Electronic Supplementary Material (ESI) for Chemical Science
This journal is © The Royal Society of Chemistry 2013

Electronic supplementary information (ESI) for:

One-step Construction of Caged Carbonic Anhydrase I Using Ligand-directed Acyl Imidazole-based Protein Labeling Method

Kazuya Matsuo,[a] Yoshiyuki Kioi,[a] Ryosuke Yasui,[a] Yousuke Takaoka,[a] Takayuki Miki,[a] Sho-hei Fujishima,[a] and Itaru Hamachi[*,a,b]

[a]*Department of Synthetic Chemistry and Biological Chemistry, Graduate School of Engineering Kyoto University, Katsura, Kyoto 615-8510, Japan*
[b]*Japan Science and Technology Agency (JST), CREST, 5 Sanbancho, Chiyoda-ku, Tokyo 102-0075, Japan*

[*]E-mail: ihamachi@sbchem.kyoto-u.ac.jp

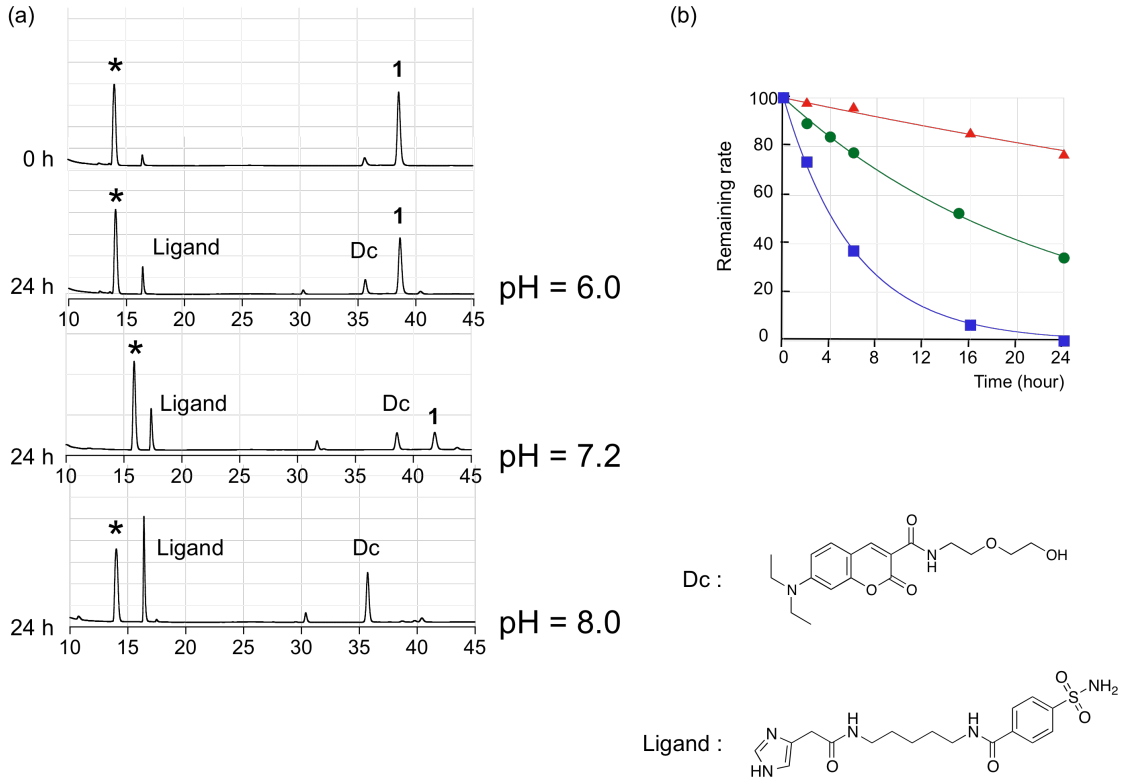


Figure S1. Evaluation of the hydrolysis stability of the LDAI reagent **1** under the different pH conditions. (a) A solution of LDAI reagent **1** (5 μM) in aqueous buffer solution (50 mM HEPES, pH 6.0, 7.2 or 8.0) was incubated at 37 °C. HPLC analysis of the hydrolysis of **1** proceeded under the different pH conditions. The peak marked with (∗) corresponds to benzenesulfonamide (10 μM) as an internal standard. The structures of "Ligand" and "Dc" generated by the hydrolysis of **1** are shown in the lower right side. (b) Time-trace plot of the remaining rate of **1** under the different pH conditions (50 mM HEPES buffer, 37 ˚C, pH = 6.0 (▲), 7.2 (●) and 8.0 (■)).

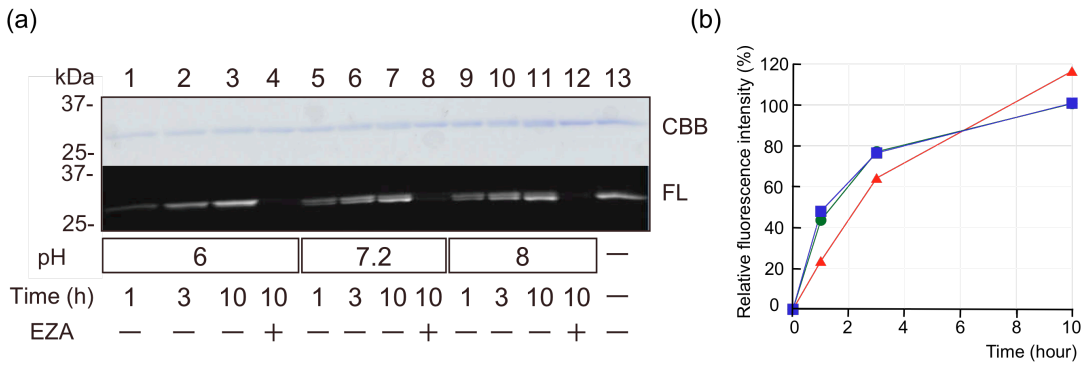


Figure S2. pH dependency of the protein labeling by the LDAI reagent **1**. (a) In-gel fluorescence analysis of the labeling reaction of CAI (5 μM) with **1** (10 μM) under the different pH conditions. EZA (100 μM) was added as a competitive inhibitor for the binding of **1** to CAI. (b) Time-trace plot of the labeling reaction of CAI with **1** under pH 6.0 (▲), pH 7.2 (●) and pH 8.0 (■). The relative fluorescence intensity (%) was obtained by the quantification of the fluorescence intensity of each band (lane 1~12 in (a)) using the band intensity in the lane 13 (loaded with the same amount of carbonic anhydrase II (CAII) that is quantitatively labeled with a Dc dye)[S1] as a standard reference.

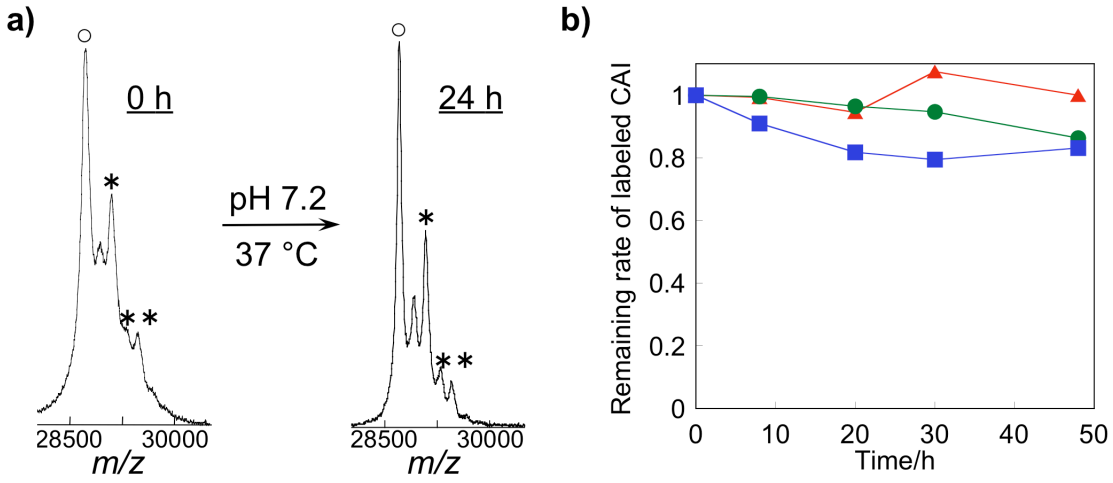


Figure S3. Evaluation of the hydrolysis stability of the labeled protein with LDAI reagent **1**. (a) MALDI-TOF mass analysis of Dc-labeled CAI incubated at pH 7.2 and 37 °C for 24 h. The peak of unlabeled CAI is marked with ○, and the peaks of CAI singly and doubly labeled by **1** are marked with ∗ and ∗∗, respectively. (b) Time-trace plot of the remaining rate of the labeled CAI (5 μM) under the three different pH conditions (50 mM HEPES buffer, 37 °C, pH = 6.0 (▲), 7.2 (●) and 8.0 (■)).

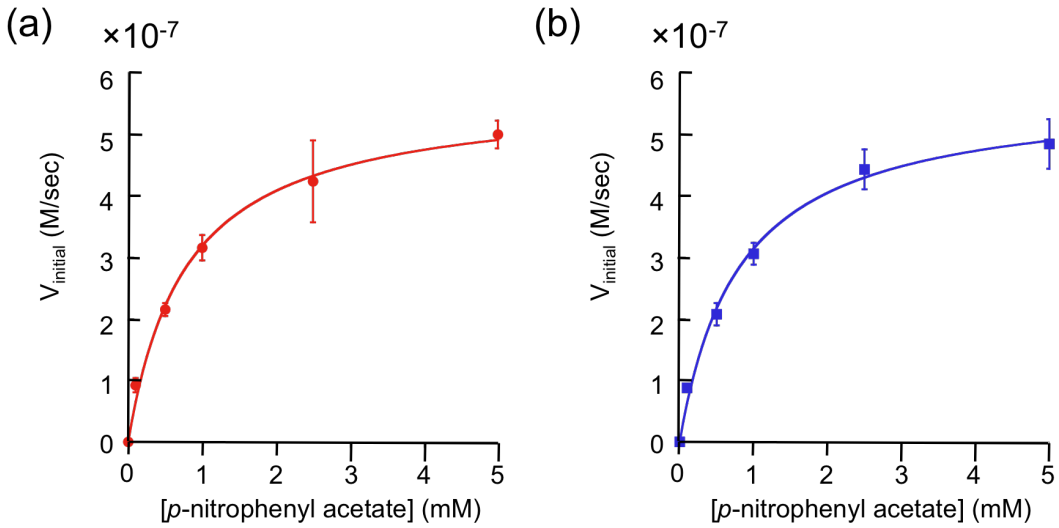
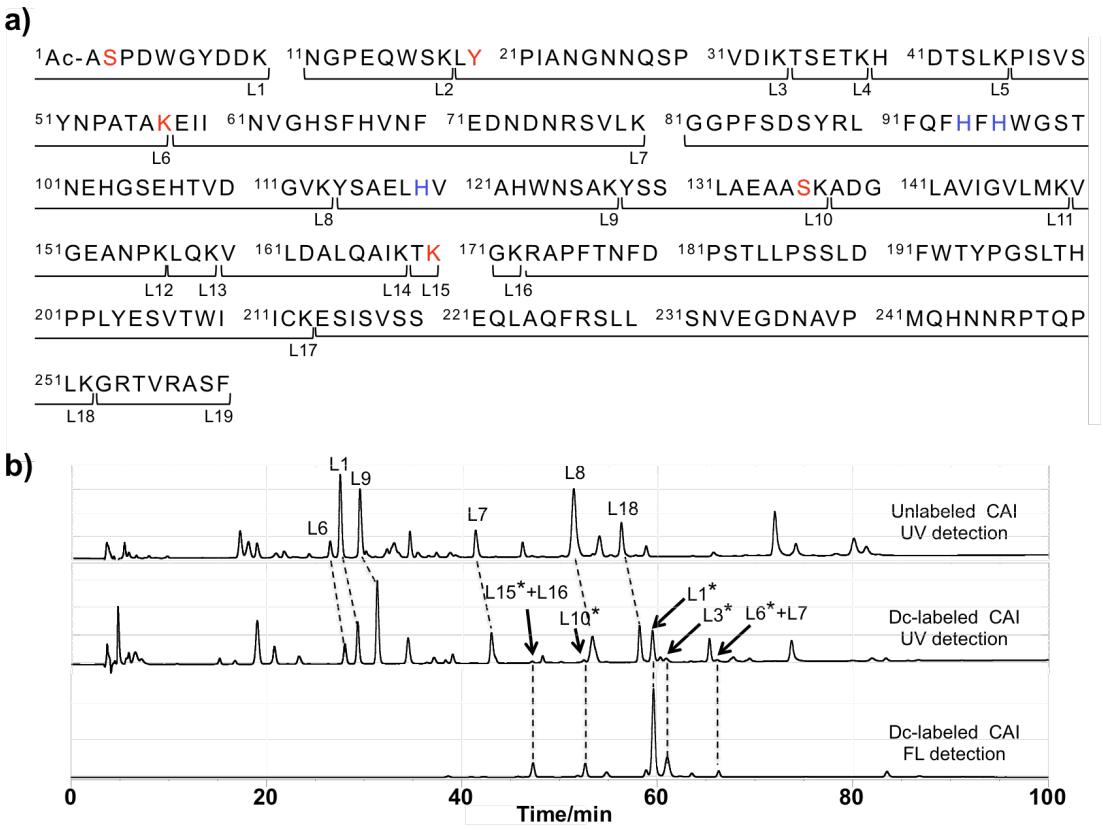


Figure S4. Evaluation of the enzyme activity of native CAI and CAI labeled with the LDAI reagent **8**. Enzyme concentration was 1.0 μM. The labeled CAI was prepared by the reaction with two equivalents of **8** in HEPES buffer (50 mM, pH 7.2) at 37 ºC for 13 h and purified by size-exclusion chromatography (TOYOPEARL HW-40F, Tosoh Corporation). The initial reaction rate was plotted as a function of the substrate concentration: ●, native CAI (a); ■, CAI labeled with **8** (b). The kinetic parameters are summarized in **Table S1**.

Table S1. Enzymatic activities of native CAI and the CAI labeled by **8** for the hydrolysis of *p*-nitrophenyl acetate.

	Native CAI	F6B-labeled CAI
k_{cat} (sec^{-1})	0.57 ± 0.02	0.57 ± 0.02
K_m (M)	$(7.8 \pm 1.0) \times 10^{-4}$	$(8.1 \pm 1.0) \times 10^{-4}$
k_{cat}/K_m ($M^{-1} \bullet sec^{-1}$)	7.3×10^{2}	7.0×10^{2}


a)
1Ac-ASPDWGYDDK 11NGPEQWSKLY 21PIANGNNQSP 31VDIKTSETKH 41DTSLKPISVS
L1 L2 L3 L4 L5
51YNPATAKEII 61NVGHSFHVNF 71EDNDNRSVLK 81GGPFSDSYRL 91FQFHFHWGST
L6 L7
101NEHGSEHTVD 111GVKYSAELHV 121AHWNSAKYSS 131LAEAASKADG 141LAVIGVLMKV
L8 L9 L10 L11
151GEANPKLQKV 161LDALQAIKTK 171GKRAPFTNFD 181PSTLLPSSLD 191FWTYPGSLTH
L12 L13 L14 L15 L16
201PPLYESVTWI 211ICKESISVSS 221EQLAQFRSLL 231SNVEGDNAVP 241MQHNNRPTQP
L17
251LKGRTVRASF
L18 L19
b)
L1 L9 L6 L7 L8 L18
Unlabeled CAI
UV detection
L15*+L16 L10* L1* L3* L6*+L7
Dc-labeled CAI
UV detection
Dc-labeled CAI
FL detection
0 20 40 60 80 100
Time/min

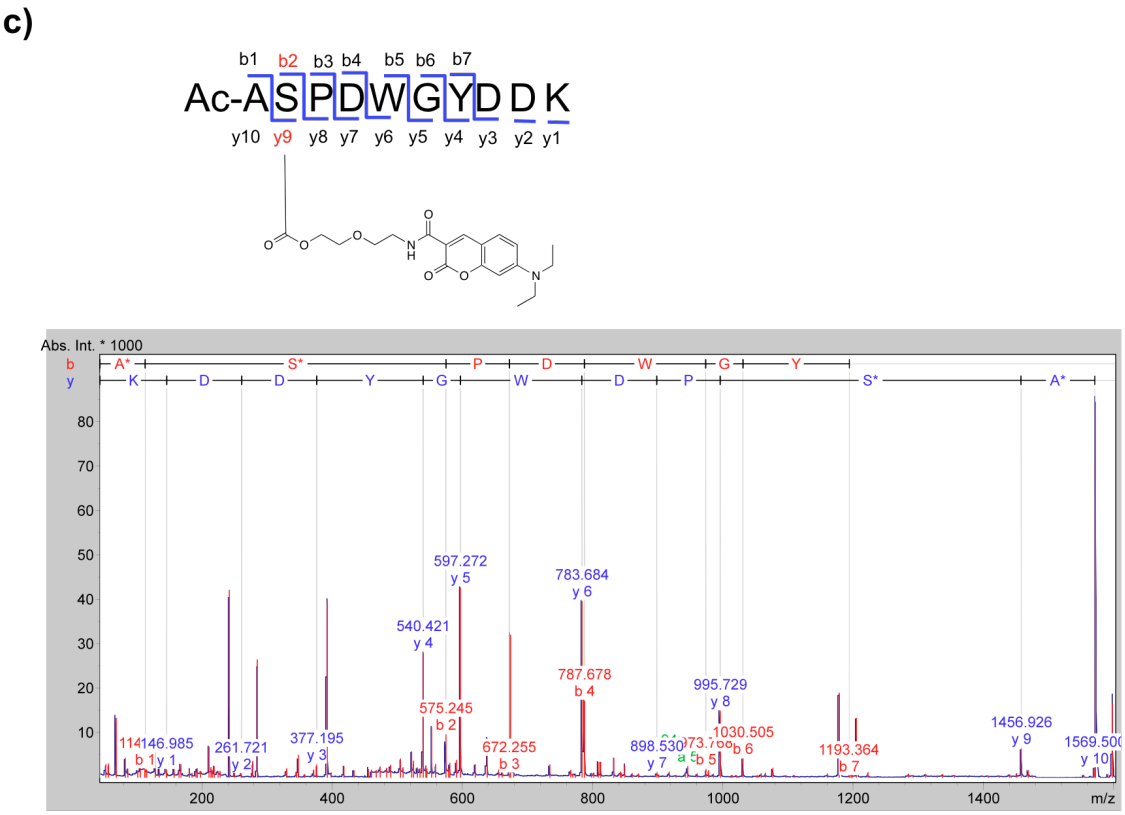

c)
b1 b2 b3 b4 b5 b6 b7
Ac-ASPDWGYDDK
y10 y9 y8 y7 y6 y5 y4 y3 y2 y1
Abs. Int. * 1000
b A* S* P D W G Y
y K D D Y G W D P S* A*
597.272 y 5
783.684 y 6
540.421 y 4
787.678 b 4
995.729 y 8
575.245 b 2
1030.505 b 6
1456.926 y 9
1114146.985 b 1 y 1
261.721 y 2
377.195 y 3
672.255 b 3
898.530 y 7
1193.364 b 7
1569.500 y 10
m/z

Figure S5. Mass spectral analysis of the labeling site of CAI by the reaction with LDAI reagent **1**. (**a**) The primary sequence of CAI and the predictable fragments generated by lysyl endopeptidase digestion. The amino acids labeled with **1** are shown in red. The histidine residues (His94, 96 and 119) coordinated to the zinc ion in the active site are shown in blue. (**b**) HPLC analysis of the digested fragments of CAI. The top figure shows the chromatogram of the fragments derived from nonlabeled CAI (UV detection at 220 nm). The middle and bottom figures show the chromatograms of the fragments derived from the labeled CAI detected with UV absorption (220 nm) and fluorescence emission (λ_{em} = 473 nm, λ_{ex} = 427 nm). The peaks marked with asterisk (∗) correspond to the labeled fragments, which were characterized by MALDI-TOF mass analysis. MALDI-TOF mass (CHCA) of peak L1*: calcd. for $[M+H]^+$ = 1569.63, obsd. 1570. peak L3*: calcd. for $[M+H]^+$ = 2117.06, obsd. 2118. peak L6*+L7: calcd. for $[M+H]^+$ = 4286.12, obsd. 4288. L10*: calcd. for $[M+H]^+$ = 1400.66, obsd. 1401. L15*+16: calcd. for $[M+H]^+$ = 807.42, obsd. 807.60. (**c**) MALDI-TOF MS/MS analysis of the dye-labeled L1*. L3*, L10*, L15*+L16 and L6*+L7 fragments were analyzed by same methods. L6*+L7 fragment was analyzed from the second digestion with the thermolysin treatment.

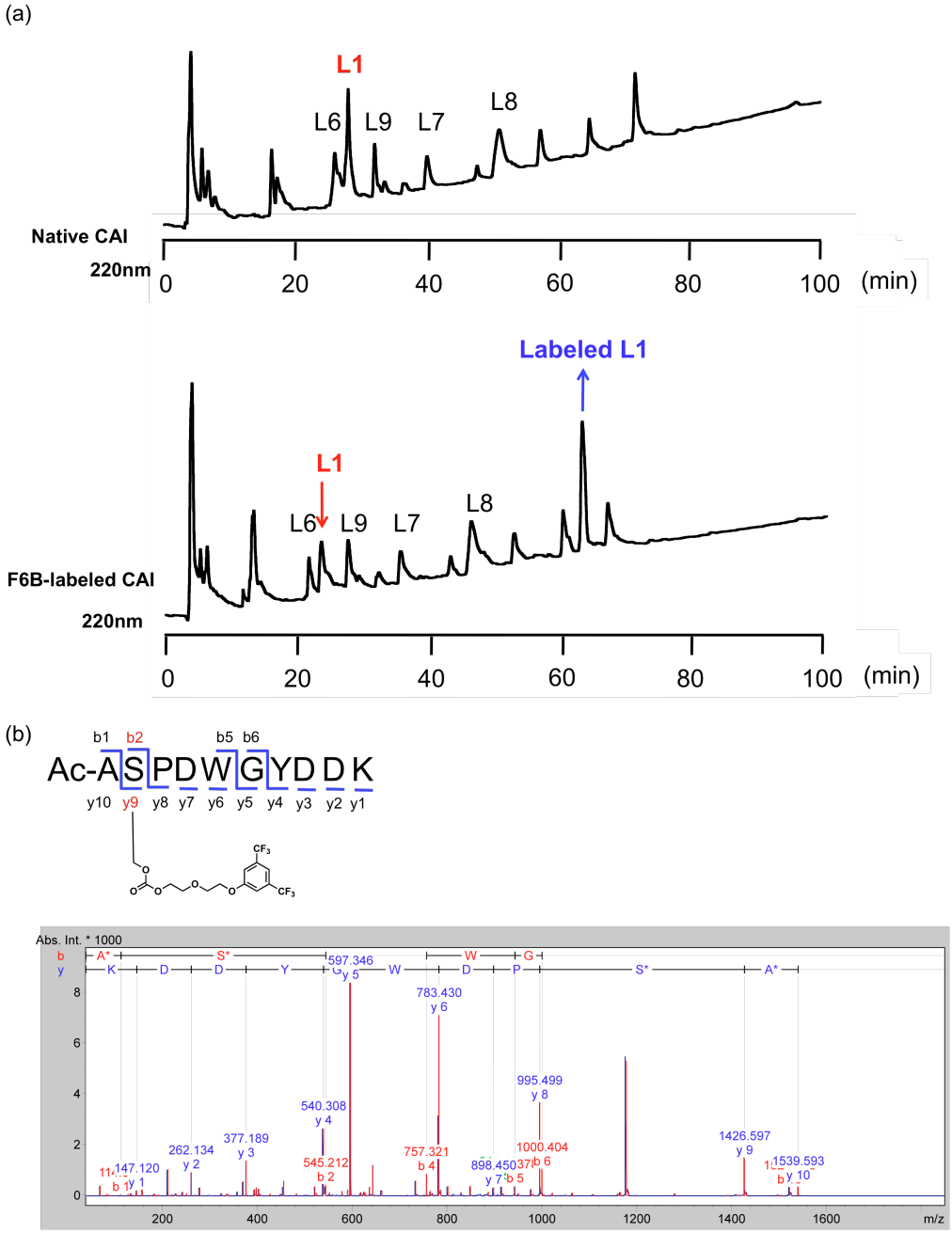


Figure S6. Identification of the labeling site of the F6B-installed CAI. (a) HPLC analysis of the digested fragments of native and F6B labeled CAI by **8**. It shows the chromatogram of the fragments derived from the labeled CAI detected with UV absorption (220 nm). The peaks marked with Labeled L1 correspond to the labeled fragments, which were characterized by MALDI-TOF MS analysis. MALDI-TOF-MS (CHCA) of peak L1*: calcd. for $[M+H]^+$ = 1539.54, obsd. 1539.6. (b) MALDI-TOF MS/MS analysis of the F6B-labeled L1 fragments.

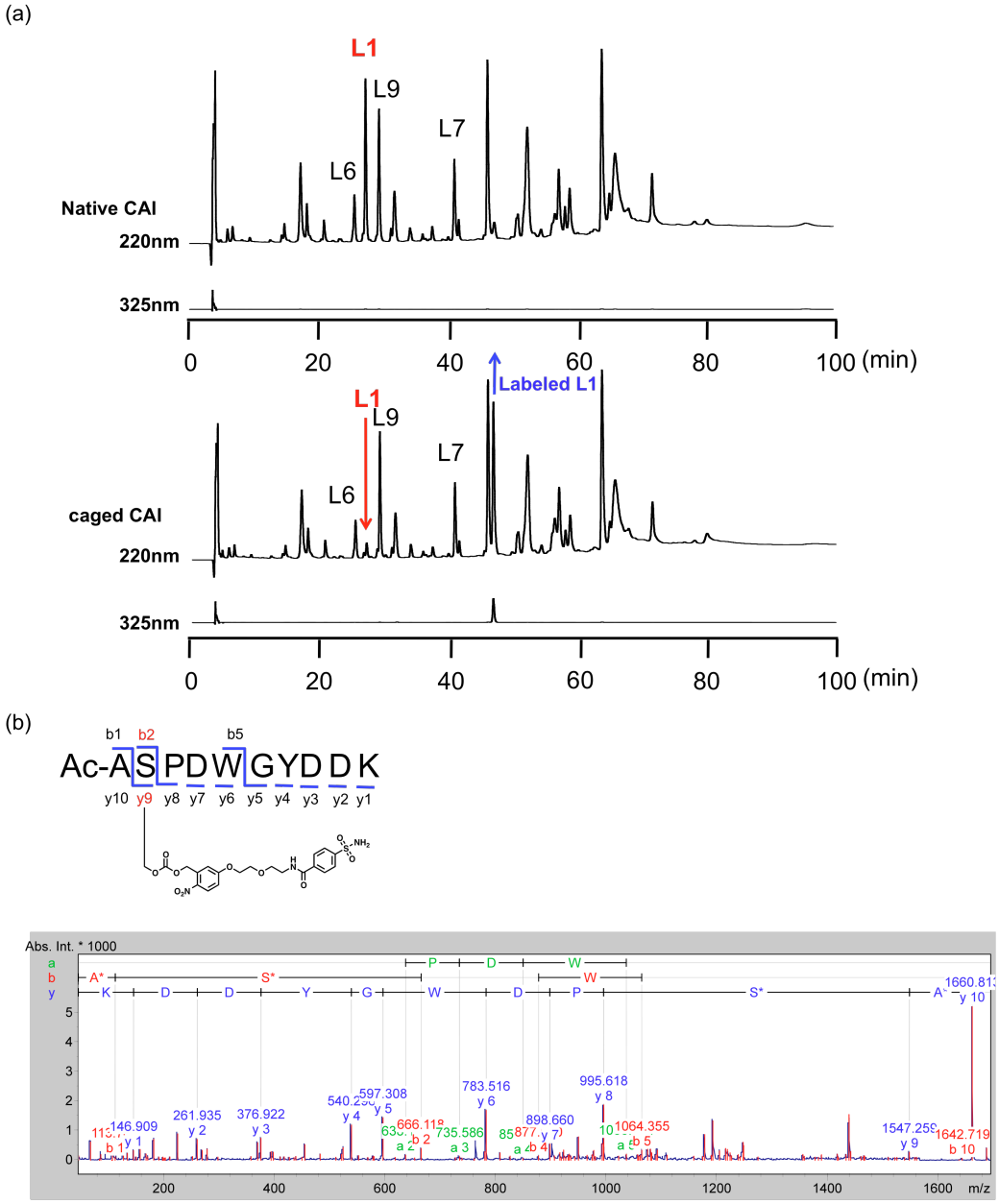


Figure S7. Identification of the labeling site of the caged-labeled CAI. (a) HPLC analysis of the digested fragments of caged labeled CAI by **9**. It shows the chromatogram of the fragments derived from the labeled CAI detected with UV absorption (220-400nm). The peaks marked with Labeled L1 correspond to the labeled fragments, which were characterized by MALDI-TOF mass analysis. MALDI-TOF mass (CHCA) of peak L1*: calcd. for $[M+H]^+$ = 1660.57, obsd. 1660.8. (b) MALDI-TOF MS/MS analysis of the caged-labeled L1 fragments.

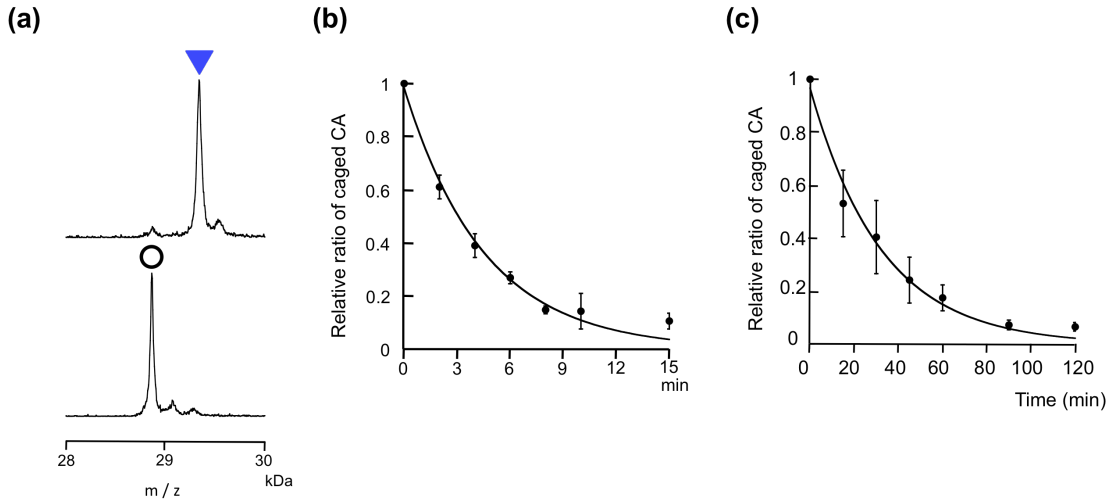


Figure S8. Uncaging efficiency of caged CAI. (a) MALDI-TOF mass charts of caged CAI (▼) and uncaged CAI (○) by photoirradiation for 90 min at 25 °C with 100W Hg lamp. Time dependency of the ratio of remaining caged CAI photoirradiated by 500W (b) and 100W (c) Hg lamp.

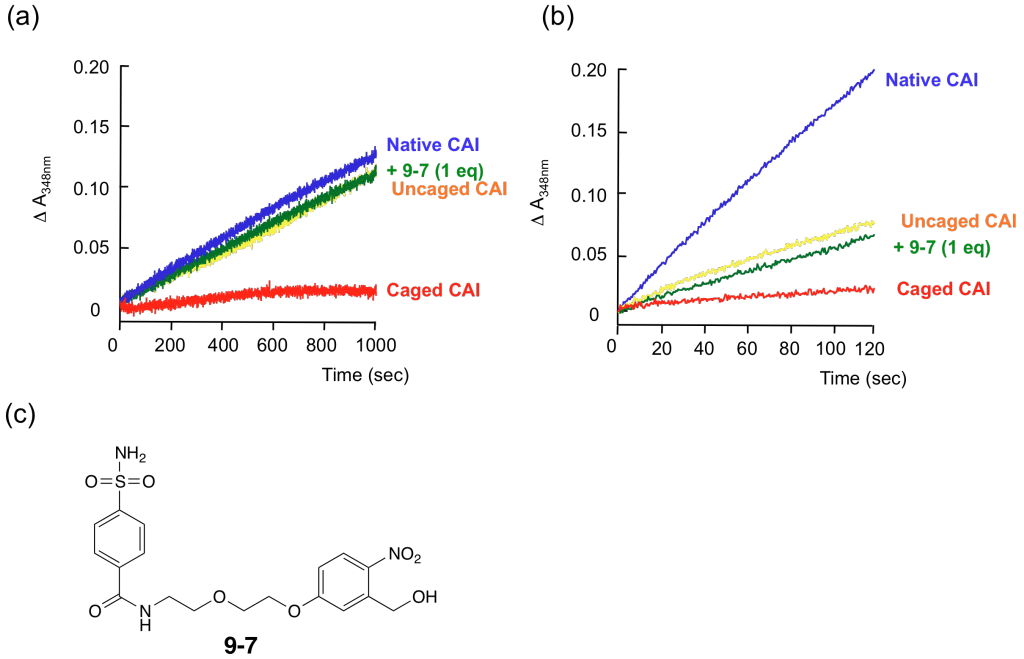


Figure S9. Concentration dependency of uncaged CAI in enzyme activity. (a) 0.10 μM CA activity of native CAI (blue line), native CAI + 1eq inhibitor **9-7** (green line), caged CAI (red line), and uncaged CAI (yellow line, photoirradiation time : 90 min with 100W Hg lamp). All measurements were performed at [substrate] = 1.0 mM at 25 °C. (b) 1 μM CA activity of native CAI (blue line), native CAI + 1eq inhibitor **9-7** (green line), and uncaged CAI (yellow line). All measurements were performed at [substrate] = 1.0 mM at 25 °C. (c) Chemical structure of inhibitor **9-7** (see **scheme S7**).

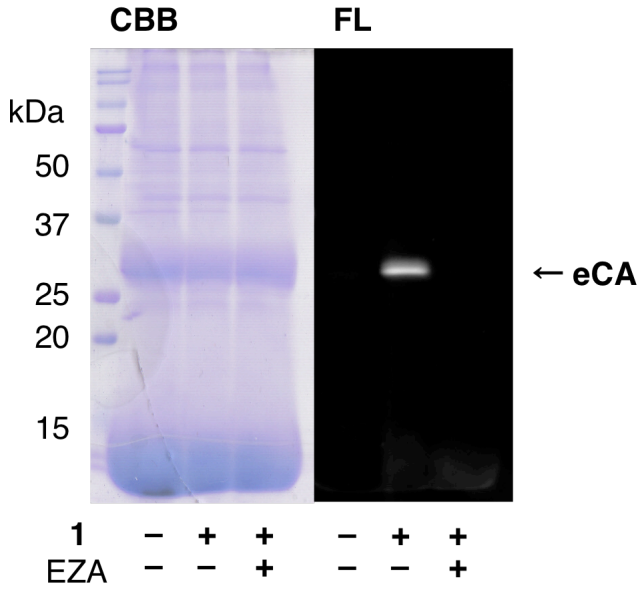


Figure S10. Labeling of eCA in human red blood cell lysate. The lysate (20-fold diluted with HEPES-buffered saline (HBS, pH 7.4)) was mixed with **1** (100 μM) and incubated at 37 °C for 24 h in the presence (lane 3) or absence (lane 2) of EZA (500 μM). The left and right image shows the CBB staining and fluorescence image of the SDS-PAGE gel, respectively.

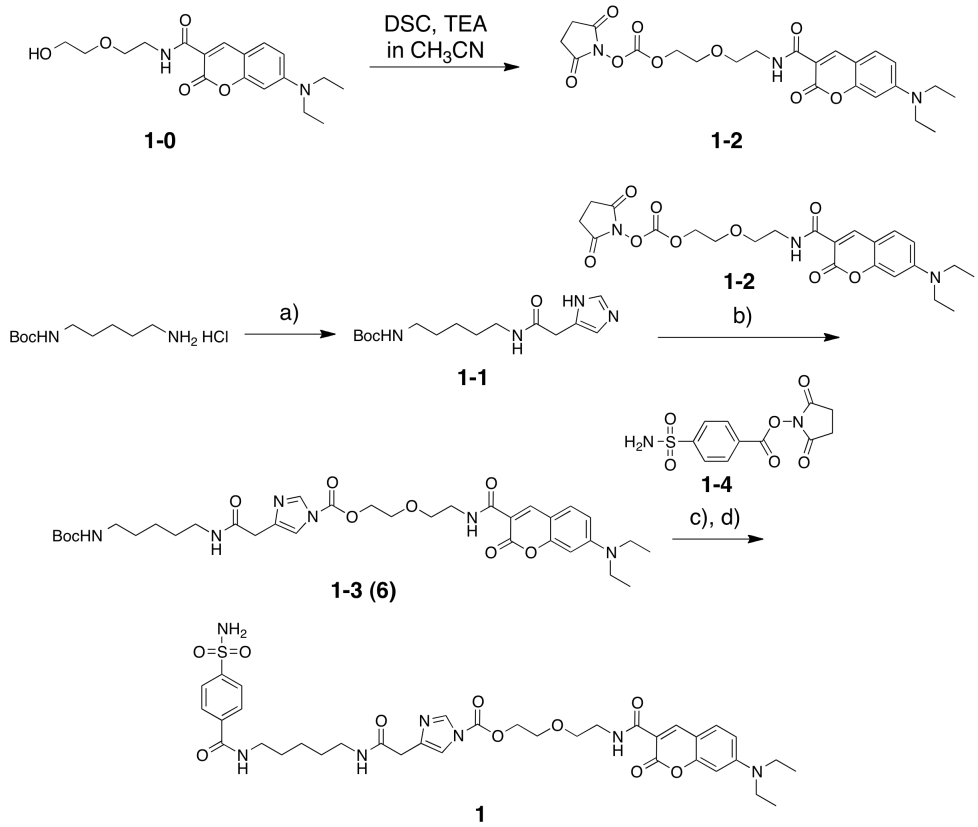
Supplementary Methods

General materials and methods for organic synthesis

All chemical reagents and solvents were obtained from commercial suppliers (Aldrich, Tokyo Chemical Industry (TCI), Wako Pure Chemical Industries, Acros Organics, Sasaki Chemical, or Watanabe Chemical Industries) and used without further purification.

Thin layer chromatography (TLC) was performed on silica gel 60 F_{254} precoated aluminium sheets (Merck) and visualized by fluorescence quenching or ninhydrin staining. Chromatographic purification was conducted by flash column chromatography on silica gel 60 N (neutral, 40–50 μm, Kanto Chemical). ^{1}H NMR spectra were recorded in deuterated solvents on a Varian Mercury 400 (400 MHz) spectrometer and calibrated to the residual solvent peak or tetramethylsilane (= 0 ppm). Multiplicities are abbreviated as follows: s = singlet, d = doublet, t = triplet, q = quartet, quin = quintet, m = multiplet, dd = double doublet, br s = broad singlet. Matrix-assisted laser desorption/ionization time-of-flight mass spectroscopy (MALDI-TOF MS) was measured by Autoflex II (Bruker Daltonics, Billerica, MA) and Ultraflex III (Bruker Daltonics, Billerica, MA) spectrometers using α-cyano-4-hydroxycinnamic acid (CHCA) or sinapinic acid (SA) as the matrix. High-resolution electrospray ionization quadrupole fourier transform mass spectroscopy (HR-ESI Qq-LTMS) were measured by a Bruker apex-ultra (7T) mass spectrometer, respectively.

Scheme S1. Synthesis of **1**



Reaction conditions: (a) *1H*-imidazole-4-acetic acid hydrochloride, EDCI•HCl, HOBt•H_2O, DIPEA in DMF, (b) **1-2**, pyridine in DCM, (c) TFA in DCM and (d) **1-4**, DIPEA in DMF.

Compound **1-2**

A solution of **1-0**[S1] (120 mg, 0.34 mmol), *N,N'*-disuccinimidyl carbonate (DSC) (264 mg, 1.02 mmol) and triethylamine (101 mg, 1.02 mmol) in anhydrous CH_3CN (2 mL) was stirred for 6 h at room temperature (rt). After dilution with water, the resulting mixture was extracted with AcOEt. The organic layer was washed with sat. $NaHCO_3$*aq* and dried over Na_2SO_4 followed by concentration *in vacuo*. The residue was purified by flash column chromatography on SiO_2 ($CHCl_3$: MeOH = 40 : 1) to give **1-2** (115 mg, 70%) as a yellow amorphous powder. ^{1}H-NMR (400 MHz, $CDCl_3$) δ 8.99 (s, 1H), 8.68 (s, 1H), 7.42 (d, J = 9.2 Hz, 1H), 6.65 (d, J = 6.4 Hz, 1H), 6.49 (s, 1H), 4.48 (t, J = 4.8 Hz, 2H), 3.79 (t, J = 4.4 Hz, 2H), 3.67-3.69 (m, 4H), 3.45 (q, J = 7.2 Hz, 4H), 1.24 (t, J = 7.2 Hz, 6H). HR-ESI MS *m/z* calcd for $[M+H]^+$ 490.1820, found 490.1817.

Compound **1-1**

A solution of *N*-(*tert*-Butoxycarbonyl)-1,5-diaminopentane hydrochloride (1.00 g, 4.94 mmol), imidazole-4-acetic acid hydrochloride (0.96 g, 5.93 mmol), HOBt•H_2O (1.14 g, 7.41 mmol), EDCI•HCl (1.42 g, 7.41 mmol) and DIPEA (2.58 mL, 14.8 mmol) in dry DMF (100 mL) was stirred for 4 h at rt. After removal of the solvent by evaporation, the residue was diluted with AcOEt. The organic layer was extracted with 5% citric acid *aq*. Then, the aqueous layer was alkalized with aqueous NH_3 solution (pH = 9) and extracted with AcOEt. The organic layer was dried over Na_2SO_4 and concentrated by evaporation. The residue was purified by flash column chromatography on SiO_2 ($CHCl_3$: MeOH : NH_3 = 100 : 10 : 1) to give **1-1** (705 mg, 50 %) as a colorless amorphous powder. ^{1}H-NMR (400 MHz, CD_3OD) δ 7.61 (s, 1H), 6.95 (s, 1H), 3.48 (s, 2H), 3.18 (t, *J* = 6.8 Hz, 2H), 3.01 (t, *J* = 6.8 Hz, 2H), 1.43-1.60 (m, 13H), 1.32 (quin, *J* = 7.2 Hz, 2H). HR-ESI MS *m/z* calcd for $[M+H]^+$ 311.2078, found 311.2076.

Compound **1-3** (**6**)

A solution of **1-1** (55 mg, 180 μmol), **1-2** (83 mg, 180 μmol), and pyridine (14 mg, 180 μmol) in CH_2Cl_2 (7 mL) was stirred overnight at rt. After removal of the solvent by evaporation, the residue was dissolved in AcOEt. The organic layer was washed with sat. $NaHCO_3$ *aq* and 1 N HCl. After drying over Na_2SO_4, the solvent was removed *in vacuo* to give **1-3 (6)** (92 mg, 75%) as a yellow amorphous powder. ^{1}H-NMR (400 MHz, $CDCl_3$) δ 9.09 (br s, 1H), 8.67 (s, 1H), 8.19 (s, 1H), 7.44 (d, *J* = 3.6 Hz, 1H), 7.41 (s, 1H), 6.65 (d, *J* = 9.2 Hz, 1H), 6.55 (s, 1H), 4.57 (t, *J* = 4.4 Hz, 2H), 3.83 (t, *J* = 4.4 Hz, 2H), 3.66-3.71 (m, 4H), 3.53 (s, 2H), 3.47 (q, *J* = 7.2 Hz, 4H), 3.22 (q, *J* = 6.0 Hz, 2H), 3.07 (q, *J* = 6.0 Hz, 2H), 1.43-1.50 (m, 13H), 1.23-1.33 (m, 8H). HR-ESI MS *m/z* calcd for $[M+H]^+$ 685.3556, found 685.3560.

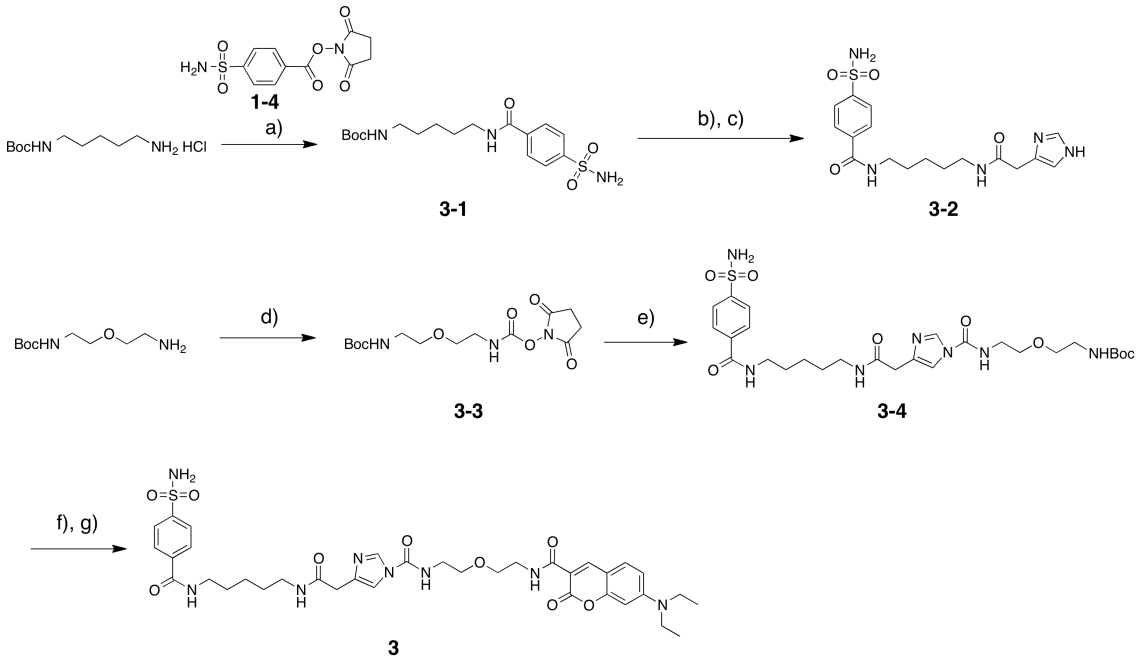
Compound **1**

A solution of **1-3** (50 mg, 80 μmol) in CH_2Cl_2 (2 mL) / TFA (1.5 mL) was stirred for 2 h at rt. After removal of the solvent in vacuo, the residual TFA was removed with toluene (twice) to give deprotected **1-3**.

A mixture of deprotected **1-3**, **1-4**[S1] (24 mg, 80 μmol), DIPEA (55 μL, 320 μmol) in

dry DMF (5 mL) was stirred for 10 h at rt. After removal of the solvent *in vacuo*, the residue was dissolved in $CHCl_3$. The organic layer was washed with sat. $NaHCO_3$ *aq* and dried over Na_2SO_4 followed by concentration *in vacuo*. The residue was purified by flash column chromatography on SiO_2 ($CHCl_3$: MeOH = 10 : 1 to 5 : 1) to give **1** (45 mg, 59%) as a yellow amorphous powder. ^{1}H-NMR (400 MHz, $CDCl_3$: CD_3OD = 1 : 1) δ 9.19 (br s, 1H), 8.62 (s, 1H), 8.19 (s, 1H), 7.97 (d, *J* = 8.4 Hz, 2H), 7.94 (d, *J* = 8.4 Hz, 2H), 7.50 (d, *J* = 9.2 Hz, 1H), 7.49 (s, 1H), 6.75 (d, *J* = 9.2 Hz, 1H), 6.55 (s, 1H), 4.61 (m, 2H), 3.88 (t, *J* = 4.8 Hz, 2H), 3.73 (t, *J* = 4.8 Hz, 2H), 3.64 (q, *J* = 4.4 Hz, 2H), 3.52 (q, *J* = 7.2 Hz, 4H), 3.48 (s, 2H), 3.40 (t, *J* = 6.4 Hz, 2H), 3.22 (q, *J* = 6.0 Hz, 2H), 1.64 (quin, *J* = 7.2 Hz, 2H), 1.56 (quin, *J* = 7.6 Hz, 2H), 1.40 (quin, *J* = 7.2 Hz, 2H), 1.27 (t, *J* = 7.2 Hz, 6H). HR-ESI MS *m/z* calcd for $[M+H]^+$ 768.3021, found 768.3029.

Scheme S2. Synthesis of **3**



Reaction conditions: (a) **1-4**, DIPEA in DMF, (b) TFA in DCM, (c) *1H*-imidazole-4-acetic acid hydrochloride, EDCI•HCl, HOBt•H_2O, DIPEA in DMF, (d) DSC, TEA in DMF at 40 °C, (e) **3-2**, pyridine in DMF, (f) TFA in DCM and (g) 7-diethylaminocoumarin-3-carboxylic acid, EDCI•HCl, HOBt•H_2O, DIPEA in DMF.

Compound **3-2**

A solution of **1-4** (200 mg, 0.67 mmol), *N*-(*tert*-butoxycarbonyl)-1,5-diaminopentane (136 mg, 0.67 mmol), and DIPEA (0.446 mL, 2.68 mmol) in dry DMF (20 mL) was stirred for 2 h at rt. After removal of the solvent *in vacuo*, the residue was dissolved in $CHCl_3$. The organic layer was washed with water and dried over Na_2SO_4 followed by concentration *in vacuo* to give **3-1** (247 mg, 95%) as a white solid.

A solution of **3-1** (150 mg, 390 μmol) in TFA (4 mL) was stirred for 1.5 h at rt. After removal of the solvent *in vacuo*, the residual TFA was further removed with toluene (twice). The residue was dissolved in dry DMF (17 mL), and the solution was mixed with DIPEA (389 μL, 1.16 mmol), imidazole-4-acetic acide hydrochloride (59 mg, 85 μmol), HOBt•H_2O (88 mg, 580 μmol) and EDCI•HCl (111 mg, 580 μmol). The reaction mixture was stirred overnight at rt. After removal of the solvent in vacuo, the residue was purified by flash column chromatography on SiO_2 ($CHCl_3$ / MeOH = 3 / 1) to give **3-2** (103 mg, 67%) as a colorless oil. ^{1}H-NMR (400 MHz, CD_3OD) δ 7.96 (m,

4H), 7.79 (s, 1H), 7.00 (s, 1H), 3.52 (s, 2H), 3.38 (t, *J* = 7.2 Hz, 2H), 3.19-3.23 (m, 2H), 1.64 (m, 2H), 1.56 (m, 2H), 1.40 (m, 2H). HR-ESI MS *m/z* calcd for $[M+H]^+$ 394.1544, found 394.1539.

Compound **3-3**

To a solution of *N,N'*-disuccinimidyl carbonate (279 mg, 1.09 mmol) in dry DMF (2 mL) was added dropwise the solution of **3-1** (44 mg, 0.215 mmol) and TEA (0.151 mL, 1.09 mmol) in dry DMF (10 mL) at 40 °C. The reaction mixture was stirred for 2 h at rt. After removal of the solvent *in vacuo*, the residue was dissolved in AcOEt. The organic layer was washed with 1N HCl and dried over Na_2SO_4. After removal of the solvent *in vacuo*, **3-3** (102 mg, containing a small amount of impurities) was obtained as a white solid. HR-ESI MS *m/z* calcd for $[M+Na]^+$ 368.1428, found 368.1427.

A solution of **3-3** (27 mg, 79 μmol), **3-2** (31 mg, 79 μmol), and pyridine (6 μL, 79 μmol) in dry DMF (2 mL) was stirred for 3 days at rt. After removal of the solvent *in vacuo*, the residue was purified by flash column chromatography on SiO_2 ($CHCl_3$: MeOH = 10 : 1 to 8 : 1) to give **3-4** (15 mg, 42% in 2 steps) as a white powder. ^{1}H-NMR (400 MHz, $CDCl_3$: CD_3OD = 1 : 1) δ 8.12 (s, 1H), 7.91-7.97 (m, 5H), 7.49 (br s, 1H), 7.42 (s, 1H), 5.66 (br s, 1H), 3.61-3.68 (m, 6H), 3.52 (q, *J* = 5.2 Hz, 4H), 3.28 (q, *J* = 5.2 Hz, 2H), 3.24 (q, *J* = 5.2 Hz, 2H), 1.65 (quin, *J* = 7.2 Hz, 2H), 1.56 (quin, *J* = 7.2 Hz, 2H), 1.39-1.43 (m, 11H). HR-ESI MS *m/z* calcd for $[M+H]^+$ 624.2810, found 624.2811.

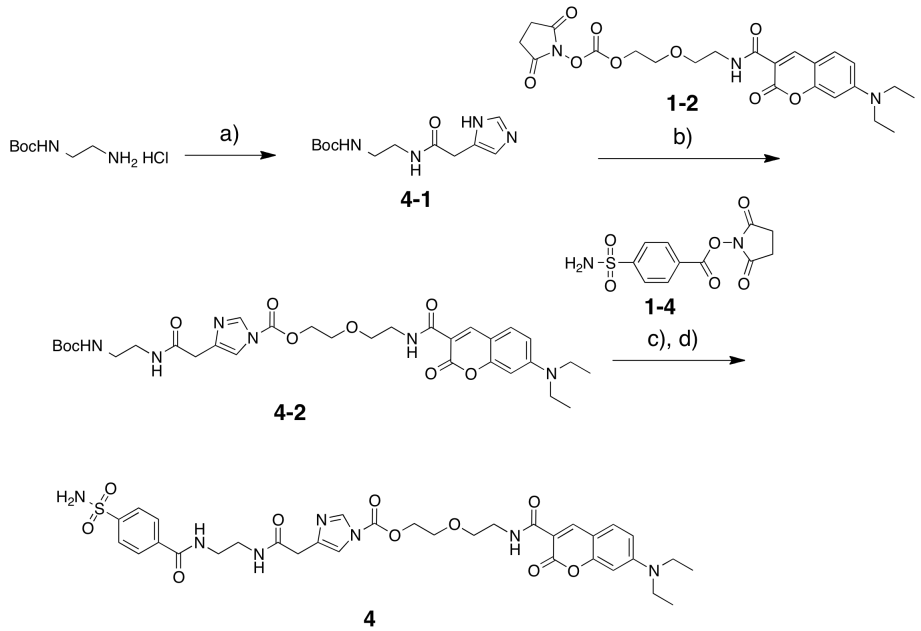
Compound **3**

A solution of **3-4** (5 mg, 8.0 μmol) in TFA (2 mL) was stirred for 1 h at rt. After removal of the solvent *in vacuo*, the residual TFA was further removed with toluene (twice) to give the deprotected **3-4**.

A solution of the deprotected **3-4**, 7-diethylaminocoumarin-3-carboxylic acid (3 mg, 10 μmol), EDCI•HCl (2 mg, 12 μmol), HOBt •H_2O (2 mg, 12 μmol), and DIPEA (20 μL) in dry DMF (2 mL) was stirred for 10 h at rt. After removal of the solvent *in vacuo*, the residue was purified by flash column chromatography on SiO_2 ($CHCl_3$: MeOH = 8 : 1) to give **3** (2 mg, 33 % in 2 steps) as a yellow powder. ^{1}H-NMR (400 MHz, $CDCl_3$: CD_3OD = 4 : 1) δ 9.21 (br s, 1H), 8.64 (s, 1H), 8.19 (s, 1H), 7.93-7.98 (m, 4H), 7.55 (s, 1H), 7.48 (d, *J* = 8.8 Hz, 1H), 6.70 (d, *J* = 8.8 Hz, 1H), 6.55 (s, 1H), 3.59 (t, *J* = 4.4 Hz,

2H), 3.37-3.52 (m, 14H), 3.23 (q, *J* = 6.4 Hz, 2H), 1.63 (quin, *J* = 7.6 Hz, 2H), 1.54 (quin, *J* = 7.6 Hz, 2H), 1.37 (quin, *J* = 7.6 Hz, 2H), 1.26 (t, *J* = 6.0 Hz, 6H). HR-ESI MS *m/z* calcd for $[M+H]^+$ 767.3181, found 767.3188.

Scheme S3. Synthesis of **4**



Reaction conditions: (a) *1H*-imidazole-4-acetic acid hydrochloride, EDCI•HCl, HOBt•H_2O, DIPEA in DMF, (b) **1-2**, pyridine in DCM, (c) TFA in DCM and (d) **1-4**, DIPEA in DMF.

Compound **4-1**

N-(*tert*-Butoxycarbonyl)-1,2-diaminoethane hydrochloride (100 mg, 0.51 mmol) was used as a starting material. By the same procedure described for the synthesis of **1-1**, **4-1** (75 mg, 55%) was obtained as a colorless amorphous powder. ^{1}H-NMR (400 MHz, CD_3OD) δ 7.61 (s, 1H), 6.95 (s, 1H), 3.50 (s, 2H), 3.27 (t, *J* = 6.0 Hz, 2H), 3.14 (t, *J* = 6.0 Hz, 2H), 1.43 (s, 9H). HR-ESI MS *m/z* calcd for $[M+H]^+$ 269.1608, found 269.1608.

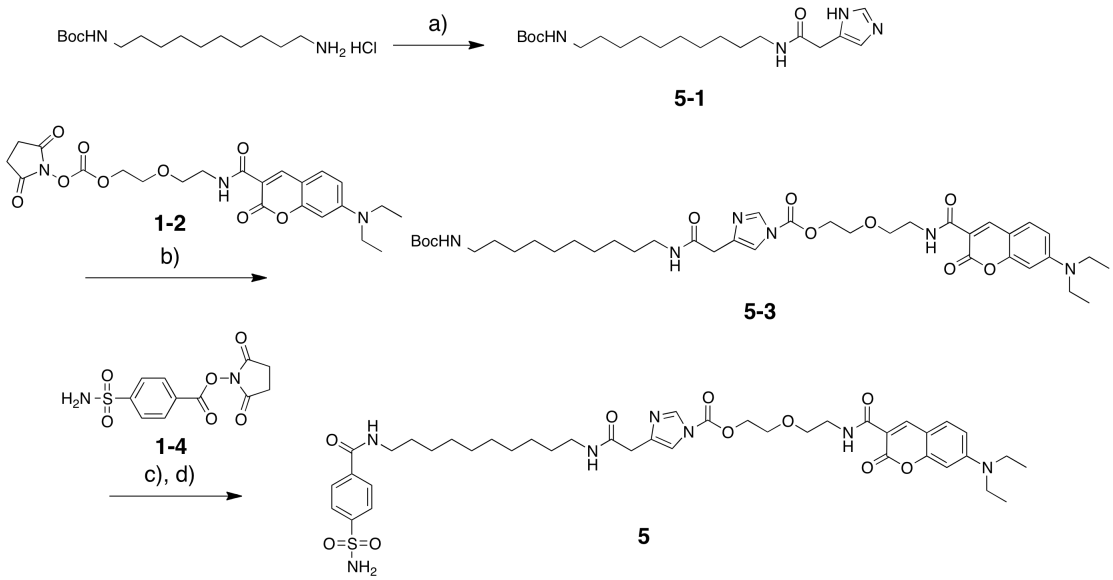
Compound **4-2**

4-1 (30 mg, 0.11 mmol) was used as a starting material. By the same procedure described for the synthesis of **1-3**, **4-2** (72 mg, quant.) was obtained as a yellow amorphous powder. ^{1}H-NMR (400 MHz, $CDCl_3$) δ 9.09 (br s, 1H), 8.67 (s, 1H), 8.17 (s, 1H), 7.45 (d, *J* = 5.2 Hz, 1H), 7.42 (s, 1H), 6.65 (d, *J* = 9.2 Hz, 1H), 6.56 (s, 1H), 4.57 (t, *J* = 4.4 Hz, 2H), 3.83 (t, *J* = 4.4 Hz, 2H), 3.65-3.71 (m, 4H), 3.55 (s, 2H), 3.49 (q, *J* = 7.2 Hz, 4H), 3.36 (q, *J* = 6.0 Hz, 2H), 3.24 (t, *J* = 5.6 Hz, 2H), 1.40 (s, 9H), 1.25 (t, *J* = 7.2 Hz, 6H). HR-ESI MS *m/z* calcd for $[M+H]^+$ 643.3086, found 643.3080.

Compound **4**

4-2 (72 mg, 110 μmol) was used as a starting material. By the same procedure described for the synthesis of **1**, compound **4** (22 mg, 27%) was obtained as a yellow amorphous powder. ^{1}H-NMR (400 MHz, $CDCl_3$) δ 9.03 (br s, 1H), 8.59 (s, 1H), 8.02 (s, 1H), 7.85 (d, *J* = 8.0 Hz, 2H), 7.77 (d, *J* = 8.4 Hz, 2H), 7.58 (br s, 1H), 7.40-7.43 (m, 2H), 6.65 (d, *J* = 8.8 Hz, 1H), 6.52 (s, 1H), 5.68 (br s, 1H), 4.56 (br s, 2H), 3.83 (br s, 2H), 3.70 (t, *J* = 4.8 Hz, 2H), 3.64 (t, *J* = 5.2 Hz, 2H), 3.53-3.55 (m, 6H), 3.47 (q, *J* = 7.2 Hz, 4H), 1.25 (t, *J* = 7.2 Hz, 6H). HR-ESI MS *m/z* calcd for $[M+H]^+$ 726.2552, found 726.2558.

Scheme S4. Synthesis of **5**



Reaction conditions: (a) *1H*-imidazole-4-acetic acid hydrochloride, EDCI•HCl, HOBt•H_2O, DIPEA in DMF, (b) **1-2**, pyridine in DCM, (c) TFA in DCM and (d) **1-4**, DIPEA in DMF.

Compound **5-1**

N-(*tert*-Butoxycarbonyl)-1,10-diaminodecane hydrochloride (1.00 g, 3.23 mmol) was used as a starting material. By the same procedure described for the synthesis of **1-1**, **5-1** (700 mg, 57%) was obtained as a colorless amorphous powder. ^{1}H-NMR (400 MHz, CD_3OD) δ 7.61 (s, 1H), 6.94 (s, 1H), 3.48 (s, 2H), 3.17 (t, J = 6.8 Hz, 2H), 3.01 (t, J = 7.2 Hz, 2H), 1.43-1.52 (m, 13H), 1.30 (br s, 12H). HR-ESI MS *m/z* calcd for $[M+H]^+$ 381.2860, found 381.2859.

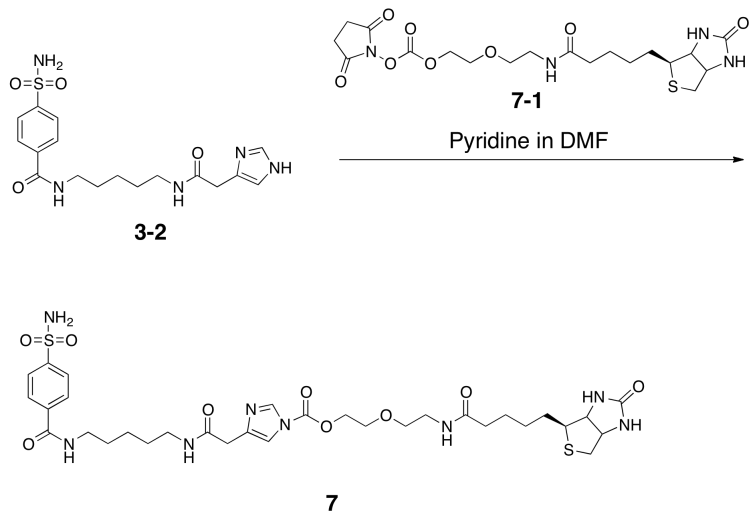
Compound **5-2**

5-1 (60 mg, 0.15 mmol) was used as a starting material. By the same procedure described for the synthesis of **1-3**, **5-2** (98 mg, 86%) was obtained as a yellow amorphous power. ^{1}H-NMR (400 MHz, $CDCl_3$) δ 9.10 (t, J = 4.8 Hz, 1H), 8.67 (s, 1H), 8.18 (s, 1H), 7.45 (d, J = 5.6 Hz, 1H), 7.42 (s, 1H), 6.65 (d, J = 8.8 Hz, 1H), 6.56 (s, 1H), 4.57 (t, J = 4.4 Hz, 2H), 3.83 (t, J = 4.4 Hz, 2H), 3.70 (t, J = 4.8 Hz, 2H), 3.67 (t, J = 4.8 Hz, 2H), 3.53 (s, 2H), 3.47 (q, J = 7.2 Hz, 4H), 3.21 (q, J = 6.8 Hz, 2H), 3.09 (q, J = 6.0 Hz, 2H), 1.44 (m, 13H), 1.23-1.28 (m, 18H). HR-ESI MS *m/z* calcd for $[M+H]^+$ 755.4338, found 755.4340.

Compound **5**

5-2 (50 mg, 66 μmol) was used as a starting material. By the same procedure described for the synthesis of **1**, compound **5** (42 mg, 76%) was obtained as a yellow amorphous powder. ^{1}H-NMR (400 MHz, $CDCl_3$: CD_3OD = 1 : 1) δ 9.19 (br s, 1H), 8.63 (s, 1H), 8.20 (s, 1H), 7.93 (d, *J* = 8.4 Hz, 2H), 7.90 (d, *J* = 8.4 Hz, 2H), 7.48 (s, 1H), 7.47 (d, *J* = 9.2 Hz, 1H), 6.69 (d, *J* = 9.2 Hz, 1H), 6.53 (s, 1H), 4.59 (m, 2H), 3.71 (t, *J* = 5.2 Hz, 2H), 3.65 (t, *J* = 6.0 Hz, 2H), 3.46-3.51 (m, 4H), 3.36-3.42 (m, 6H), 3.18 (q, *J* = 6.0 Hz, 2H), 1.62 (m, 2H), 1.46 (m, 2H), 1.21-1.34 (m, 20H). HR-ESI MS *m/z* calcd for $[M+H]^+$ 838.3804, found 838.3806.

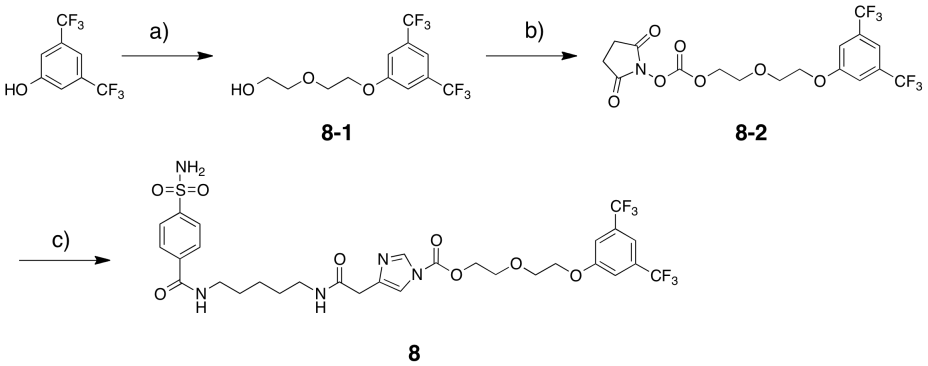
Scheme S5. Synthesis of **7**



Compound **7**

To a stirred solution of **3-2** (22 mg, 56 μmol) in dry DMF (2 mL) was added **7-1**[S2] (32 mg, 67 μmol), pyridine (5.1 μL) at room temperature. The reaction mixture was allowed to stir for 12 h at rt. The reaction mixture was purified by column chromatography ($CHCl_3$: MeOH = 10 : 1 to 5 : 1) to yield **7** (239 mg). The reaction mixture was purified by HPLC [(0.1 % TFA MeCN : 0.1 % TFA H_2O, linear gradient) = 5 : 95 (0 min), 30 : 70 (50 min), 100 : 0 (60 min) to 5 : 95 (70 min)] to yield **7** (13 mg, 30 %) as a white powder. ^{1}H-NMR (400 MHz, CD_3OD) δ 8.20 (s, 1H), 7.95 (m, 4H), 7.45 (s, 1H), 4.59 – 4.55 (m, 2H), 4.47 (dd, J = 4.8, 7.6 Hz, 1H), 4.28 (dd, J = 4.8, 7.6 Hz, 1H), 3.81 (t, J = 5.6 Hz, 2H), 3.57 (t, J = 5.6 Hz, 2H), 3.48 (s, 2H), 3.40 – 3.33 (m, 3H), 3.18 – 3.15 (m, 4H), 2.90 (dd, *J* = 4.8, 12.8 Hz, 1H), 2.69 (d, *J* = 12.8 Hz, 1H), 2.17 (t, *J* = 7.2 Hz, 2H), 1.73 – 1.54 (m, 8H), 1.45 – 1.38 (m, 4H). HR-ESI MS *m/z* calcd for $[M+H]^+$ 751.2902, found 751.2902.

Scheme S6. Synthesis of **8**



Reaction conditions: (a) ethylene glycol mono-2-chloroethyl ether, K_2CO_3, KI in DMF at 60 °C (b) DSC, TEA in DMF and (c) **3-2**, pyridine in DMF.

Compound **8-1**

To a stirred solution of 3,5-bis(trifluoromethyl)phenol (800 mg, 3.48 mmol), K_2CO_3 (1.92 g, 13.9 mmol), KI (1.16 g, 6.96 mmol) in dry *N,N*-dimethylformamide (3 mL) was added ethylene glycol mono-2-chloroethyl ether (1.30 g, 10.4 mmol). The reaction mixture was allowed to stir for 40 h at 60 °C. To the reaction mixture was added sat. $NaHCO_3$ *aq* to quench redundant ethylene glycol mono-2-chloroethyl ether. The reaction mixture was allowed to stir for 7 h at 60 °C. The reaction mixture was diluted with AcOEt. The solution was washed with *sat.* $NaHCO_3$, 5 % citric acid *aq*, brine. The organic layer was dried over $MgSO_4$, filtered, evaporated to yield **8-1** (1.26 g, quantitative yield) as pale yellow oil. ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.47 (s, 1H), 7.35 (s, 1H), 4.23 (t, J = 4.4 Hz, 2H), 3.91 (t, J = 4.4 Hz, 2H), 3.78 (t, J = 4.4 Hz, 2H), 3.69 (t, J = 4.4 Hz, 2H), 1.95 (br, 1H).

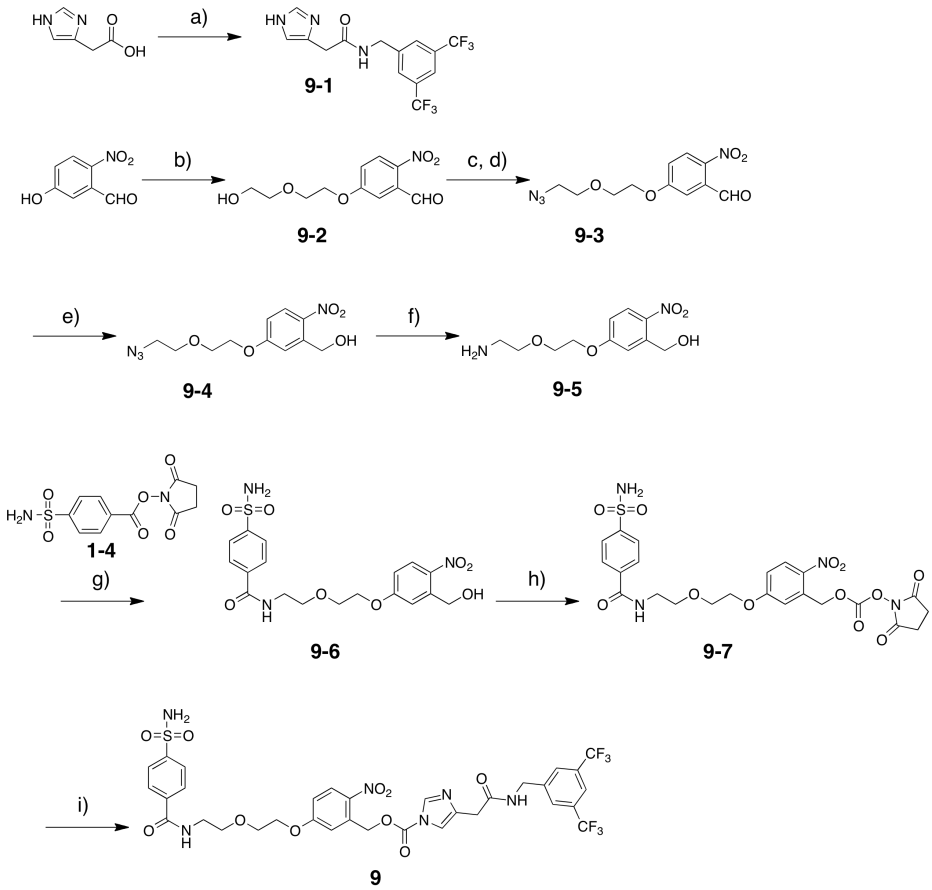
Compound **8-2**

To a stirred solution of **8-1** (60 mg, 0.189 mmol) in dry DMF (3 mL) was added *N,N′*-disuccinimidyl carbonate (145 mg, 0.567 mmol), triethylamine (78.6 μL, 0.567 mmol) at room temperature. The reaction mixture was allowed to stir for 6 h at room temperature. The reaction mixture was purified by column chromatography ($CHCl_3$) to yield **8-2** (70 mg, 80%) as pale yellow oil. ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.47 (s, 1H), 7.37 (s, 1H), 4.50 (t, J = 4.4 Hz, 2H), 4.24 (t, J = 4.4 Hz, 2H), 3.92 (t, J = 4.4 Hz, 2H), 3.87 (t, J = 4.4 Hz, 2H), 2.84 (s, 4H).

Compound **8**

To a stirred solution of **3-2** (16 mg, 41 μmol) in dry DMF (3 mL) was added **8-2** (23 mg, 49 μmol, 1.2 eq), pyridine (3.7 μL, 49 mmol) at rt. The reaction mixture was allowed to stir for 6 h at rt. The reaction mixture was purified by column chromatography ($CHCl_3$: MeOH = 20 : 1 to 10 : 1) to yield **8** (12 mg, 41%). ^{1}H-NMR (400 MHz, $CDCl_3/CD_3OD$ = 7/3) δ 8.12 (s, 1H), 7.94 (d, *J* = 4.4 Hz, 4H), 7.56 (s, 1H), 7.45 (s, 2H), 7.40 (s, 1H), 4.60 (t, *J* = 4.4 Hz, 2H), 3.94 (t, *J* = 4.4 Hz, 2H), 3.45 (t, *J* = 4.4 Hz, 2H), 3.42 (t, *J* = 4.4 Hz, 2H), 3.36 (t, *J* = 1.6 Hz, 1H), 3.27 – 3.22 (m, 2H), 1.65 – 1.63 (m, 2H), 1.57 – 1.54 (m, 2H), 1.47 – 1.40 (m, 2H). HR-ESI MS *m/z* calcd for $[M+H]^+$ 738.2027, found 738.2025.

Scheme S7. Synthesis of **9**



Reaction conditions: (a) 3,5-bis(trifluoromethyl)benzyl amine, EDCI•HCl, HOBt•H_2O, DIPEA in DMF, (b) 2-(2-chloroethoxy)ethanol, K_2CO_3 in DMF at 50 °C, (c) MsCl, TEA in DCM at 0 °C, (d) NaN_3 in DMF at 50 °C, (e) $NaBH_4$ in THF, (f) PPh_3 in THF/H_2O, (g) **1-4**, DIPEA in DMF, (h) DSC, DIPEA in DMF and (i) **9-1**, DIPEA in DMF

Compound **9-1**

To a stirred solution of *1H*-imidazole-4-acetic acid hydrochloride (446 mg, 3.0 mmol), EDCI•HCl (578 mg, 3.0 mmol), HOBt•H_2O (105 mg, 3.0 mmol) and DIPEA (1433 µL) in DMF (20 mL) was added 3,5-bis(trifluoromethyl)benzyl amine (734 mg, 2.74 mmol). The reaction mixture was stirred at rt for 1 hour. The mixture was diluted with sat. $NaHCO_3$ *aq* and then extracted with $CHCl_3$. The combined organic layer was washed with brine and dried over $MgSO_4$. Filtration and concentration *in vacuo* gave **9-1** (597 mg, 62%) as a light yellow solid. ^{1}H-NMR (400 MHz, $CDCl_3$: CD_3OD = 1 : 1) δ 7.75 (s, 1H), 7.72 (s, 2H), 7.60 (s, 1H), 6.96 (s, 1H), 4.51 (s, 1H), 3.57 (s, 1H).

Compound **9-2**

To a stirred solution of 5-hydroxy-2-nitrobenzaldehyde (1000 mg, 6.0 mmol), KI (166 mg, 1.0 mmol) and K_2CO_3 (1244mg, 9.0 mmol) in DMF (20 mL) was added 2-(2-chloroethoxy)ethanol (1246 mg, 10 mmol). The reaction mixture was stirred at 50 °C for 24h. The mixture was extracted with AcOEt. The combined organic layer was washed with brine and dried over $MgSO_4$. Filtration, concentration *in vacuo* and purification by silica gel flash column chromatography ($CHCl_3$) gave 1409 mg (92 %) of **9-2** as a white solid. ^{1}H-NMR (400 MHz, $CDCl_3$) δ 10.48 (s, 1H), 8.17 (dd, J = 1.6, 9.2 Hz, 1H), 7.38 (d, J = 2.8 Hz, 1H), 7.18 (dd, J = 2.8, 9.2 Hz, 1H), 4.30 (t, J = 4.4 Hz, 1H), 3.91(t, J = 4.8 Hz, 1H), 3.78 (t, J = 4.4 Hz, 1H), 3.67 (t, J = 4.0 Hz, 2H).

Compound **9-3**

A solution of **9-2** (1400 mg, 5.5 mmol) in DCM (20 mL) is treated at 0 °C with TEA (1.5 mL) and methanesulfonyl chloride (387 μL, 5.5 mmol). The reaction mixture is stirred at 0 °C for 1h. The mixture was partitioned between $CHCl_3$ and water. The collected organic layer was washed with brine, collected, dried over $MgSO_4$. Filtration, and concentration *in vacuo* gave mesylated **9-2** as a yellow solid. Without further purification, mesylated **9-2** was used in next step.

To a solution of mesylated **9-2** in DMF (20 mL) was added NaN_3 (650 mg, 10 mmol) and K_2CO_3 (1382 mg, 10 mmol). The mixture was stirred for 8h at 50 °C. The reaction mixture was poured into water and the whole was extracted with AcOEt. The AcOEt layer was separated, washed with brine, and was dried over $MgSO_4$. Filtration, concentration *in vacuo*, purification by silica gel flash column chromatography ($CHCl_3$: AcOEt = 10 : 1) gave 1314 mg (85%) of **9-3** as a white solid. ^{1}H-NMR (400 MHz, $CDCl_3$) δ 10.49 (s, 1H), 8.15 (d, J = 9.2 Hz, 1H), 7.36 (d, J = 2.8 Hz, 1H), 7.19 (dd, J = 3.2, 9.3 Hz, 1H), 4.29 (t, J = 4.4 Hz, 1H), 3.91 (t, J = 4.4 Hz, 1H), 3.74 (t, J = 4.8 Hz, 1H), 3.67 (t, J = 5.2 Hz, 2H).

Compound **9-4**

To a stirred suspension of **9-3** (1314 mg, 4.69 mmol) in THF (15 mL) was added $NaBH_4$ (190 mg, 5.0 mmol). The reaction mixture was stirred at rt for 30 min. The solvent was removed by evaporation. The residue was diluted with water, and then

extracted with AcOEt. The combined organic layer was washed with sat. $NaHCO_3$ *aq* and brine and dried over $MgSO_4$. Filtration, and concentration *in vacuo* gave **9-4** (1324 mg, 100%) as a white solid. ^{1}H-NMR (400 MHz, $CDCl_3$) δ 8.18 (d, *J* = 8.8 Hz, 1H), 7.26 (overlapped with $CHCl_3$, 1H), 6.92 (dd, *J* = 2.8, 9.2 Hz, 1H), 5.00 (s, 2H), 4.27 (t, *J* = 4.4 Hz, 1H), 3.91 (t, *J* = 4.4 Hz, 1H), 3.75 (t, *J* = 4.4 Hz, 1H), 3.42 (t, *J* = 4.8 Hz, 2H).

Compound **9-5**

To a stirred suspension of **9-4** (779 mg, 2.76 mmol) in THF/H_2O (10 mL /1 mL) was added PPh_3 (796 mg, 3.04 mmol). The reaction mixture was stirred at rt for 24h. The solvent was removed by evaporation. The residue was purified by silica gel flash column chromatography ($CHCl_3$ saturated by NH_3*aq* : MeOH = 4 : 1) gave 587 mg (83 %) of **9-5** as a white solid. ^{1}H-NMR (400 MHz, $CDCl_3$) δ 8.16 (d, *J* = 9.2 Hz, 1H), 7.39 (d, *J* = 2.8 Hz, 1H), 6.90 (dd, *J* = 3.2, 9.2 Hz, 1H), 5.01 (s, 2H), 4.30 (t, *J* = 4.8 Hz, 1H), 3.86 (t, *J* = 4.8 Hz, 1H), 3.86 (t, *J* = 4.8 Hz, 1H), 2.91 (t, *J* = 4.4 Hz, 2H).

Compound **9-6**

To a stirred solution of **9-5** (254 mg, 1.0 mmol) and DIPEA (348 μL) in DMF (2 mL) and DCM (10mL) was added **1-4** (298 mg, 1.0 mmol). The reaction mixture was stirred at rt for 2 hours. The mixture was diluted with sat. $NaHCO_3$ *aq* and then extracted with AcOEt. The combined organic layer was washed with brine and dried over $MgSO_4$. Filtration, concentration *in vacuo* and purification by silica gel flash column chromatography ($CHCl_3$: MeOH = 8 : 1) gave 369 mg (84 %) of **9-6** as a white solid. ^{1}H-NMR (400 MHz, CD_3OD) δ 8.12 (dd, *J* = 3.6, 9.2 Hz, 1H), 7.94 – 7.87 (m, 4H), 7.48 (m, 1H), 6.88 (d, *J* = 8.8 Hz, 1H), 4.99 (s, 2H), 4.29 (m, 2H), 3.91 (m, 1H), 3.75 (m, 1H), 3.63 (m, 2H).

Compound **9-7**

To a stirred solution of **9-6** (50 mg, 114 μmol) and DIPEA (75 μL) in DMF (3 mL) was added *N,N′*-disuccinimidyl carbonate (143 mg, 563 μmol). The reaction mixture was stirred at rt for 5 hours. The mixture was diluted with sat. $NaHCO_3$ *aq* and then extracted with AcOEt. The combined organic layer was washed with brine and dried over $MgSO_4$. Filtration, concentration *in vacuo* gave 35.1 mg (53 %) of **9-7** as a white

solid. ^{1}H-NMR (400 MHz, CD_3OD) δ 8.20 (d, *J* = 9.6 Hz, 1H), 7.97 – 7.85 (m, 4H), 7.45 (m, 1H), 7.16 (m, 1H), 5.72 (s, 2H), 4.28 (m, 1H), 3.80 (m, 1H), 3.62 (m, 1H), 3.46 (m, 1H), 2.80 (s, 4H).

Compound **9**

To a stirred solution of **9-7** (30.2 mg, 52 μmol) and DIPEA (32 μL) in DMF (2 mL) was added **9-1** (18.6 mg, 53 μmol). The reaction mixture was stirred at rt for 2 hours. The mixture was diluted with sat. $NaHCO_3$ *aq*, and then extracted with AcOEt. The combined organic layer was washed with brine and dried over $MgSO_4$. Filtration, concentration *in vacuo* and purification by silica gel flash column chromatography ($CHCl_3$: MeOH = 1 : 0 to 10 : 1) gave 21.4 mg (50 %) of **9** as a white solid. ^{1}H-NMR (400 MHz, $CDCl_3$: CD_3OD = 1 : 1) δ 8.23 (m, 1H), 8.21 (s, 1H), 7.95 – 7.88 (m, 4H), 7.79 (s, 1H), 7.77 (s, 1H), 7.58 (d, *J* = 2.8 Hz, 1H), 7.49 (s, 1H), 7.05 (m, 1H), 5.83 (s, 2H), 4.53 (s, 2H), 4.29 (m, 1H), 3.91 (m, 1H), 3.63 (m, 1H). HR-ESI MS *m/z* calcd for $[M+H]^+$ 817.1721, found 817.1710.

Peptide mapping of the Dc labeled CAI

A solution of CAI (30 μM) and **1** (60 μM) in HEPES buffer (50 mM, pH 7.2) was incubated for 12 h at 37 °C. The labeled CAI was purified by size-exclusion chromatography (TOYOPEARL HW-40F) and dialyzed against Tris buffer (100 mM, pH 8.0) using a Spectra/Por® dialysis membrane (MWCO: 10,000). After concentration by ultrafiltration (Centricon Ultracel YM-10, Millipore), the labeling yield of CAI (63%) was determined based on the relative absorbance at 280 nm and 430 nm. For protein digestion, the solution was incubated with lysyl endopeptidase (LEP) (LEP/substrate ratio = 1/3 (w/w)) at 37 °C for 4 h in the presence of urea (2 M). Native (unlabeled) CAI was also subjected to LEP digestion under the same conditions. The digested sample was subjected to RP-HPLC (column; YMC-pack ODS-A, 250 × 4.6 mm, mobile phase; CH_3CN (containing 0.1% TFA) : H_2O (containing 0.1% TFA) = 5 : 95 to 55 : 45 (linear gradient over 100 min), flow rate; 1.0 mL/min, detection; UV (220 nm) and fluorescence (excitation at 427 nm, emission at 473 nm)). For the second digestion, the fraction containing the labeled L6+L7 was lyophilized, redissolved in 200 mL Tris buffer (100 mM, pH = 8.0), and the solution was digested with thermolysin (1 mg) for 3 h at 65 °C. All of the purified digested fragments were analyzed by MALDI-TOF mass. For the peptide mapping, the labeled fragments were further analyzed by MALDI-TOF MS/MS.

Peptide mapping of the F6B labeled CAI

A solution of CAI (10 μM) and **8** (20 μM) in HEPES buffer (50 mM, pH 7.2) was incubated for 8 h at 37 °C. The labeled CAI was purified by size-exclusion chromatography (TOYOPEARL HW-40F). After concentration by ultrafiltration (Centricon Ultracel YM-10, Millipore), the labeling yield of CAI (100%, 17 μM) was determined based on the relative absorbance at 280 nm and ^{19}F-NMR. For protein digestion, the solution was incubated with lysyl endopeptidase (LEP) (LEP/substrate ratio = 1/5 (w/w)) at 37 °C for 4.5 h in the presence of urea (2 M). Native (unlabeled) CAI was also subjected to LEP digestion under the same conditions. The digested sample was subjected to RP-HPLC (column; YMC-pack ODS-A, 250 × 4.6 mm, mobile phase; CH_3CN (containing 0.1% TFA) : H_2O (containing 0.1% TFA) = 5 : 95 to 55 : 45 (linear gradient over 100 min), flow rate; 1.0 mL/min, detection; UV (220 nm). All of

the purified digested fragments were analyzed by MALDI-TOF mass. For the peptide mapping, the labeled fragments were further analyzed by MALDI-TOF MS/MS.

References

S1) S. Tsukiji, M. Miyagawa, Y. Takaoka, T. Tamura, I. Hamachi, *Nat. Chem. Biol.,* 2009, **5**, 341.

S2) S. Fujishima, R. Yasui, T. Miki, A. Ojida, I. Hamachi, *J. Am. Chem. Soc.,* 2012, **13***4*, 3961.