m, 5H), 4.25 (s, 2H), 4.08 (s, 2H). ^{13}C-NMR (60 MHz, DMSO) δ 192.70, 136.89, 128.79, 128.50, 127.32, 46.72, 32.09. MS (ESI): m/z 182.0 $[(M+1)]^+$.

7-Hydroxycoumarin-4-acetic acid (**9**):

Resorcinol (5 g, 45 mmol) was dissolved in 50 ml of 70% H_2SO_4 under ice-cold conditions, and the solution was allowed to stir for 30 minutes. Acetone carboxylic acid (6.6 g, 45 mmol) was added in 5 portions. The mixture was allowed to stir further for 4 hrs, before pouring onto crushed ice pieces. The precipitate formed was washed with water, ethyl acetate and dried overnight under reduced pressure to afford a pure white solid (9 g, 91%). ^{1}H NMR (300 MHz, DMSO) δ 7.53 (d, *J* = 9.1 Hz, 1H), 6.80 (dd, *J* = 8.7 & 2.1 Hz, 1H), 6.73 (d, *J* = 2.1 Hz, 1H), 6.22 (s, 1H), 3.82 (s, 2H). ^{13}C-NMR (60 MHz, DMSO) δ 170.56, 161.10, 158.16, 154.94, 150.05, 127.65, 126.58, 112.93, 111.90, 102.23, 39.30. MS (ESI): m/z 221.0 $[(M+1)]^+$.

7-Acetoxycoumarin-4-acetic acid (**10**):

To **9** (1 g, 4.53 mmol) was added 25 ml of acetic anhydride and 5 ml of pyridine, and the reaction was allowed to stir at RT for 30 minutes. Upon removal of solvent *in vacuo*, the residue was taken into ethyl acetate and washed with 1 N HCl (2 x 30 ml), water (2 x 30 ml) and brine (2 x 30 ml). The organic layer was dried over anhydrous Na_2SO_4 and concentrated *in vacuo* to afford **10** (30% yield). ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.71 (d, *J* = 8.9 Hz, 1H), 7.23 (d, *J* = 5.6 Hz, 1H), 7.24(s, 1H), 6.47(s, 1H), 2.94 (s, 2H), 2.33 (s, 3H). MS (ESI): m/z 263.0 $[(M+H)]^+$.

Acetic acid 4-[9-benzyl sulfanylcarbonylmethyl-carbamoyl)-methyl]-2-oxo-2H-chromen-7-yl ester (**1**):

To a solution of **10** (0.162 g, 0.62 mmol) in dry THF was added EDC (0.13 g, 0.68 mmol) and HOBt (0.104 g, 0.68 mmol). The reaction was stirred for 30 minutes under ice-cold conditions, followed by addition of **8** (0.135 g, 0.62 mmol) dissolved in minimal amount of DMF. The resulting solution was stirred further for 6 hours at RT. Upon removal of the solvent *in vacuo*, the resulting mixture was taken into ethyl acetate and washed with 1 N HCl (2 x 30 ml), water (2 x 30 ml) and brine (2 x 30 ml). The organic layer was dried over anhydrous Na_2SO_4 and concentrated *in vacuo* to afford a white solid, which was further purified by flash chromatography (silica gel, dichloromethane/ethanol = 9:1) to afford **1** (0.15 g, 56%). ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.67 (d, *J* = 9.0 Hz, 1H), 7.27-7.23 (m, 7H), 6.38 (s, 1H), 4.1 (s, 2H), 3.73 (s, 2H), 2.91 (s, 2H), 2.32 (s, 3H). MS (ESI): m/z 448.0 $[(M+Na)]^+$.

4-(3,6-diacetoxy-3H-xanthen-9-yl)-isophthalic acid (**13**):

Resorcinol **11** (5.7 g, 52 mmol) was dissolved in 50 ml of CH_3SO_3H. To this solution, trimellitic anhydride (5 g, 26 mmol) was added, and the reaction was heated at 80-85 ºC for 12 hours. The highly viscous solution was cooled to RT then poured into 10 volumes of ice-cold water. The resulting precipitate was collected and dried *in vacuo* to give compound **12** as a crude yellow solid which was used without further purifications. Compound **12** was dissolved in 50 ml of acetic anhydride and 10 ml of pyridine. The solution was heated at 85 ºC for 15 minutes. The resulting solution was poured into ice-cold water to give the crude precipitate of **13**, which, upon recystallization with DCM, afforded pure **13** as white crystals (5.17 g, 53%). ^{1}H-NMR (300 MHz, $CDCl_3$) δ 8.75 (s, 1H), 8.38 (d, *J* = 8.04 Hz, 1H), 7.29 (d, *J* = 8.04, 1H), 7.15-7.09 (m, 2H), 6.85-6.81 (m, 4H), 2.31 (s, 6H). ^{13}C-NMR (60 MHz, $CDCl_3$) δ 168.71, 167.46, 165.85, 155.48, 152.12, 150.69, 133.24, 132.47, 130.31, 129.23, 128.45, 126.14, 115.48, 110.41, 109.30, 81.21, 19.85. MS (ESI): m/z 461 $[(M+1)]^+$.

N-Benzylsulfanylcarbonylmethyl-6-(3,6-dihydroxy-3H-xanthen-9-yl)-isophthalamic acid (FL) (**2**):

To a solution of **13** (0.424 g, 0.92 mmol) in dry THF was added EDC (0.176 g, 0.92 mmol) and HOBt (0.14 g, 0.92 mmol). The solution was stirred for 30 minutes under ice-cold conditions, followed by addition of **8** (0.2523 g, 1.16 mmol) dissolved in minimal amount of DMF. The resulting solution was stirred further for 6 hours at RT. Upon removal of the solvent *in vacuo*, the residue was taken into ethyl acetate and washed with 1 N HCl (2 x 30 ml), water (2 x 30 ml) and brine (2 x 30 ml). The organic layer was dried over anhydrous Na_2SO_4 and concentrated under reduced pressure to afford a crude mixture, which was further purified by flash chromatography (silica gel, dichloromethane/methanol = 9:1) to afford **2** (0.32 g, 56%). ^{1}H-NMR (300 MHz, $CDCl_3$) δ 8.52 (s, 1H), 8.22 (d, *J* = 8.04 Hz, 1H), 7.25-7.21 (m, 6H), 7.12-7.09 (m, 2H), 6.81-6.75 (m, 4H), 4.39 (d, *J* = 5.61 Hz, 2H), 4.15 (s, 2H), 2.3 (s, 6H). ^{13}C-NMR (60 MHz, DMSO) δ 196.52, 168.78, 168.37, 165.84, 155.31, 152.20, 151.42, 136.72, 134.08, 129.58, 127.74, 127.50, 127.33, 126.45,125.61, 124.50, 123.91, 115.53, 110.46, 82.09, 49.48, 32.90, 20.98. MS (ESI): m/z 623.9 $[(M+1)]^+$.

5-Carboxy-tetramethylrhodamine (**15**):

Trimellitic anhydride (1.00 g, 5.2 mmol) and 3-dimethylaminophenol, **14** (0.72 g, 10.5 mmol) was refluxed in toluene (50 ml) for 12 hours. Upon cooling to RT, the resulting precipitate was collected by filtration, and further purified by flash chromatography (silica gel, dichloromethane/methanol/acetic acid = 8:1.9:0.1) to afford pure **15** as a dark purple solid (0.811 g, 36%). ^{1}H-NMR (300 MHz, $CDCl_3$) δ 8.52 (s, 1H), 8.12 (d, *J* = 8.61 Hz, 1H), 7.59 (d, *J* = 7.62 Hz, 1H), 6.82 (m, 2H), 6.51 (d, *J* = 5.61 Hz, 2H). 6.09 (s, 2H), 3.28 (s, 12H). MS (ESI): m/z 431.2 $[(M+1)]^+$.

Amino thioacetic acid S-benzyl ester of 5-carboxy-tetramethylrhodamine (TMR) (**3**):

To a solution of **15** (0.05 g, 0.116 mmol) in dry THF was added EDC (0.022 g, 0.116 mmol) and HOBt (0.018 g, 0.116mmol). The solution was stirred for 30 minutes under ice-cold conditions, followed by addition of **8** (0.025 g, 0.116 mmol) dissolved in minimal amount of DMF. The resulting solution was stirred further for 6 hours at RT. Upon removal of the solvent *in vacuo*, the residue was taken into ethyl acetate and washed with 1 N HCl (2 x 30 ml), water (2 x 30 ml) and brine (2 x 30 ml). The organic layer was dried over anhydrous Na_2SO_4 and concentrated under reduced pressure to afford a crude mixture, which was further purified by flash chromatography (silica gel, dichloromethane/ethanol = 9:1) to afford **3** (0.024 g, 35%) as a dark purple solid. ^{1}H-NMR (300 MHz, CD_3OD) δ 8.78 (s, 1H), 8.27 (d, *J* = 8.01 Hz, 1H), 7.55 (d, *J* = 7.83 Hz, 1H), 7.28 (d, *J* = 6.45, 2H), 7.20-7.26 (m, 5H), 7.09-6.99 (m, 4H), 4.40 (s, 2H), 4.15 (s, 2H), 3.30 (s, 12H). MS (ESI): m/z 594.2 $[(M+1)]^+$.

Acetoxynapthofluorescein (**18**):

Trimellitic anhydride (1.93 g, 10 mmol) was dissolved in CH_3SO_3H (1 M). **16** (3.2 g, 20 mmol) was added and the solution was heated at 100 ^{o}C for 12 hrs. Upon cooling to RT, the solution was poured into 8 volumes of ice-cold water. The resulting red precipitate was collected and dried *in vacuo* to give crude **17**, which was used directly without further purifications. MS (ESI): m/z 477.4 $[(M+1)]^+$. Crude **17** was dissolved in 50 ml of acetic anhydride and 10 ml of pyridine. The resulting mixture was heated at 80-90 ^{o}C for 15 minutes**.** Upon concentration to dryness, the residue was taken into ethyl acetate and washed with 1 N HCl (2 x 50 ml), water (2 x 50 ml) and brine (2 x 50 ml). The organic layer was dried over anhydrous Na_2SO_4, and concentrated *in vacuo* to give crude **18** (46% yield for 2 steps based on the crude product), which was confirmed by MS ((ESI): m/z 561.0 $[(M+1)]^+$) and used without further purifications.

Amino thioacetic acid S-benzyl ester of acetoxynapthofluorescein (CF) (**4**):

To a THF solution of crude **18** (0.065 g, 0.116 mmol), obtained from above reaction, was added EDC (0.022 g, 0.116 mmol) and HOBt (0.018 g, 0.116 mmol). The solution was stirred for 30 minutes under ice-cold conditions, followed by addition of **8** (0.025 g, 0.116 mmol) dissolved in minimal amount of DMF. The resulting solution was stirred further for 6 hours at RT. Upon removal of the solvent *in vacuo*, the residue was taken into ethyl acetate and washed with 1 N HCl (2 x 30 ml), water (2 x 30 ml) and brine (2 x 30 ml). The organic layer was dried over anhydrous Na_2SO_4 and concentrated under reduced pressure to afford a crude mixture, which was further purified by reverse phase HPLC to afford pure **4** (0.025 g, 30% yield). ^{1}H-NMR (300 MHz, $CDCl_3$) δ 8.62

(s, 1H), 8.22(d, *J* = 6.42 Hz, 1H), 8.14-8.11 (m, 2H), 7.64 (s, 2H), 7.48 (d, *J* = 8.03 Hz, 1H), 7.31-7.27 (m, 9H), 7.19 (d, *J* = 7.62 Hz, 2H), 4.2 (d, *J* = 6.0 Hz, 2H), 4.01 (s, 2H), 2.37 (s, 6H). MS (ESI): m/z 724.1 $[(M+1)]^+$.

[5-(2-Oxo-hexahydro-thieno[3,4-d]imidazol-6-yl)-pentanoylamino]-thioacetic acid S-benzyl ester (**5**):

Biotin **19** (0.156 g, 0.64 mmol) was dissolved in 40 mL of dry THF. To this solution was added EDC (0.123 g, 0.64 mmol) and HOBt (0.098 g, 0.64 mmol) under ice-cold conditions. The reaction was allowed to stir for 30 minutes, followed by addition of **8** (0.139 g, 0.64 mmol) dissolved in minimal amount of DMF. The resulting solution was allowed to stir overnight at RT. Upon removal of the solvent *in vacuo*, the residue was taken into ethyl acetate and washed with 1 N HCl (2 x 30 ml), water (2 x 30 ml) and brine (2 x 30 ml). The organic layer was dried over anhydrous Na_2SO_4 and concentrated under reduced pressure to afford **5** (0.110 g, 42%). ^{1}H-NMR (300 MHz, DMSO) δ 7.31-7.28 (m, 5H), 6.40-6.35 (broad, 2H), 4.37-4.32 (m, 2H), 4.22 (d, *J* = 5.61 Hz, 2H), 4.09 (s, 2H), 3.1-3.0 (m, 1H), 2.88-2.78 (m, 2H), 2.19-2.14 (m, 2H), 1.60-1.48 (m, 6H). MS (ESI): m/z 408.0 $[(M+1)]^+$.

(4-Benzoyl-benzoylamino)-thioacetic acid S-benzyl ester (**6**):

To a solution of **20** (0.136 g, 0.6 mmol) in dry THF (40 ml) was added EDC (0.115 g, 0.6 mmol) and HOBt (0.092 g, 0.6 mmol). The solution was stirred for 30 minutes under ice-cold conditions, followed by addition of **8** (0.136 g, 0.6 mmol) dissolved in minimal amount of DMF. The resulting solution was stirred further for 6 hours at RT. Upon removal of the solvent *in vacuo*, the residue was taken into ethyl acetate and washed with 1 N HCl (2 x 30 ml), water (2 x 30 ml) and brine (2 x 30 m). The organic layer was dried over anhydrous Na_2SO_4 and concentrated under reduced pressure to afford a white solid, which was further purified by flash chromatography (silica gel, dichloromethane/ethanol = 9.5:0.5) to afford **6** (0.157 g, 68%). ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.94-7.91 (m, 2H), 7.76-7.73 (m, 4H), 7.61-7.56 (m, 1H), 7.48-7.43 (m, 2H), 7.27-7.20 (m, 5H), 4.37 (d, *J* = 5.64 Hz, 2H), 4.12 (s, 2H). ^{13}C-NMR (60 MHz, $CDCl_3$) 195.88, 194.66, 166.68, 140.39, 138.69, 136.66, 134.20, 132.89, 131.82, 131.17, 130.05, 129.99, 128.99, 128.80, 128.23, 128.16, 127.45, 127.38, 127.10, 49.39, 33.00. MS (ESI): m/z 389.47 $[(M+1)]^+$.

2-Nitrobenzyl tosylate (**22**):

2-Nitrobenzyl alcohol (1 g, 6.5 mmol) was dissolved in 40 ml of dry DCM and 1.36 ml of triethylamine. *p*-Toluenesulfonyl chloride (1.24 g, 6.5 mmol) was added and the solution was stirred for 2 hours before quenching with water. The organic layer was separated and washed with 1 N HCl(1 x 50 ml), water (1 x 50 ml), 2 N NaOH (1 x 50 ml), water (1 x 50 ml) and brine (1 x 50 ml). The organic layer was dried over anhydrous Na_2SO_4 and concentrated under reduced pressure to give a dull white solid, which was further purified by flash chromatography (silica gel, dichloromethane/ethanol = 9.5:0.5) to give 2-nitrobenzyl tosylate, **22** (1.57 g, 79%). ^{1}H-NMR (300 MHz, $CDCl_3$) δ 8.08-7.99 (dd, *J* = 9.21 & 8.04 Hz, 1H), 7.81 (d, *J* = 8.4 Hz, 1H), 7.72-7.62 (m, 3H), 7.50 (d, *J* = 8.85 Hz, 1H) 7.33 (d, *J* = 8.43 Hz, 2H), 4.93 (s, 2H), 2.42 (s, 3H). ^{13}C-NMR (60 MHz,

$CDCl_3$) δ 149.33, 146.12, 137.71, 135.27, 134.11, 133.09, 132.69, 132.57, 128.12, 123.89, 42.72, 20.64. MS (ESI): m/z 329.9 $[(M+Na)]^+$.

"Caged" Fluorescein (C2FL) (**23**):

Crude **12** was first purified by recrystallization. The recrystallized product (0.602 g, 1.6 mmol) was dissolved in 30 ml of DMF containing K_2CO_3 (1.1 g, 5 mmol). 2-nitrobenzyl tosylate (0.492 g, 1.6 mmol) was added to the solution and the reaction was heated at 60 ºC for 6 hours in a RBF covered with Aluminium foil to maintain dark conditions. After 6 hrs, the solvent was removed *in vacuo* and the residual mixture was washed with 1 N HCl, water and brine followed by drying in anhydrous Na_2SO_4 and concentration under reduced pressure to afford **23** (0.64 g, 62%). The work up was strictly done in dark. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 8.44 (s, 1H), 8.16 (d, J = 8.11 Hz, 1H), 8.00-7.97 (m, 2H), 7.75-7.85 (m, 2H), 7.60-7.56 (m, 4H), 7.45 (d, J = 8.10 Hz, 1H), 7.35 (d, J = 7.86 Hz, 2H), 7.22-7.10 (m, 4H), 5.51 (s, 4H). MS (ESI): m/z 647.0 $[(M+1)]^+$.

Amino thioacetic acid S-benzyl ester of C2FL (**7**):

The reaction and workups were carried out in dark. To a solution of **23** (0.084 g, 0.13 mmol) in dry THF was added EDC (0.027 g, 0.143 mmol) and HOBt (0.022 g, 0.143 mmol). The solution was stirred for 30 minutes under ice-cold conditions, followed by addition of **8** (0.028 g, 0.13 mmol) dissolved in minimal amount of DMF. The resulting solution was stirred further for 6 hours at RT. Upon removal of the solvent *in vacuo*, the residue was taken into ethyl acetate and washed with 1 N HCl (2 x 30 mm), water (2 x 30 ml) and brine (2 x 30 ml). The organic layer was dried over anhydrous Na_2SO_4 and concentrated under reduced pressure to afford a crude mixture, which was further purified by flash chromatography (silica gel, dichloromethane/ethanol = 9.8:0.2) to afford **7** (0.040 g, 38%). ^{1}H-NMR (300 MHz, $CDCl_3$) δ 8.38 (s, 1H), 8.29 (d, J = 8.13 Hz, 1H), 8.08-8.11 (m, 2H), 7.99-7.90 (m, 2H), 7.71-7.68 (m, 4H), 7.65 (d, J = 8.10 Hz, 1H), 7.60-7.49 (m, 6H), 7.25 (s, 5H), 5.29 (s, 4H), 4.68 (s, 2H), 4.19 (s, 2H). MS (ESI): m/z 810.1 $[(M+1)]^+$.

4. Fluorescence experiments

10 µl each of the stock solutions of probes **1**, **2**, **3** and **4** (1 mM in DMSO) was added to 100 µl of 1 mM K_2CO_3 solution, and the mixture was allowed to stand at RT for > 5 min for the deacylation reaction to occur, resulting in the release of very strong fluorescence in the solution. 10 µl of the solution was subsequently diluted with 100 µl of distilled water, and the resulting solution was transferred to a black 96-well microtitre plate (Nunc, USA), where the excitation and the emission spectra of the probes were recorded using a SpectraMax™ GeminiXS fluorescence microplate reader (Molecular Devices, USA).

5. *In vitro* Protein labeling procedures

With N-terminal cysteine proteins

Cloning and expression of N-terminal cysteine proteins will be reported elsewhere. Purified enhanced green fluorescent protein (EGFP) having an N-terminal cysteine was incubated for up to 24 hours with probes **2, 3, 4** & **5**, respectively, in order to assess our labeling strategy. Probes were prepared as 200 μM stocks (25 x in DMSO) and stored at -20 ^{o}C. In a typical labeling reaction, 6 μl of each probe (final concentration: 8 μM) was added to 50 μl of pure protein (1 mg/ml) dissolved in 1 x PBS (final concentration of protein: ~2 nM), with or without 1 mM DTT, and the reaction was topped up with 1 x PBS to a final volume of 150 μl. DTT was shown in our experiments, where live cells were used, to reduce background labelings. At specific time intervals, 15 μl of the reaction was withdrawn and quenched by addition of 1.7 μl of 100 mM cysteine to the reaction mixture (final concentration of cysteine: 10 mM), followed by denaturation with SDS-PAGE loading dye at 95 ^{0}C for 3 min. The loading dye also serves to hydrolyze the diacetates groups on probes **2** and **4**, thereby releasing the fluorescence. Upon separation with a 12% SDS-PAGE gel, the fluorescence labeling of EGFP by probes **2**, **3** and **4** was conveniently visualized by scanning the resulting gel with a Typhoon™ 9200 fluorescence scanner (Amersham Biosciences, USA). Fluorescence intensity of each protein band was analysed using the software, Image Quant 5.2, preinstalled on the instrument. In all cases, more than 75% of labeling was obtained in 3 hours for all three probes. For the labeling of EGFP with probe **5**, an anti-biotin Western blot was performed to visualize the amount of biotinylated EGFP. Briefly, following SDS-PAGE separation, the resulting gel was electroblotted onto a polyvinylidene difluoride (PVDF) membrane (BioRad, USA) and blocked for 1 h with 5% non-fat dry milk in PBST (phosphate buffered saline, pH 7.4 with 0.1% Tween 20). The membrane was incubated with anti biotin-conjugated HRP (NEB, USA) in an 1:3000 dilution in milk-PBST for 1 h and then washed with PBST (3 x 15 min). Visualization was done with the Enhanced ChemiLuminescent ECL™ kit (Amersham).

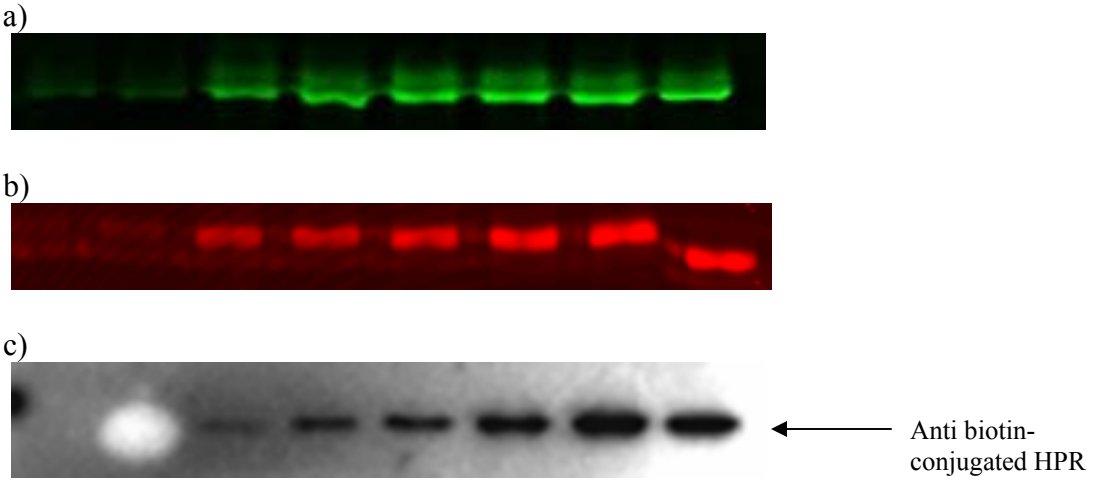


Fig. 4. SDS-PAGE of purified N-terminal cysteine EGFP labeled with a) probe **2** and b) probe **3** followed by visualization with fluorescence gel scanner; and c) probe **5** followed by Western blot with anti biotin-

conjugated HRP. For probe **4** labeling, see **Figure 1a** in main text. Lanes 1 to 8 (left to right), time of labeling reaction – 1 min, 10 min, 30 min, 1 h, 3 h, 6 h, 12 h, 24 h. See **Fig. 2** in main text for summary.

With proteins that do not contain an N-terminal cysteine (negative control)

The site-specific nature of the labeling reaction was confirmed by repeating the experiment under identical conditions with 5 control proteins which either do not have cysteine residues at all, or have only internal cysteines (e.g. papain). In all cases, no labeling was observed, thereby unambiguously supporting our design principle, in which exclusive labeling should only occur at the N-terminal cysteine of the target protein.

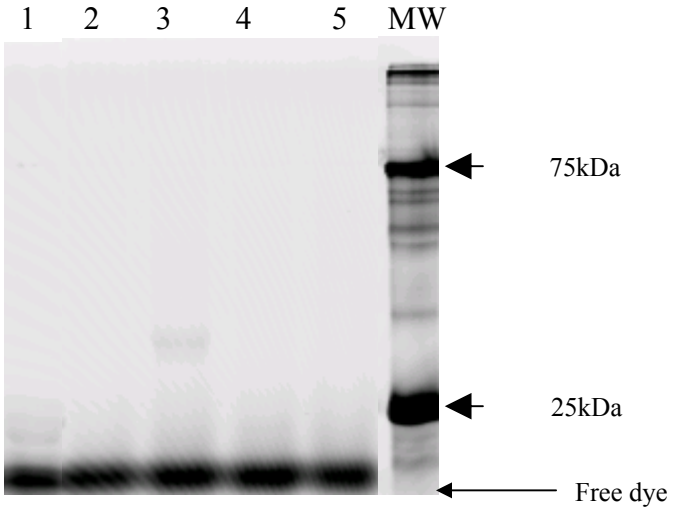


Fig 5. Control labeling reactions of 5 commercially available proteins not containing an N-terminal cysteine. Probe **3** was used in the labeling reaction, and the reaction was performed as mentioned above with 60 min incubation time. All proteins are available from Sigma (St Louis, USA). Lane #: 1. GST (cat #: G-4385); 2. Protease (P-5380); 3. Lipase (L-9031); 4. Papain (P-4762); 5. Pepsin (P-6887); 6. BioRad Precision Plus Protein Standards

6. Protein labeling inside live bacterial cells.

A DNA construct containing N-terminal Cysteine GST was transformed into the *E.coli* expression strain ER2566 (NEB) and grown in 100mg/L ampicillin containing LB media at 37°C in a 250rpm air shaker. At OD_{600} ~ 0.6, protein expression was induced with 0.3mM IPTG (isopropyl-β-D-thiogalactoside). Cells were grown for 12h at room temperature to optimize protein expression as well as self-cleavage of the intein-fusion protein to generate N-terminal cysteine proteins *in vivo*. For labeling, 8μM of the TMR probe (probe **3**) was added directly to the LB media and the cells incubated for 24h. Cells were harvested by centrifugation at 4000rpm for 10min and resuspended in 1x PBS buffer (pH 7.4) containing 10% glycerol and left standing for 30min. This procedure was repeated 3 times for complete removal of any free probe. Subsequently, cells were used for either fluorescence microscope or SDS-PAGE experiments.

For fluorescence microscope experiments, cells were mounted on clean glass slides coated with 1.5% agarose. Fluorescence images were recorded with a AxioSkop 40

fluorescent microscope (Carl Zeiss, Germany) equipped with a cooled CCD camera (AxioCam, Zeiss) using a 100x oil objective. TMR fluorescence imaging was recorded using a BP546/12 excitation and LP590 emission filter (Zeiss).

For SDS-PAGE experiments, the labeled cells, upon washes as mentioned above, were harvested by centrifugation. The cell pellet was treated directly with SDS-PAGE loading dye, and analyzed by 12% SDS-PAGE gel and fluorescence scanning (Typhoon 9200, Amersham Biosciences).