Supporting Information for

Dependence of Effective Molarity on Linker Length for an Intramolecular Protein-Ligand System

Vijay M. Krishnamurthy, Vincent Semetey, Paul J. Bracher, Nan Shen, and George M. Whitesides*

Department of Chemistry and Chemical Biology, Harvard University

12 Oxford Street, Cambridge, MA 02138

* Author to whom correspondence should be addressed.

Telephone: (617) 495-9430

Fax: (617) 495-9857

E-mail: gwhitesides@gmwgroup.harvard.edu

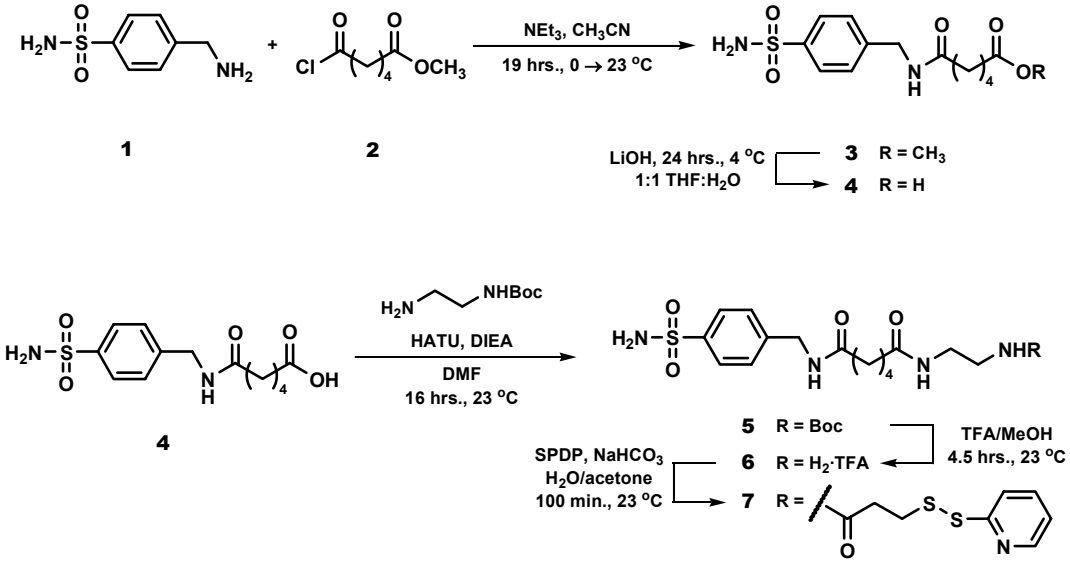
Supporting Experimental Procedures

General Considerations. Analytical HPLC was run on a Varian instrument with a C18 column (length = 250 mm, i.d. = 4.6 mm, 5 μm particle size) from Vydac using a linear gradient of water with 0.1% TFA (solvent *A*) followed by acetonitrile containing 0.08% TFA (solvent *B*), at a flow rate of 1.2 mL min^{-1} (UV detection at 214 and 254 nm). Mass spectra were obtained by matrix-assisted laser desorption ionization/time-of-flight mass spectrometry (MALDI-TOF) on a Perspective Biosystems Voyager-DE PRO (Framingham, MA) using α-cyano-4-hydroxycinnamic acid as a matrix. High resolution electrospray ionization mass spectra (ESI-MS) were obtained by using a Micromass LCT mass spectrometer (Beverly, MA) with a mass resolving power of 5,000 ($m/\Delta m$) by the Mass Spectrometry facility at the Harvard Department of Chemistry and Chemical Biology.

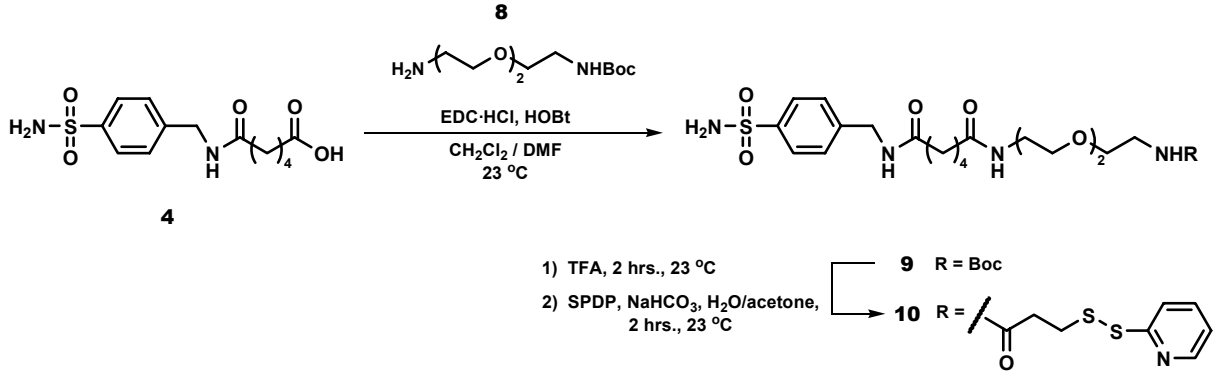
General Synthetic Considerations. Schemes S.1-S.5 show the syntheses of the molecules used in this study. The carboxylic acid **4** and mono-Boc-protected diamine **8** were prepared as previously described.[1, 2]

SA-OMe (3). A commercial sample of 4-(aminomethyl)benzenesulfonamide (2.0 g, 10.7 mmol) was suspended in 100 mL of CH_3CN in a 200-mL round-bottomed flask. Triethylamine (1.8 mL, 12.9 mmol) was added to the flask, which was then fitted with a dropping funnel and lowered into an ice-water bath. A solution of methyl adipoyl chloride (**2**, 1.8 mL, 11.6 mmol) in 10 mL of CH_2Cl_2 was added dropwise over the course of 15 min.[3] The flask was allowed to warm to room temperature and the mixture was stirred for a total of 19 h. The solvent was removed by rotary evaporation. The residual solid was suspended in 60 mL of 0.5 M HCl for 10 min, and then filtered through a sintered glass frit. The white solid was washed with 100 mL of deionized water and pumped dry under vacuum. Yield: 2.66 g (8.10 mmol, 75%). White

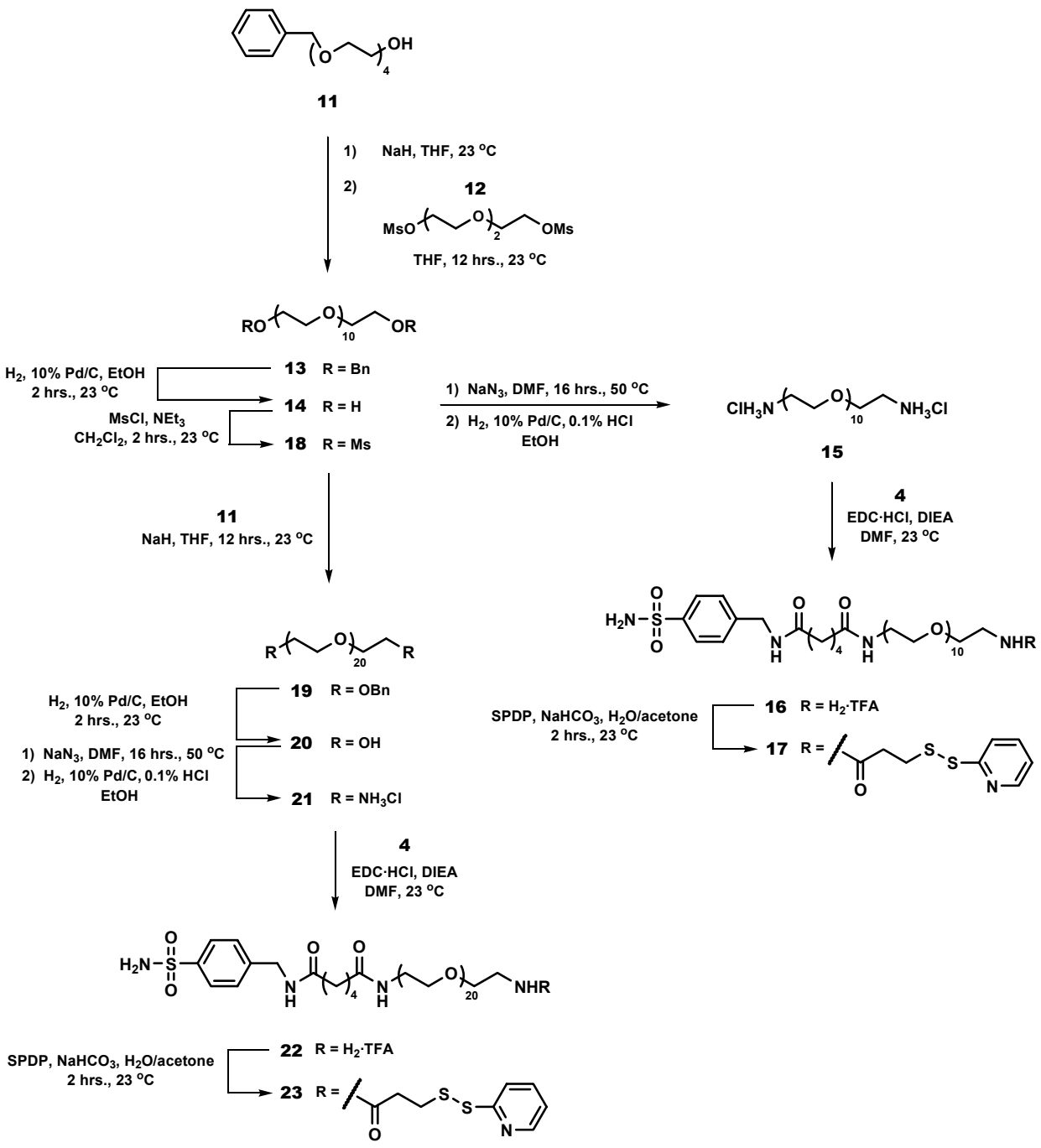
Scheme S.1. Synthesis of the SA-OMe (**3**), Pyr-SSEG$_0$SA (**7**), and their intermediates.



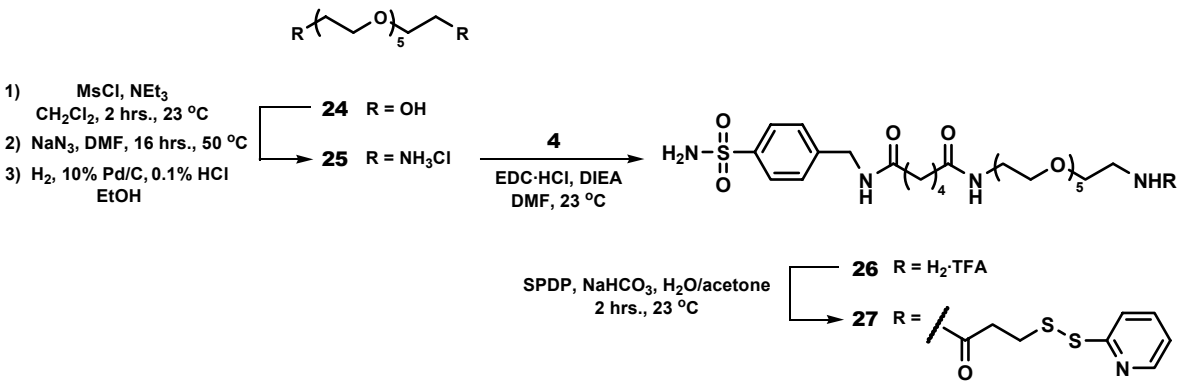
Scheme S.2. Synthesis of the Pyr-SSEG$_2$SA (**10**) and its intermediates.



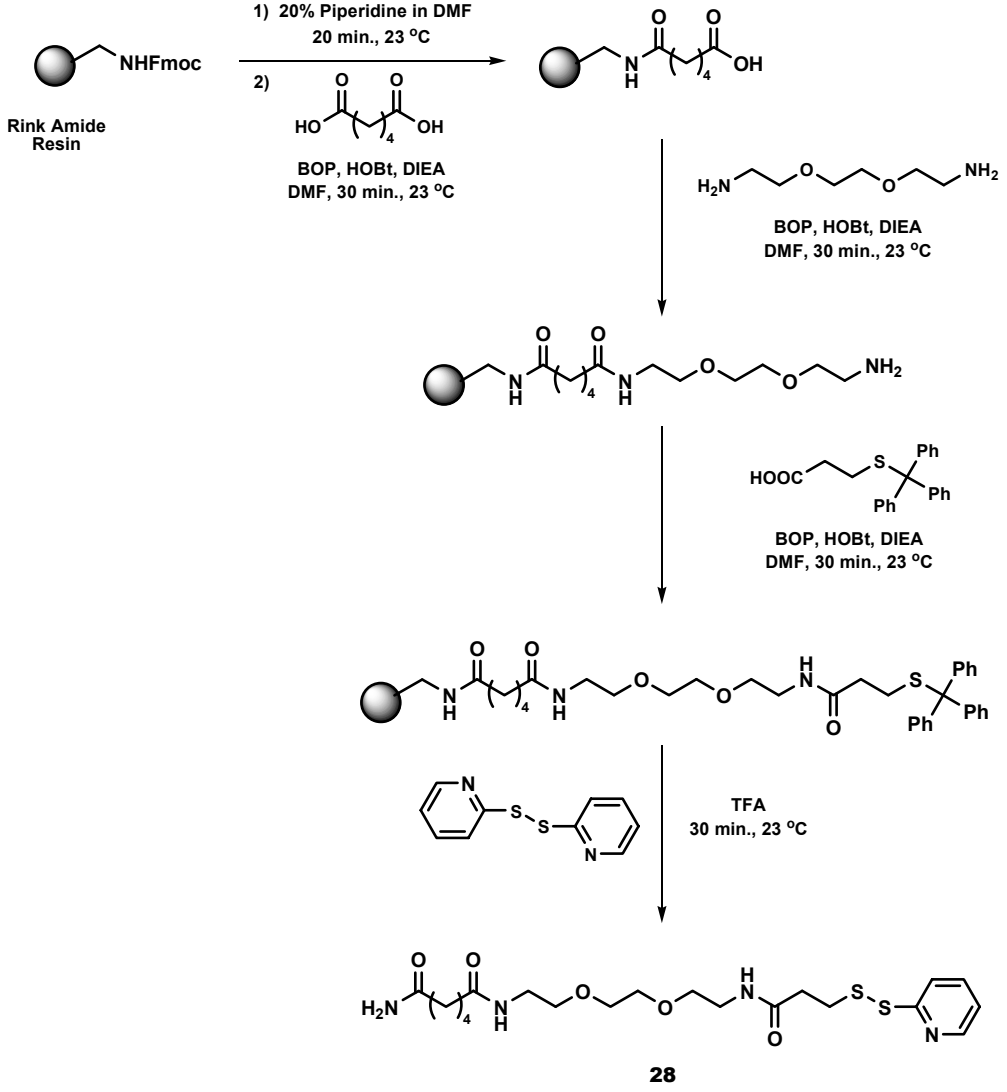
Scheme S.3. Synthesis of Pyr-SSEG$_{10}$SA (**17**), Pyr-SSEG$_{20}$SA (**23**), and their intermediates.



Scheme S.4. Synthesis of Pyr-SSEG$_5$SA (**27**) and its intermediates.



Scheme S.5. Synthesis of Pyr-SSEG$_2$CONH$_2$ (**28**).



powder. ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.39 (t, *J* = 6.0 Hz, 1H), 7.74 (d, *J* = 8.2 Hz, 2H), 7.38 (d, *J* = 8.5 Hz, 2H), 7.29 (s, 2H), 4.29 (d, *J* = 6.0 Hz, 2H), 3.57 (s, 3H), 2.30 (t, *J* = 7.0 Hz, 2H), 2.14 (t, *J* = 7.0 Hz, 2H), 1.51 (m, 4H). ^{13}C NMR (126 MHz, DMSO-d_6): δ 173.96, 172.70, 144.54, 143.23, 128.12, 126.35, 51.92, 42.32, 35.58, 33.68, 25.36, 24.76. High Resolution ESI-MS: 329.1168. Calculated for $C_{14}H_{21}N_2O_5S^+$ [M+H$^+$]: 329.1166.

Compound 4. Methyl ester **3** (2.5 g, 7.6 mmol) was suspended in 140 mL of THF in a 500-mL round-bottomed flask. In a separate flask, 2.5 g of lithium hydroxide was dissolved in 140 mL of deionized water. Both mixtures were chilled to 4 °C and combined to form a turbid white mixture. After 1 h of stirring, the mixture had become homogeneous. After 24 h, 50 mL of 3 M HCl was added, and the mixture was allowed to warm to room temperature. Following the addition of a 100-mL portion of saturated aqueous NaCl solution, the mixture was extracted four times with 100-mL portions of EtOAc, and the combined organic layers were evaporated. Yield: 2.28 g (7.25 mmol, 95%). White powder. ^{1}H NMR (400 MHz, DMSO-d_6): δ 11.99 (s, 1H), 8.39 (t, *J* = 5.9 Hz, 1H), 7.74 (d, *J* = 8.5 Hz, 2H), 7.38 (d, *J* = 8.2 Hz, 2H), 7.28 (s, 2H), 4.29 (d, *J* = 6.0 Hz, 2H), 2.20 (t, *J* = 6.9 Hz, 2H), 2.14 (t, *J* = 7.0 Hz, 2H), 1.49 (m, 4H). ^{13}C NMR (126 MHz, DMSO-d_6): δ 175.09, 172.76, 144.55, 143.22, 128.11, 126.36, 42.32, 35.68, 34.07, 25.47, 24.84. High Resolution ESI-MS: 315.1005. Calculated for $C_{13}H_{19}N_2O_5S^+$ [M+H$^+$]: 315.1009.

Compound 5. Carboxylic acid **4** (500 mg, 1.59 mmol) was dissolved in 15 mL of dry DMF with stirring. *O*-(7-Azabenzotriazole-1-yl)-*N*,*N*,*N*,*N*'-tetramethyluronium hexafluoro-phosphate (HATU, 665 mg, 1.75 mmol) was added as a solid and the resulting clear solution was stirred for 10 min at room temperature.[4] *N*-Boc-ethylenediamine (0.5 mL, 3.2 mmol) was injected and the resulting yellow solution was stirred for 20 min before 0.8 mL (4.6 mmol) of

diisopropylethylamine (DIEA) was added by syringe. The mixture was stirred for 16 h at room temperature, at which point 150 mL of a saturated solution of NaCl was added. The mixture was cooled to 4 °C and the white precipitate was isolated by vacuum filtration over sintered glass and washed with 100 mL of deionized water. Yield: 695 mg (1.5 mmol, 96%). White powder. ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.44 (t, J = 6.0, 1H), 7.83 (t, J = 4.9, 1H), 7.74 (d, J = 8.0, 2H), 7.38 (d, J = 8.0, 2H), 7.30 (s, 2H), 6.79 (t, J = 5.3, 1H), 4.28 (d, J = 5.9, 2H), 3.02 (m, 2H), 2.93 (m, 2H), 2.12 (t, J = 6.6, 2H), 2.03 (t, J = 6.6, 2H), 1.46 (m, 4H), 1.35 (s, 9H). ^{13}C NMR(101 MHz, DMSO-d_6): δ 172.84, 172.79, 156.29, 144.55, 143.22, 128.12, 126.35, 78.31, 42.32, 40.78, 39.33, 35.84, 35.78, 28.92, 25.62 (the signals from the two central adipoyl methylene carbons appear to be accidentally equivalent). High Resolution ESI-MS: 457.2118. Calculated for $C_{20}H_{33}N_4O_6S^+$ [M+H$^+$]: 457.2115.

Compound 6. The Boc-protected amine **5** (486 mg, 1.1 mmol) was suspended in 10 mL of trifluoroacetic acid with vigorous stirring at room temperature. After 3 h, 5 mL of methanol was added and the solution stirred for an additional 1.5 h. The solvent was removed by rotary evaporation under high vacuum. The remaining solid was redissolved in 10 mL of methanol, evaporated, and placed on high vacuum overnight to remove the last vestiges of excess trifluoroacetic acid. Yield: 565 mg (100% based on loss of *t*-butyl signal in NMR). Clear oil with a faint brown hue. ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.40 (t, J = 5.9, 1H), 7.97 (t, J = 5.8, 1H), 7.77 (s, br, 3H), 7.74 (d, J = 8.2, 2H), 7.38 (d, J = 8.5, 2H), 7.29 (s, 2H), 4.29 (d, J = 6.0, 2H), 3.25 (dd, J = 12.1, 6.3, 2H), 2.82 (dd, J = 12.2, 6.2, 2H), 2.14 (t, J = 6.9, 2H), 2.08 (t, J = 6.9, 2H), 1.48 (m, 4H). ^{13}C NMR (101 MHz, DMSO-d_6): δ 173.57, 172.84, 155.88 (q, J = 34.3), 144.53, 143.22, 128.12, 126.35, 116.81 (q, J = 294.5), 42.33, 39.40, 37.07, 35.79, 35.76, 25.60, 25.41. High Resolution ESI-MS: 357.1595. Calculated for $C_{15}H_{25}N_4O_4S^+$ [M+H]$^+$: 357.1591.

Pyr-SSEG$_0$SA (7). The full quantity of the trifluoroacetate salt of amine **6** (565 mg) was dissolved in 8 mL of 250 mM $NaHCO_3$. A 5-mL solution of *N*-succinimidyl-3-(2-pyridyldithio)-propionate (SPDP, 0.38 mmol) in acetone was added dropwise with rapid stirring. A precipitate formed within 10 min and the solution was stirred at room temperature for a total of 100 min. The precipitate was isolated by vacuum filtration over a frit of sintered glass and washed with 25 mL of deionized water (76 mg, 0.14 mmol). A second batch of product (20 mg, 0.036 mmol) was isolated by allowing the acetone from the filtrate to evaporate and filtering the subsequent precipitate in the same manner as before. Combined yield: 96 mg (0.17 mmol, 45%). White powder. ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.44 (m, 1H), 8.39 (t, J = 5.8, 1H), 7.99 (m, 1H), 7.85-7.71 (m, 3H), 7.74 (d, J = 8.4, 2H), 7.38 (d, J = 8.4, 2H), 7.30 (s, 2H), 7.23 (ddd, J = 7.3, 4.8, 1.1, 1H), 4.29 (d, J = 5.8, 2H), 3.06 (m, 4H), 2.99 (t, J = 7.1, 2H), 2.47 (t, J = 7.0, 2H), 2.12 (t, J = 6.6, 2H), 2.03 (t, J = 6.6, 2H), 1.46 (m, 4H). ^{13}C NMR (101 MHz, DMSO-d_6): δ 172.85, 172.78, 170.63, 159.86, 150.28, 144.53, 143.22, 138.51, 128.12, 126.36, 121.86, 119.81, 42.34, 39.16, 38.89, 35.89, 35.80, 35.34, 34.65, 25.61 (the signals from the two central adipoyl methylene carbons appear to be accidentally equivalent). High Resolution ESI-MS: 554.1554. Calculated for $C_{23}H_{32}N_5O_5S_3^+$ $[M+H]^+$: 554.1560.

Compound 9. Carboxylic acid **4** (0.98 g, 3.1 mmol) and Boc-protected amine **8** (0.77 g, 3.1 mmol) were dissolved in 40 mL of methylene chloride containing 5 mL of DMF. Hydroxybenzotriazole (HOBt, 0.43 g, 3.1 mmol) and 1-Ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride (EDC·HCl, 0.595 g, 3.1 mmol) were added with stirring and the reaction was allowed to proceed at room temperature. The solution was evaporated under reduced pressure, and the residue was taken up in 100 mL of water, then extracted four times with 100-mL portions of ethyl acetate. The combined organic phases were

washed sequentially with saturated solutions of ammonium chloride, sodium bicarbonate, and sodium chloride, then dried over magnesium sulfate and evaporated under reduced pressure. Purification by flash chromatography using 2:1 acetone/ethyl acetate as the eluent gave 0.953 g (57%) of a white powder. RP-HPLC t_R = 12.57 min (linear gradient, 0-100% *B*, 20 min). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.38 (t, *J* = 6.0, 1H), 7.82 (t, *J* = 5.8, 1H), 7.74 (d, *J* = 8.2, 2H), 7.39 (d, *J* = 8.1, 2H), 7.29 (s, 2H), 6.75 (t, *J* = 6.0, 1H), 4.30 (d, *J* = 5.9, 2H), 3.48 (s, 4H), 3.37 (m, 4H), 3.18 (m, 2H), 3.04 (m, 2H), 2.13 (t, *J* = 6.8, 2H), 2.06 (t, *J* = 6.8, 2H), 1.47 (m, 4H), 1.36 (s, 9H). ^{13}C NMR (101 MHz, DMSO-d_6): δ 172.12, 172.02, 155.56, 143.84, 142.48, 127.40, 125.65, 77.56, 69.50, 69.42, 69.14, 41.62, 38.42, 35.12, 35.07, 28.20, 24.97, 24.94. MS (MALDI-TOF) *m/z*: 567.82 and 583.85. Calculated for $C_{24}H_{40}N_4NaO_8S^+$ $[M+Na]^+$: 567.65 and for $C_{24}H_{40}KN_4O_8S^+$ $[M+K]^+$: 583.76.

Pyr-SSEG$_2$SA (10). Boc-protected amine **9** (0.42 g, 0.93 mmol) was dissolved in 8 mL of trifluoroacetic acid and stirred for 2 h at room temperature. Water (25 mL) was added, and the solution was lyophilized to give the deprotected amine. The crude amine was dissolved in 19 mL of 250-mM aqueous sodium bicarbonate. A solution of *N*-succinimidyl-3-(2-pyridyldithio)-propionate (SPDP, 0.29 g, 0.93 mmol) in 9 mL of acetone was added with stirring. The reaction was allowed to proceed for 2 h, then the acetone was removed under vacuum. The product precipitated from the aqueous solution as a sticky solid; sonication converted the solid to a powder. The powder was filtered, dried, and recrystallized from hot acetone to give 0.47 g (78 %) of a white solid. RP-HPLC t_R = 11.64 min (linear gradient, 0-100% *B*, 20 min). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.43 (m, 1H), 8.42 (t, *J* = 6.0, 1H), 8.02 (m, 2H), 7.80 (m, 2H), 7.72 (d, *J* = 8.2, 2H), 7.36 (d, *J* = 8.3, 2H), 7.27 (s, 2H), 7.20 (m, 1H), 4.27 (d, *J* = 5.9, 2H), 3.46 (s, 4H), 3.35 (m, 4H), 3.16 (m, 4H), 2.97 (t, *J* = 6.9, 2H), 2.49 (t, *J* = 7.0, 2H), 2.11 (t, *J* = 6.7, 2H), 2.03

(t, J = 6.8, 2H), 1.45 (m, 4H). ^{13}C NMR (101 MHz, DMSO-d_6): δ 172.13, 172.04, 169.86, 151.48, 143.84, 142.52, 137.76, 127.36, 125.66, 121.18, 121.08, 119.11, 69.49, 69.14, 69.02, 41.61, 38.59, 38.41, 35.12, 35.07, 34.47, 34.04, 24.97. MS (MALDI-TOF) *m/z*: 664.80. Calculated for $C_{27}H_{39}N_5NaO_7S_3^+$ $[M+Na]^+$: 664.81.

$BnO(CH_2CH_2O)_{11}Bn$ (13). To a solution of tetra(ethylene glycol) monobenzyl ether **11** (6.16 g, 21.66 mmol) in THF (20 mL) were added portions of NaH (a total of 1.82 g, 43.32 mmol) as a 57% dispersion in oil. Next, a solution of **12** (3.06 g, 10 mmol) in THF (10 mL) was added to the mixture. The mixture was stirred for 12 h at room temperature and then evaporated to dryness. The residue was chromatographed (SiO_2; EtOAc → 10% MeOH in EtOAc) to yield **13** (2.2 g, 3.22 mmol, 32%) as an oil. RP-HPLC t_R = 17.80 min (linear gradient, 0-100% *B*, 20 min). ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.25-7.38 (m, 10H), 4.56 (s, 4H), 3.72-3.60 (m, 44H). MS (MALDI-TOF) *m/z*: 705.10 and 721.15. Calculated for $C_{36}H_{58}NaO_{12}^+$ $[M+Na]^+$: 705.84 and for $C_{36}H_{58}KO_{12}^+$ $[M+K]^+$: 721.95.

$HO(CH_2CH_2O)_{11}H$ (14). The compound **13** (2.2 g, 3.22 mmol) and 10% Pd/C (250 mg) were combined in ethanol (100 mL). This mixture was hydrogenated at 23 °C for 2 h. The mixture was then filtered through Celite, and the filter cake washed with 2 x 10 mL EtOH. The resulting solution was evaporated to yield **14** (1.44 g, 2.86 mmol, 89%) as an oil. ^{1}H NMR (500 MHz, $CDCl_3$): δ 3.67-3.49 (m, 44H), 3.32 (s, 2H). ^{13}C NMR (126 MHz, $CDCl_3$): δ 72.94, 70.73, 70.70, 70.65, 70.59, 70.31, 61.67. MS (MALDI-TOF) *m/z*: 525.10 and 540.89. Calculated for $C_{22}H_{46}NaO_{12}^+$ $[M+Na]^+$: 525.59 and for $C_{22}H_{46}KO_{12}^+$ $[M+K]^+$: 541.70.

$BnO(CH_2CH_2O)_{21}Bn$ (19). The oligoethyleneglycol **14** (598 mg, 1.19 mmol) in CH_2Cl_2 (10 mL) was first converted to its dimesylate (**18**) by adding triethylamine (365 µL, 2.62 mmol) and mesyl chloride (207 µL, 2.62 mmol) dropwise. The solution was stirred for 2 h and then

washed with water (3 x 50 mL), evaporated, and dried. To a solution of penta(ethylene glycol) monobenzyl ether **11** (861 mg, 2.62 mmol) in THF (20 mL) was added NaH, in portions (111 mg, 2.75 mmol, total), as a 57% dispersion in oil. Next, a solution of the dimesylate **18** in THF (10 mL) was added to the mixture. The mixture was stirred for 12 h at room temperature and then evaporated to dryness. The residue was purified by HPLC (30-100%) to yield **19** (724 mg, 0.64 mmol, 54%) as an oil. RP-HPLC t_R = 13.92 min (linear gradient, 0-100% *B*, 20 min). ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.38-7.24 (m, 10H), 4.56 (s, 4H), 3.73-3.61 (m, 84H). MS (MALDI-TOF) *m/z*: 1146.4 and 1162.7 $[M+K]^+$. Calculated for $C_{56}H_{98}NaO_{22}^+$ $[M+Na]^+$: 1146.38 and for $C_{56}H_{98}KO_{22}^+$ $[M+K]^+$: 1162.49.

$HO(CH_2CH_2O)_{21}H$ (20). The compound **19** (674 mg, 0.6 mmol) and 10% Pd/C (50 mg) were combined in ethanol (40 mL). This mixture was hydrogenated at 23 °C for 2 h. The mixture was then filtered through Celite, the filter cake was washed with 2 x 10 mL ethanol, and the resulting solution was evaporated to yield **20** (560 mg, 0.59 mmol, 99%). ^{1}H NMR (500 MHz, CD_3OD): δ 3.80-3.55 (m, 84H). ^{13}C NMR (126 MHz, $CDCl_3$): δ 72.74, 70.75, 70.72, 70.48, 61.84. MS (MALDI-TOF) *m/z*: 965.64 and 981.68. Calculated for $C_{42}H_{86}NaO_{22}^+$ $[M+Na]^+$: 966.12 and for $C_{42}H_{86}KO_{22}^+$ $[M+K]^+$: 982.23.

General Procedure for the Conversion of Alcohols to Amines (14 to 15, 20 to 21, and 24 to 25). Our basic procedure for converting alcohols to amines was adapted from an example in the literature.[5] Oligoethylene glycols were converted to their mesylates (mesyl chloride, triethylamine, dichloromethane), and the mesylates were displaced with sodium azide (DMF, overnight, 50ºC) and then hydrogenated (EtOH with 0.1% HCl, Pd/C, H_2).

$H_2N(CH_2CH_2O)_5CH_2CH_2NH_2$•2HCl (25). ^{1}H NMR (500 MHz, $CDCl_3$): δ 3.80-3.48 (m, 24H). ^{13}C NMR (126 MHz, $CDCl_3$): δ 70.84, 70.81, 70.76, 70.49, 70.39, 70.28, 70.22, 70.04,

67.05, 50.85. MS (MALDI-TOF) *m/z*: 281.94. Calculated for $C_{12}H_{29}N_2O_5^+$ $[M+H]^+$: 281.37.

$H_2N(CH_2CH_2O)_{10}CH_2CH_2NH_2$•2HCl (15). ^{1}H NMR (500 MHz, CD_3OD): δ 3.76 (t, *J* = 5.1 Hz, 4H), 3.73-3.63 (m, 36H), 3.16 (t, *J* = 5.1 Hz, 4H). ^{13}C NMR (126 MHz, $CDCl_3$): δ 70.38, 70.30, 70.28, 70.26, 70.20, 69.95, 66.93, 39.99. MS (MALDI-TOF) *m/z*: 502.01. Calculated for $C_{22}H_{49}N_2O_{10}^+$ $[M+H]^+$: 501.64.

$H_2N(CH_2CH_2O)_{20}CH_2CH_2NH_2$ (21). ^{1}H NMR (500 MHz, CD_3OD): δ 3.55-3.81 (m, 80H), 3.22-3.16 (m, 4H). ^{13}C NMR (126 MHz, $CDCl_3$): δ 70.38, 70.30, 70.28, 70.26, 70.20, 69.95, 66.93. MS (MALDI-TOF) *m/z*: 942.54. Calculated for $C_{42}H_{89}N_2O_{20}^+$ $[M+H]^+$: 942.17.

Compound 16. To a solution of the acid **4** (10 mg, 32 μmol), 1-Ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride (EDC·HCl, 6 mg, 32 μmol) and *N*,*N*-diisopropylethylamine (DIEA, 7 μL, 64 μmol) in 5 mL of DMF, was added **15** (64 mg, 127 μmol) with stirring and the reaction was allowed to proceed at room temperature. The crude product was purified by HPLC (linear gradient, 0-80% *B*, 40 min) and lyophilized to afford **16** (10 mg, 11 μmol, 35%) as a clear oil. RP-HPLC t_R = 11.52 min (linear gradient, 0-100% *B*, 20 min). ^{1}H NMR (500 MHz, CD_3OD): δ 7.85 (d, *J* = 8.3 Hz, 2H), 7.44 (d, *J* = 8.3 Hz, 2H), 4.43 (s, 2H), 3.80-3.74 (m, 2H), 3.73-3.57 (m, 36H), 3.56-3.50 (m, 2H), 3.37-3.32 (m, 2H), 3.22-3.16 (m, 2H), 2.28 (t, *J* = 6.8 Hz, 2H), 2.21 (t, *J* = 6.8 Hz, 2H), 1.71-1.57 (m, 4H). MS (MALDI-TOF) *m/z*: 798.32. Calculated for $C_{35}H_{65}N_4O_{14}S^+$ $[M+H]^+$: 797.98.

Pyr-SSEG_{10}SA (17). The compound **16** (9 mg, 11.3 μmol) was dissolved in 230 μL of 250 mM aqueous sodium bicarbonate. A solution of *N*-succinimidyl-3-(2-pyridyldithio)-propionate (SPDP, 3.5 mg, 11.3 μmol) in 109 μL of acetone was added with stirring. The reaction was allowed to proceed for 2 h, at which point the acetone was removed under vacuum. The aqueous solution was neutralized and purified by RP-HPLC (linear gradient, 0-100% *B*, 40

min) to yield **17** (4 mg, 4 μmol, 36%) as an oil. RP-HPLC t_R = 14.54 min (linear gradient, 0-100% *B*, 20 min). ^{1}H NMR (500 MHz, CD_3OD): δ 8.42 (d, *J* = 4.9 Hz, 1H), 7.88-7.83 (m, 4H), 7.45 (d, *J* = 8.3 Hz, 2H), 7.29-7.23 (m, 1H), 4.43 (s, 2H), 3.72-3.59 (m, 36H), 3.55-3.49 (m, 4H), 3.37-3.32 (m, 4H), 3.07 (t, *J* = 6.8 Hz, 2H), 2.63 (t, *J* = 6.8 Hz, 2H), 2.28 (t, *J* = 6.8 Hz, 2H), 2.22 (t, *J* = 6.8 Hz, 2H), 1.70-1.58 (m, 4H). MS (MALDI-TOF) *m/z*: 994.22 and 1016.04. Calculated for $C_{43}H_{72}N_5O_{15}S_3^+$ $[M+H]^+$: 995.25 and for $C_{43}H_{71}N_5NaO_{15}S_3^+$ $[M+Na]^+$: 1017.19.

Compound 22. To a solution of the acid **4** (6 mg, 19 μmol), 1-Ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride (EDC·HCl, 4 mg, 19 μmol) and *N*,*N*-diisopropylethylamine (DIEA, 2 μL, 19 μmol) in 5 mL of DMF, were added **21** (72 mg, 76 μmol) with stirring and the reaction was allowed to proceed at room temperature. The crude product was purified by HPLC (linear gradient, 0-80% *B*, 40 min) and lyophilized to afford **22** (8 mg, 6 μmol, 31%) as a clear oil. RP-HPLC t_R = 16.68 min (linear gradient, 0-100% *B*, 20 min). ^{1}H NMR (500 MHz, CD_3OD): δ 7.85 (d, *J* = 8.3 Hz, 2H), 7.44 (d, *J* = 8.3 Hz, 2H), 4.43 (s, 2H), 3.80-3.74 (m, 2H), 3.73-3.57 (m, 76H), 3.55-3.49 (m, 2H), 3.37-3.32 (m, 2H), 3.20-3.12 (m, 2H), 2.28 (t, *J* = 6.8 Hz, 2H), 2.22 (t, *J* = 6.8 Hz, 2H), 1.69-1.59 (m, 4H). MS (MALDI-TOF) *m/z*: 1238.78. Calculated for $C_{55}H_{105}N_4O_{24}S^+$ $[M+H]^+$: 1238.51.

Pyr-SSEG$_{20}$SA (23). Amine **22** (6 mg, 4.5 μmol) was dissolved in 92 μL of 250 mM aqueous sodium bicarbonate. A solution of *N*-succinimidyl-3-(2-pyridyldithio)-propionate (SPDP, 1.4 mg, 4.5 μmol) in 45 μL of acetone was added with stirring. The reaction was allowed to proceed for 2 h, at which point the acetone was removed under vacuum and the aqueous solution was neutralized. The product was purified by RP-HPLC (linear gradient, 0-100% *B*, 40 min) to yield **23** (2 mg, 1.4 μmol, 31%) as an oil. RP-HPLC t_R = 18.72 min (linear gradient, 0-100% *B*, 20 min). ^{1}H NMR (500 MHz, CD_3OD): δ 8.45-8.41 (m, 1H), 7.90-7.81 (m,

4H), 7.47 (d, J = 8.3 Hz, 2H), 7.28-7.23 (m, 1H), 4.45 (s, 2H), 3.70-3.60 (m, 76H), 3.57-3.51 (m, 4H), 3.41-3.36 (m, 4H), 3.08 (t, J = 6.8 Hz, 2H), 2.65 (t, J = 6.8 Hz, 2H), 2.30 (t, J = 6.8 Hz, 2H), 2.24 (t, J = 6.8 Hz, 2H), 1.72-1.62 (m, 4H). MS (MALDI-TOF) *m/z*: 1455.56 and 1471.52. Calculated for $C_{63}H_{111}N_5NaO_{25}S_3^+$ $[M+Na]^+$: 1457.76 and $C_{63}H_{111}KN_5O_{25}S_3^+$ $[M+K]^+$: 1473.87.

Compound 26. To a solution of acid **4** (23 mg, 73 mmol), 1-Ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride (EDC·HCl, 14 mg, 73 μmol), and *N*,*N*-diisopropylethylamine (DIEA, 13 μL, 146 μmol) in 5 mL of DMF was added **25** (82 mg, 292 μmol) with stirring, and the reaction was allowed to proceed at room temperature. The solution was evaporated to dryness. The crude product was purified by HPLC (linear gradient, 0-80% *B*, 40 min) and lyophilized to afford **26** (20 mg, 29 μmol, 40%) as a clear oil. RP-HPLC t_R = 8.93 min (linear gradient, 0-100% *B*, 20 min). ^{1}H NMR (500 MHz, CD_3OD): δ 7.87 (d, J = 8.3 Hz, 2H), 7.47 (d, J = 8.3 Hz, 2H), 4.45 (s, 2H), 3.75 (t, J = 5.4 Hz, 2H), 3.72-3.61 (m, 16H), 3.56 (t, J = 5.4 Hz, 2H), 3.38 (t, J = 4.9 Hz, 2H), 3.15 (t, J = 4.9 Hz, 2H), 2.33-2.28 (m, 2H), 2.26-2.20 (m, 2H), 1.71-1.59 (m, 4H). ^{13}C NMR (126 MHz, CD_3OD): δ 70.27, 70.22, 70.12, 70.06, 70.04, 69.96, 69.80, 69.47, 66.72, 42.50, 42.39, 39.45, 39.00, 35.55, 35.51, 25.36, 25.31. MS (MALDI-TOF) *m/z*: 577.38 and 599.44. Calculated for $C_{25}H_{45}N_4O_9S^+$ $[M+H]^+$: 577.72 and for $C_{25}H_{44}NaN_4O_9S^+$ $[M+Na]^+$: 599.69.

Pyr-SSEG$_{20}$SA (27). The amine **26** (9 mg, 15.6 μmol) was dissolved in 320 μL of 250 mM aqueous sodium bicarbonate. A solution of *N*-succinimidyl-3-(2-pyridyldithio)-propionate (SPDP, 5 mg, 15.6 μmol) in 150 μL of acetone was added with stirring. The reaction was allowed to proceed for 2 h. The acetone was removed under vacuum and the aqueous solution was neutralized and purified by RP-HPLC (linear gradient, 0-100% *B*, 40 min) to yield **27** (6 mg, 7.7 μmol, 50%) as an oil. RP-HPLC t_R = 10.89 min (linear gradient, 0-100% *B*, 20 min). ^{1}H

NMR (500 MHz, CD_3OD): δ 8.49-8.45 (m, 1H), 7.93-7.90 (m, 2H), 7.87 (d, *J* = 8.3, 2H), 7.46 (d, *J* = 8.3, 2H), 7.34-7.30 (m, 1H), 4.45 (s, 2H), 3.70-3.60 (m, 16H), 3.58-3.53 (m, 4H), 3.41-3.35 (m, 4H), 3.10 (t, *J* = 6.8 Hz, 2H), 2.65 (t, *J* = 6.8 Hz, 2H), 2.30 (t, *J* = 6.8 Hz, 2H), 2.23 (t, *J* = 6.8 Hz, 2H), 1.71-1.61 (m, 4H). MS (MALDI-TOF) *m/z*: 774.95 and 797.07. Calculated for $C_{33}H_{52}N_5O_{10}S_3^+$ $[M+H]^+$: 774.98 and for $C_{33}H_{51}N_5NaO_{10}S_3^+$ $[M+Na]^+$: 796.96.

Pyr-SSEG$_2$CONH$_2$ (28). Pyr-SSEG$_2$CONH$_2$ (**28**) was synthesized using a stepwise solid-phase methodology. Assembly of the molecule was carried out on a 50-μmol scale starting from Rink-amide MBHA resin.[6] The Fmoc group was removed using 20% piperidine in DMF (1 × 5 min, 1 × 15 min) under nitrogen bubbling. The resin was then filtered and washed with DMF (6 × 3 min). A solution of adipic acid (10 equiv.), BOP (10 equiv.), and HOBt (10 equiv.) in DMF and DIEA were added successively to the resin, and suspension was mixed for 30 min, followed by extensive washings with DMF. 2,2'-(Ethylenedioxy)diethylamine (10 equiv.) was coupled using BOP (10 equiv.), HOBt (10 equiv.), and DIEA for 30 min and the resin was washed with DMF, and then Trt$S(CH_2)_2CO_2H$ (5 equiv.) was coupled with BOP (5 equiv.), HOBt (5 equiv.), and DIEA (9 equiv.) in DMF for 30 min. The resin was washed with CH_2Cl_2 and Et_2O, then dried under nitrogen. Thiol deprotection/activation and cleavage from the resin was performed by treatment with a mixture of TFA (5 mL) and 2,2'-Dipyridyl disulfide (10 equiv.) for 30 min. After evaporation, purification by RP-HPLC (linear gradient, 0-80% *B*, 40 min) and lyophilization gave **28** (3 mg, 13%) as a white powder. RP-HPLC t_R = 12.76 min (linear gradient, 0-100% *B*, 20 min). ^{1}H NMR (400 MHz, CD_3OD): δ 8.45-8.42 (m, 1H), 7.91-7.87 (m, 2H), 7.32-7.27 (m, 1H), 3.59 (s, 4H), 3.55-3.49 (m, 4H), 3.39-3.31 (m, 4H), 3.07 (t, *J* = 6.8, 2H), 2.63 (t, *J* = 6.8, 2H), 2.20 (m, 4H), 1.66-1.55 (m, 4H). MS (MALDI-TOF) *m/z*: 473.9. Calculated for $C_{20}H_{33}N_4O_5S_2^+$ $[M+H]^+$: 473.62.

Purification of HCA-SSEG$_{10}$SA for Isothermal Titration Calorimetry.** Crude HCA**-SSEG$_{10}$SA (~5 mg) was incubated (with gentle shaking at 25 °C) with 200 μL of benzenesulfonamide-conjugated agarose (Sigma) in 2 mL of sodium phosphate buffer pH 7.5. After ~2 h, the supernatant was isolated by centrifugation (~200×g, 15 min). UV analysis of the supernatant revealed that ~10% of total protein was lost by this procedure. The sample was analyzed by fluorescence spectroscopy for the binding of DNSA (see next section) to determine the amount of unmodified HCA** present. HCA** was present as <1% of total protein.

Determination of Dissociation Constants for the Binding of DNSA to HCA.** To the wells of a black microwell plate were added DNSA (which was two-fold diluted across the wells of the plate) and HCA**, HCA**-SCH$_2$CO$_2^-$, or HCA**-SSEG$_2$CONH$_2$ (50–100 nM) in a final volume of 200 μL of 20 mM sodium phosphate pH 7.5. The plate was allowed to incubate at 25 °C for 1 h, and then its fluorescence was measured (excitation wavelength = 290 nm and emission wavelength = 460 nm, with a 455 nm cut-off filter). Wells were read ~30 times. Fluorescence intensities (F) were fit to eq S.1 (originally derived by Burton et al. in ref 18 of main text); this equation does not make the assumption that the concentration of free DNSA (that not bound to CA) is equal to that of total DNSA.

$$F = F_{\min} + (F_{\max} - F_{\min})\left(\frac{K_d + [\mathrm{CA}]_{\mathrm{total}} + [\mathrm{DNSA}]_{\mathrm{total}} - \sqrt{(K_d + [\mathrm{CA}]_{\mathrm{total}} + [\mathrm{DNSA}]_{\mathrm{total}})^2 - 4[\mathrm{CA}]_{\mathrm{total}}[\mathrm{DNSA}]_{\mathrm{total}}}}{2[\mathrm{CA}]_{\mathrm{total}}}\right) \quad \text{(S.1)}$$

$[\mathrm{CA}]_{\mathrm{total}}$ and $[\mathrm{DNSA}]_{\mathrm{total}}$ are the total concentrations of HCA** and DNSA, respectively. $F_{\min}$ is a constant to take into background fluorescence, $F_{\max}$ is the maximum fluorescence at total saturation of CA by DNSA, and K_d is the dissociation constant for the HCA-DNSA complex. $[\mathrm{CA}]_{\mathrm{total}}$ was contrained to its known value, and the other parameters were allowed to vary to

optimize the non-linear least squares fit (Origin). The value of F_{min} was very close to zero in all of the experiments.

Determination of Unmodified HCA Contaminating HCA**-SSEG$_n$SA.** To the wells of a black microwell plate were added 0-100 nM HCA** and 5 μM DNSA, in 200 μL of 20 mM sodium phosphate buffer pH 7.5. The fluorescence of the plate was measured under conditions in which only the HCA**-DNSA complex was fluorescent (excitation wavelength = 290 nm and emission wavelength = 460 nm, with a 455 nm cut-off filter). Figure S.4 shows these data with a linear fit. To determine the contamination by unmodified HCA** in the samples of HCA**-SSEG$_n$SA, HCA**-SSEG$_n$SA was added to the wells of the plate containing 5 μM DNSA, to a final concentration of 0.5 or 1.0 μM in 200 μL of 20 mM sodium phosphate buffer pH 7.5, and the fluorescence measured (as above). After subtraction of the fluorescence of the HCA**-SSEG$_n$SA protein alone (i.e., no DNSA added), this fluorescence intensity was used with the linear calibration plot to determine the amount of unmodified HCA** in the sample. Dividing this concentration by the total concentration of protein (as measured by UV spectroscopy), gave the percentage of contamination by unmodified HCA**.

Determination of Dissociation Constants of Ethox and SA-OMe for HCA-SCH$_2$CO$_2^-$ by Competition with DNSA.** To the wells of a black microwell plate were added dilutions of Ethox or SA-OMe, DNSA to a final concentration of 5 μM, and HCA**-SCH$_2$CO$_2^-$ to a final concentration of 25-50 nM in 200 μL of 20 mM sodium phosphate buffer pH 7.5. The plate was allowed to incubate covered (to prevent photobleaching of DNSA) at 25 °C for 1 h. The fluorescence of the plate was measured (as above). The data were fit to eq S.2, which only makes the assumption that the concentration of free DNSA is equal to the concentration of total DNSA (a reasonable assumption given that $[DNSA]_{total} >> [CA]_{total}$), and is thus a simplification

of the general equation derived by Wang et al.[7]

$$F = F_{\min} + (F_{\max} - F_{\min})\left(\frac{-A + \sqrt{A^2 + 4(K_d^{DNSA} + (K_d^{DNSA})^2/[\mathrm{DNSA}]_{total})[\mathrm{CA}]_{total} K_d^{L}[\mathrm{L}]_{total}}}{2[\mathrm{CA}]_{total}(K_d^{DNSA} + (K_d^{DNSA})^2/[\mathrm{DNSA}]_{total})}\right) \quad \text{(S.2a)}$$

$$A = K_d^{DNSA}[\mathrm{L}]_{total} - K_d^{DNSA}[\mathrm{CA}]_{total} + K_d^{L}[\mathrm{DNSA}]_{total} + K_d^{DNSA} K_d^{L} \quad \text{(S.2b)}$$

In this equation, K_d^{DNSA} and K_d^{L} are the dissociation constants of the HCA-DNSA and HCA-L complexes, respectively, $[\mathrm{L}]_{total}$ is the total concentration of the ligand (Ethox or SA-OMe), and the other terms are as defined in eq S.1. The values of $[\mathrm{DNSA}]_{total}$ and K_d^{DNSA} were constrained to their known values, and the other parameters were allowed to vary to optimize the non-linear least squares fit (Origin).

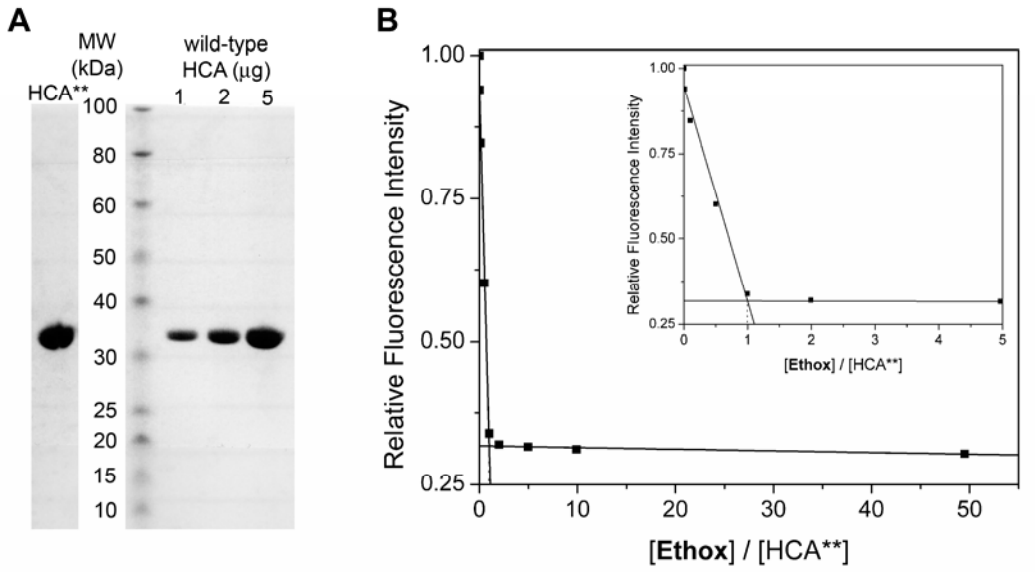


Figure S.1. Assessment of purity of HCA**. (**A**) SDS-PAGE of over-expressed and purified HCA**. No contaminant proteins were present in the "overloaded" gel. The effective molecular weight of HCA** is the same as for commercially available wild-type HCA. The HCA** lane was digitally moved to be adjacent to the molecular-weight standards to facilitate comparison between lanes (care was taken to ensure that this lane was not manipulated vertically). (**B**) Sulfonamide-binding activity of HCA**. The intrinsic fluorescence of 100 nM HCA** was measured (excitation wavelength = 290 nm, emission wavelength = 340 nm) as the high-affinity, fluorescence quencher Ethox (K_d = 0.20 nM; see main text) was titrated into the solution. The data are shown after background subtraction, correction for dilution and for the inner-filter effect (<5% correction at these concentrations; see ref 36 of main text), and normalization to a maximum signal of unity. The *x*-coordinate at the intersection of the two straight lines (shown as a vertical dotted line) shows that all (99.8%) of the HCA** can bind sulfonamides. The residual fluorescence (~0.30) at saturating concentrations of Ethox is consistent with a value reported by Kernohan et al. (ref 17 of main text). The inset shows a magnification of the main plot.

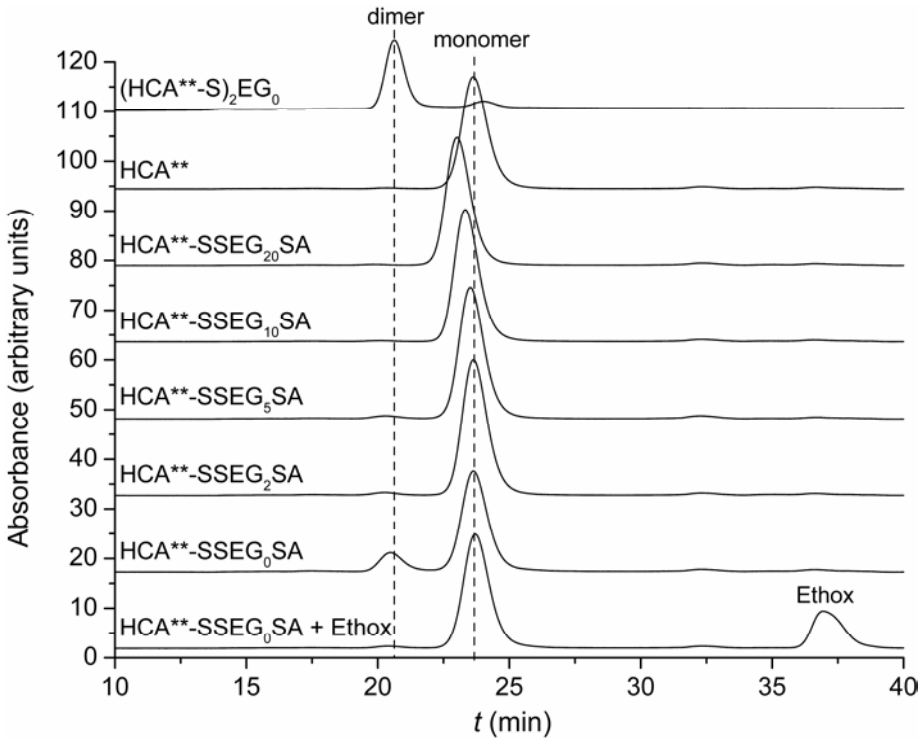


Figure S.2. Characterization of HCA**-SSEG$_n$SA proteins by size-exclusion high-performance liquid chromatography (SE-HPLC). Onto a Superdex 75 10/300 GL column (GE Healthcare Bio-Sciences, Piscataway, NJ) was injected 25 μL of a 20 μM sample of the indicated HCA** protein. The running buffer was 50 mM Tris-sulfate pH 8.0 with a flow rate of 0.5 mL min^{-1} and UV detection at 214 nm. The (HCA**-S)$_2$EG$_0$ sample was prepared by treating HCA** with a bis(maleimide) (Mack and Whitesides, manuscript in preparation) and serves as a control for the presence of HCA** dimer. HCA**-SSEG$_0$SA was incubated with a large excess of Ethox (0.67 mM; HCA**-SSEG$_0$SA + Ethox) and demonstrates that the dimer in the HCA**-SSEG$_0$SA sample is non-covalently bound. The decrease in retention time (t) of the monomer peak as the length (n) of the linker of the HCA**-SSEG$_n$SA proteins increased is consistent with the increasing size of the HCA**-SSEG$_n$SA protein with increasing n.

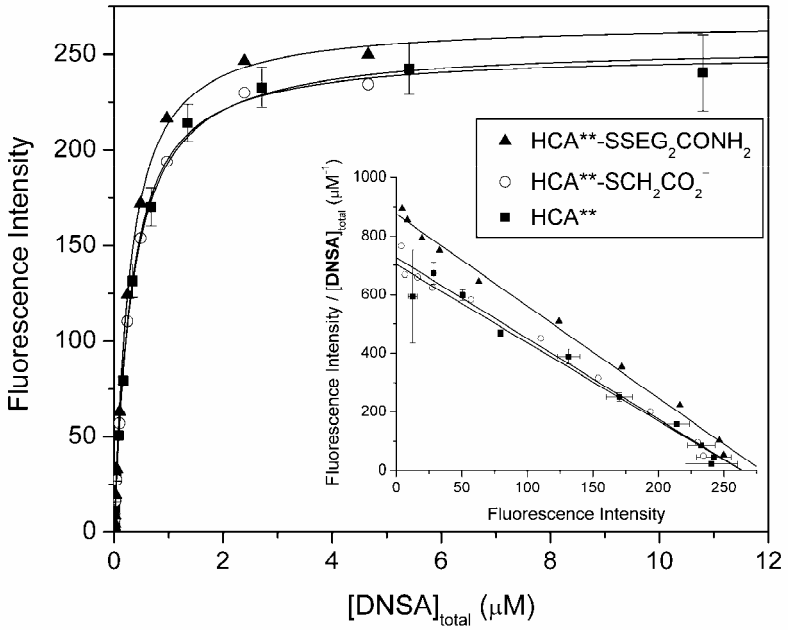


Figure S.3. Binding of dansylamide (DNSA) to HCA**, HCA**-$SCH_2CO_2^-$, and HCA**-$SSEG_2CONH_2$. The fluorescence of the HCA-DNSA complex was followed (excitation wavelength = 290 nm, emission wavelength = 460 nm) as 50-100 nM protein was treated with different concentrations of DNSA (refs 16 and 18 of main text). The data are shown after background subtraction. The solid curves are fits to the data using the full quadratic equation for binding (eq S.1; see *Supporting Experimental Procedures* and ref 18 of main text). The inset shows a Scatchard plot for the data; the assumption that $[DNSA]_{free} \approx [DNSA]_{total}$ is not true, which precludes the extraction of values of K_d from the linear fits.

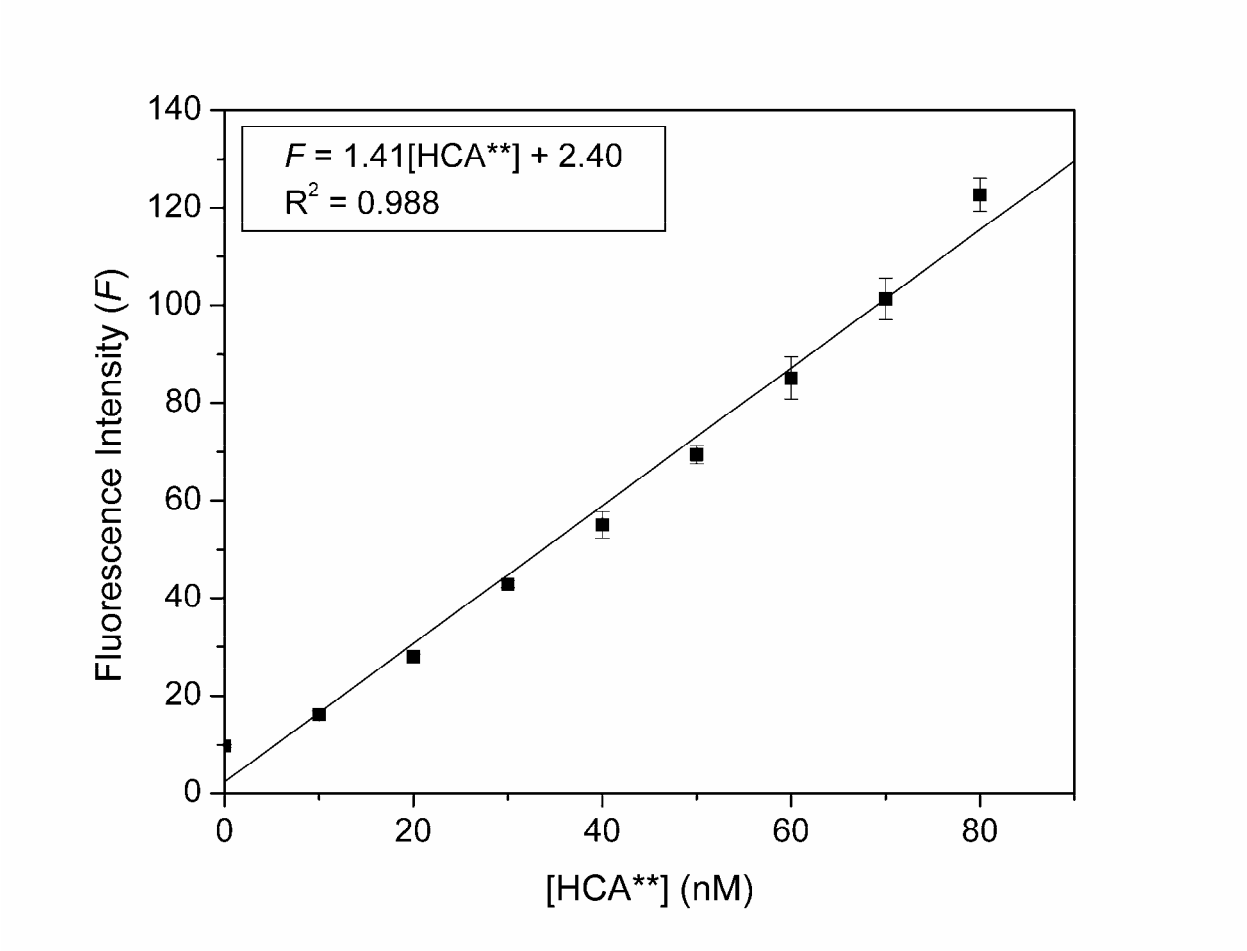


Figure S.4. Calibration plot to determine the concentration of unmodified HCA** in samples of HCA**-SSEG$_n$SA. The fluorescence of wells containing 5 μM of dansylamide (DNSA) and different concentrations of HCA** with an excitation wavelength = 290 nm, emission wavelength = 460 nm (with 455 nm cut-off filter) and a linear fit to the data are shown (see *Supporting Experimental Procedures* for details). The error bars represent the maximum variation of one measurement from the mean of four replicates.

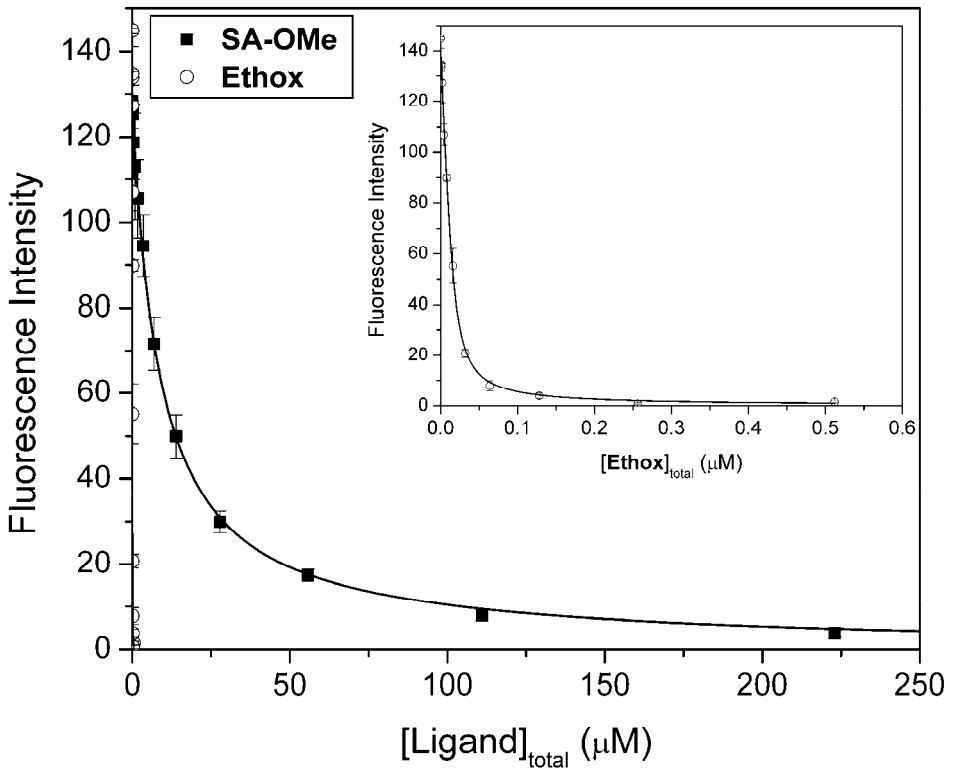


Figure S.5. Determination of dissociation constants of Ethox and SA-OMe for HCA**-$SCH_2CO_2^-$ through competition with dansylamide (DNSA). The fluorescence of HCA-DNSA complex (50 nM protein and 5 μM DNSA) equilibrated with different concentrations of ligand (Ethox or SA-OMe) was measured (excitation wavelength = 290 nm, emission wavelength = 460 nm; see refs 16, 24, and 25 of main text). The data are shown after background subtraction. The solid curves are fits to the data using the full quadratic equation for binding (eq S.2) but assuming that $[DNSA]_{free} \approx [DNSA]_{total}$ (see *Supporting Experimental Procedures*). The inset shows a magnification of the main plot to show the binding for the high-affinity ligand Ethox.

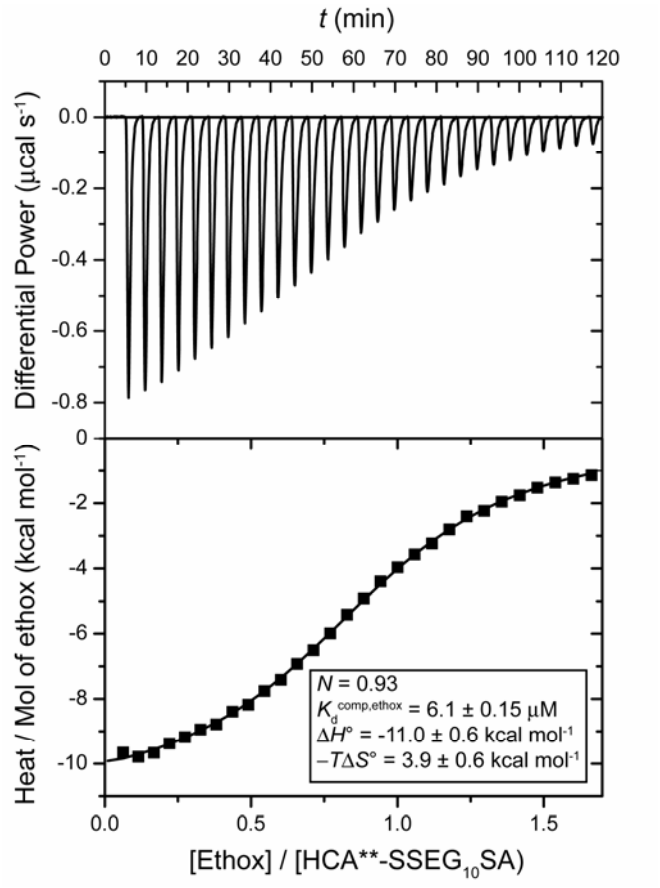


Figure S.6. Thermogram from ITC for the binding of Ethox to HCA**-$SSEG_{10}$SA at 25 °C. The sample cell contained 60.5 µM HCA**-$SSEG_{10}$SA in 20 mM sodium phosphate buffer pH 7.5 and 0.6% DMSO-d_6 (v/v) (to solubilize the arylsulfonamide and to allow for NMR quantitation, see Experimental Section). The injection syringe contained 440 µM Ethox in the same buffer. One injection of 2.0 µL preceded 29 injections of 10.0 µL. The interval between injections was 4 min. Top panel: Data after baseline correction. Bottom panel: Data after peak integration, blank subtraction, and normalization to moles of injectant. The solid line shows a sigmoid fit to a single-site binding model (with the first datum omitted). The fitting parameters are shown in the box.

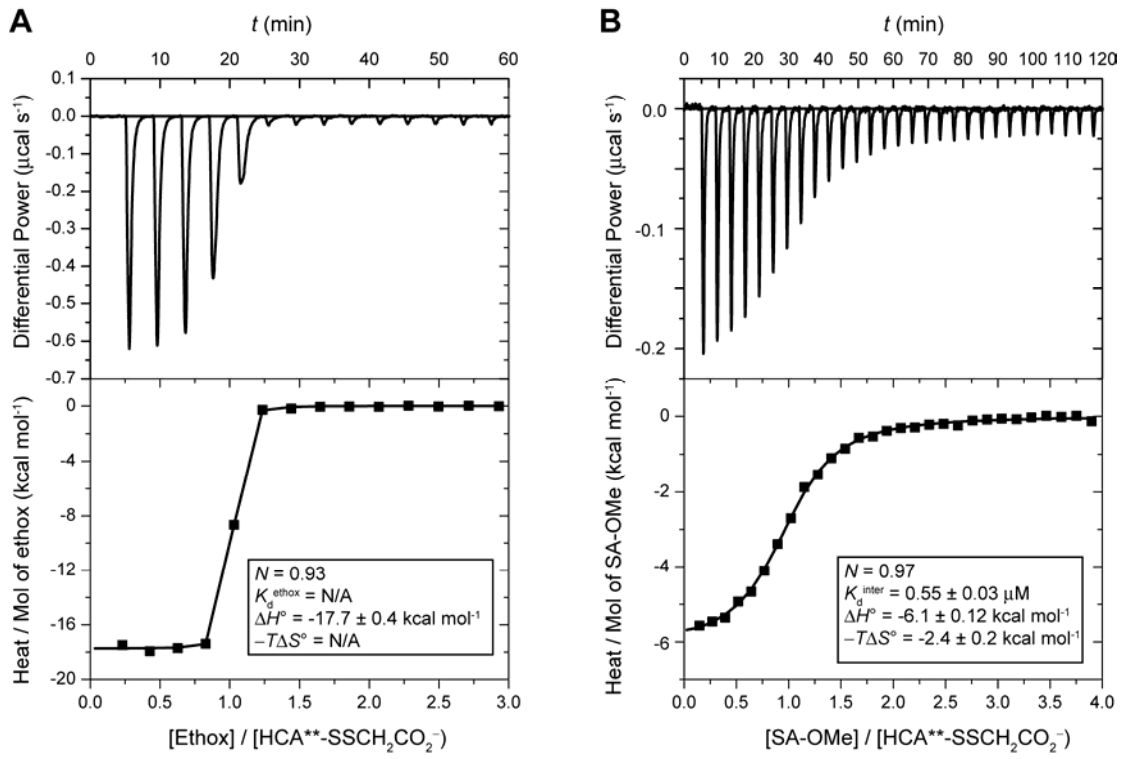


Figure S.7. Thermograms from ITC for the binding of Ethox (**A**) or SA-OMe (**B**) to HCA**-$SCH_2CO_2^-$ at 25 °C. The sample cell contained ~6 µM HCA**-$SCH_2CO_2^-$ in 20 mM sodium phosphate buffer pH 7.5 and 0.6% DMSO-d_6 (v/v) (to solubilize the arylsulfonamide and to allow for NMR quantitation, see Experimental Section). The injection syringe contained ~120 µM of the appropriate ligand in the same buffer. In (**A**), one injection of 2.0 µL preceded 14 injections of 12.0 µL. In (**B**), one injection of 2.0 µL preceded 29 injections of 10.0 µL. The interval between injections was 4 min. Top panels: Data after baseline correction. Bottom panels: Data after peak integration, blank subtraction, and normalization to moles of injectant. The solid lines show sigmoid fits to a single-site binding model (with the first datum omitted). The fitting parameters are shown in the boxes. In (**A**), a reliable value for K_d could not be estimated by ITC because the affinity of Ethox for HCA**-$SCH_2CO_2^-$ is too high to be measured directly (i.e., the thermogram lacks curvature and is a step function).

Table S.1. Calculated and observed masses of HCA** proteins.

Protein	Calculated mass (Da)	ESI-MS (Da)[c]
HCA**	29 057[a]	29 068
HCA**-$SCH_2CO_2^-$	29 138[b]	29 151
HCA**-$SSEG_2CONH_2$	29 419[a]	29 428
HCA**-$SSEG_0SA$	29 500[a]	29 504
HCA**-$SSEG_2SA$	29 588[a]	29 592
HCA**-$SSEG_5SA$	29 720[a]	29 725
HCA**-$SSEG_{10}SA$	29 940[a]	29 945
HCA**-$SSEG_{20}SA$	39 381[a]	39 385

[a] Calculated by summing the mass of HCA** lacking both the Zn^{2+} cofactor and the first residue (Met), which is believed to be post-translationally removed by *E. coli* (ref 29 of main text) and the mass of the coupled molecule. [b] Calculated for HCA** as in *a*, but including the mass of a Na^+ ion. [c] Deconvoluted mass of the major peak from electrospray ionization mass spectrometry (ESI-MS). Two minor peaks with masses greater by ~25 Da and ~62 Da than the major peak were present in all samples; these peaks presumably represent the Na^+-adduct and Zn^{2+}-adduct (holo-CA), respectively. ESI-MS was performed on a Sciex triple quadrupole mass spectrometer (Applied Biosystems, Foster City, CA) as previously described by Shaw et al.[8] Spectra were deconvoluted and a mass spectrum was generated using Max Entropy software (part of MassLynx from Waters, Milford, MA). Mass values were determined from at least six charge states.

Supporting References

1. Chu, Y. H.; Avila, L. Z.; Biebuyck, H. A.; Whitesides, G. M. *J. Med. Chem.* **1992,** *35*, 2915-2917.
2. Sigal, G. B.; Mammen, M.; Dahmann, G.; Whitesides, G. M. *J. Am. Chem. Soc.* **1996,** *118*, 3789-3800.
3. Ilies, M. A.; Masereel, B.; Rolin, S.; Scozzafava, A.; Câmpeanu, G.; Cîmpeanu, V.; Supuran, C. T. *Bioorg. Med. Chem.* **2004,** *12*, 2717-2726.
4. Carpino, L. A. *J. Am. Chem. Soc.* **1993,** *115*, 4397-4398.
5. Snider, B. B.; Ahn, Y.; O'Hare, S. M. *Org. Lett.* **2001,** *3*, 4217-4220.
6. Rink, H. *Tetrahedron Lett.* **1987,** *28*, 3787-3790.
7. Wang, Z.-X. *FEBS Lett.* **1995,** *360*, 111-114.
8. Shaw, B. F.; Durazo, A.; Nersissian, A. M.; Whitelegge, J. P.; Faull, K. F.; Valentine, J. S. *J. Biol. Chem.* **2006,** *281*, 18167-18176.