SUPPORTING INFORMATION

Genetically Encoded Optochemical Probes for Simultaneous Fluorescence Reporting and Light-Activation of Protein Function with Two-Photon Excitation

Ji Luo[1,2], Rajendra Uprety[2], Yuta Naro[1], Chungjung Chou[2], Jason W. Chin[3], and Alexander Deiters[1,*]

[1]University of Pittsburgh, Department of Chemistry, Pittsburgh, PA 15260. [2]North Carolina State University, Department of Chemistry, Raleigh, NC 27695. [3]Medical Research Council Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge CB2 0QH, United Kingdom.

**To whom correspondence should be addressed. Email: deiters@pitt.edu.*

Supporting Movies:

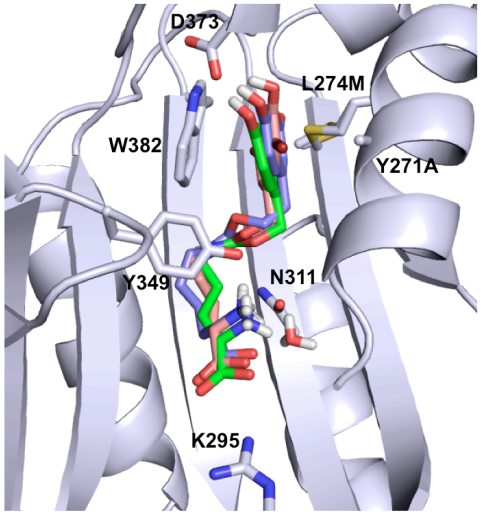
Movie S1. Fluorescent coumarin nuclear localization by Z-stack analysis. CHO K1 cells transiently expressing NLS-**1**-EGFP in the presence of BhcKRS/tRNA$_{CUA}$ and 0.25 mM of **1** were imaged by Z-stack.

Movie S2. EGFP fluorescence upon 365 nm light irradiation. HEK 293T cells transiently expressing EGFP-KTAG-mCherry in the presence of BhcKRS/tRNA$_{CUA}$ and 0.25 mM of **1** were photolysed using 365 nm light (30 s, DAPI filter) at time t_0 = 0 min. The EGFP fluorescence was followed.

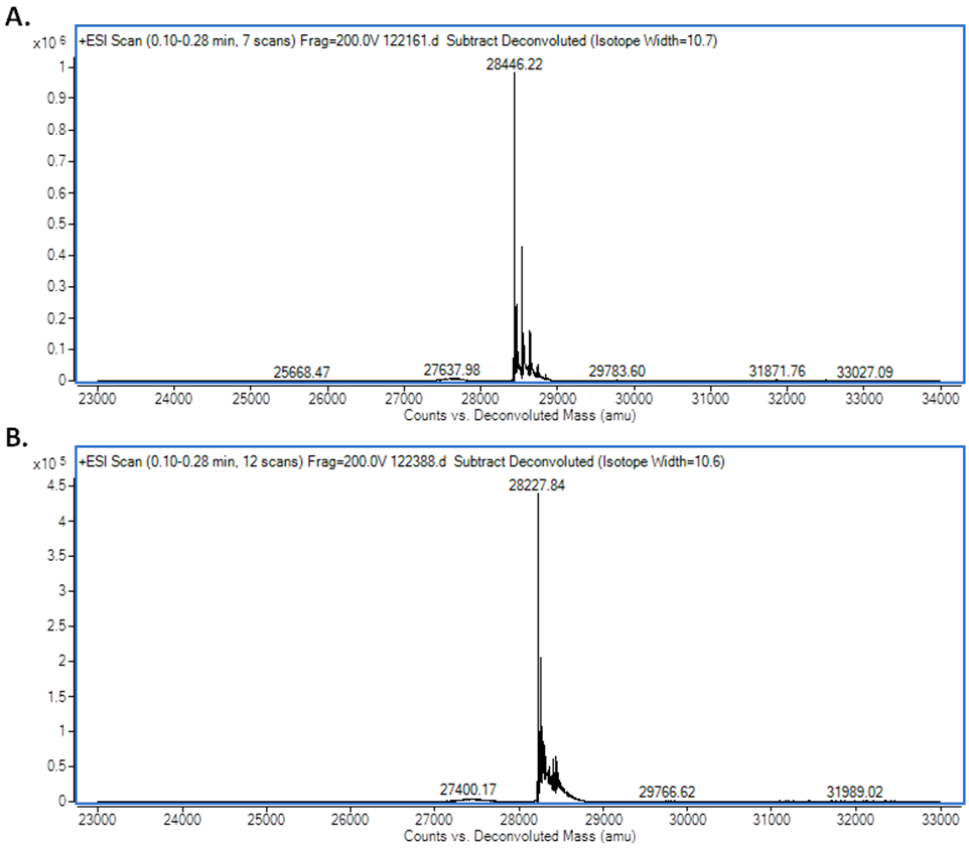
Table S1. Primer list

Primer No.	Sequence
G1	gtcctatacaagcttcgatggcttcaccatggcactag
G2	gatctagagtcgcggccgcggatccttatcattaagcg
N1	gactcagatcaattactcgagatggccaccgtcctgaagcgccccgctgccaccaaaaaggctggccaggccaaa
N2	tcgaagcttgagctcgtctgagtccggacttgtacagtgaagcgtccagcttcttctttttggcctggccagccttttgg
QC1	gaaggcttcaccatggcacaagcaattagccatggtgag
QC2	ctcaccatggctaattgcttgtgccatggtgaagccttc
GL1	gatctactggtctgccttagggtgtcgctctgcctcatag
GL2	ctatgaggcagagcgacaccctaaggcagaccagtagatc

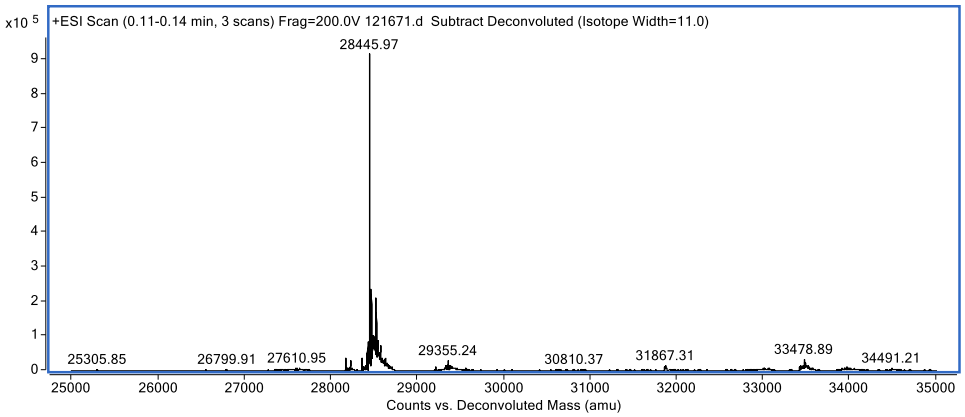
Supporting Figures:



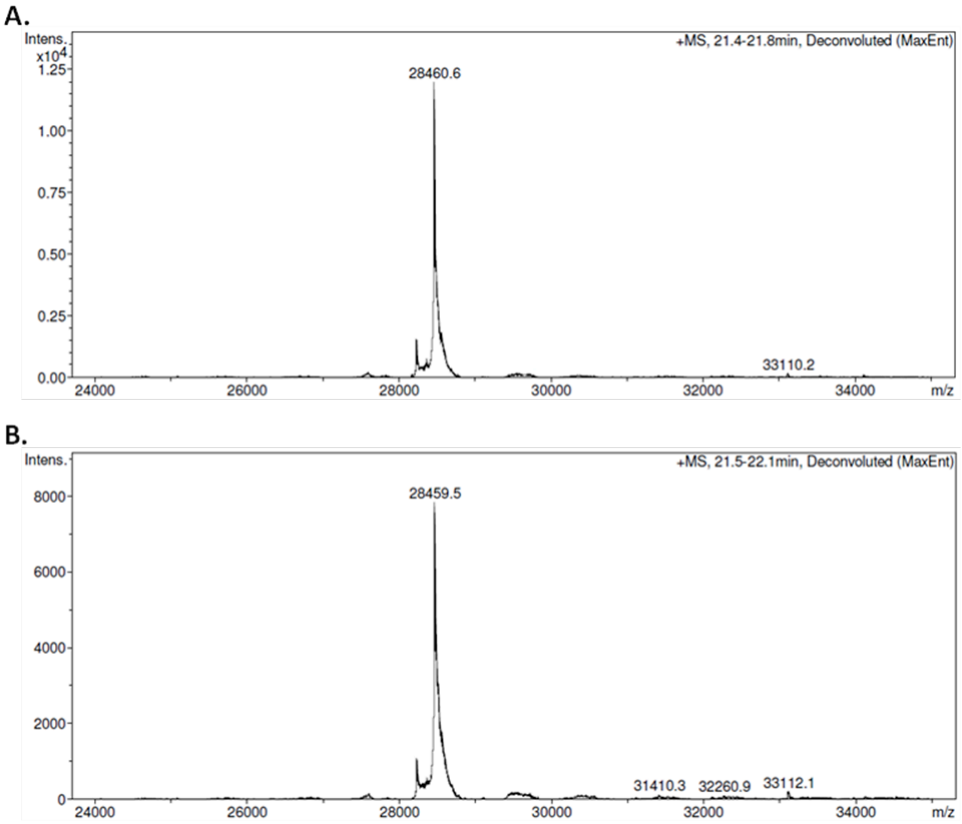
Supporting Figure S1. Docked poses of **1** (green), **2** (pink), and **3** (purple) superimposed within the BhcKRS mutant model.



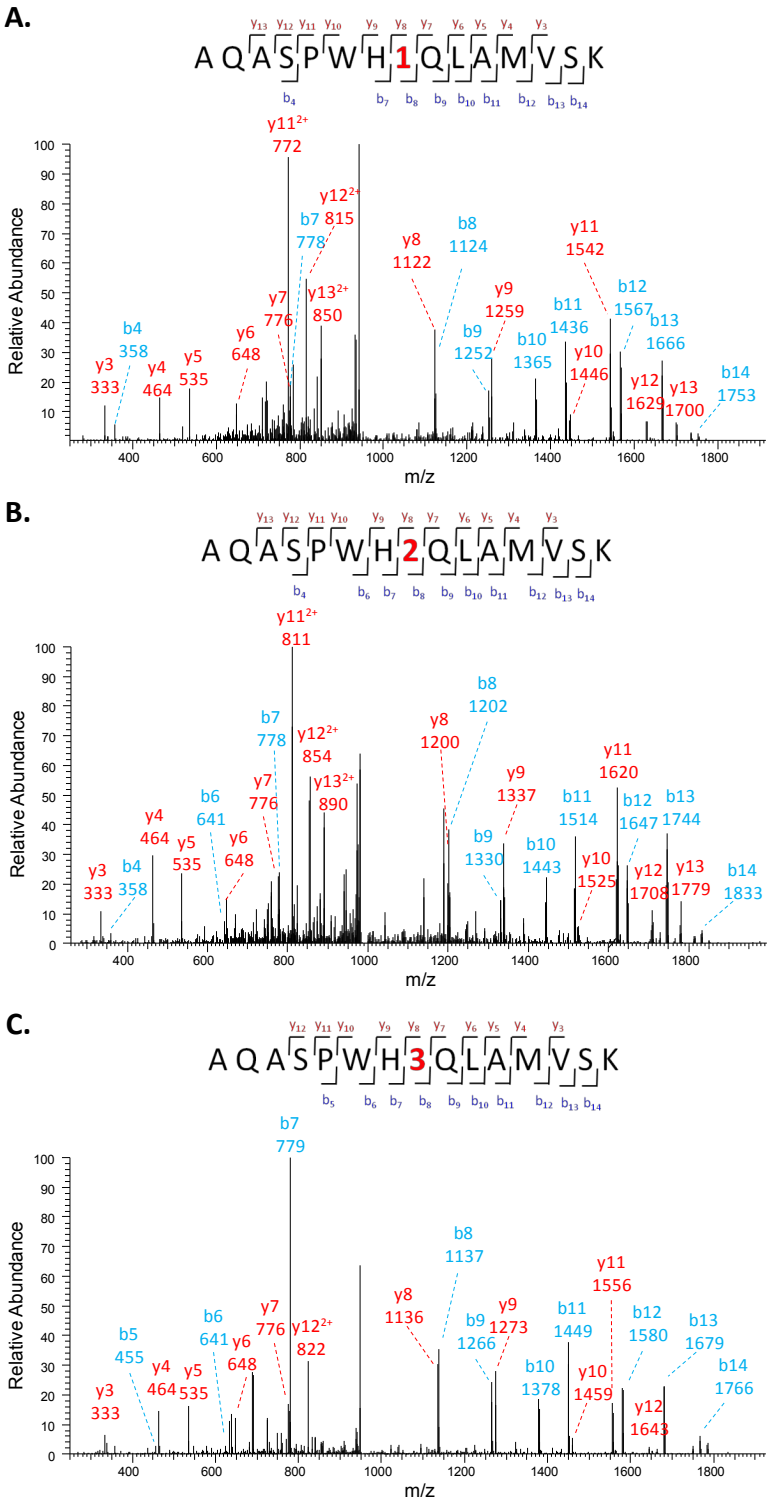
Supporting Figure S2. ESI-MS analysis of sfGFP-**1** expressed in *E. coli*. A) sfGFP-**1**: observed MS: 28446.22 Da ± 0.19 Da, expected MS: 28446.03 Da; B) decaged sfGFP-**1**: observed MS: 28227.84 Da ± 0.22 Da, expected MS: 28228.06 Da.



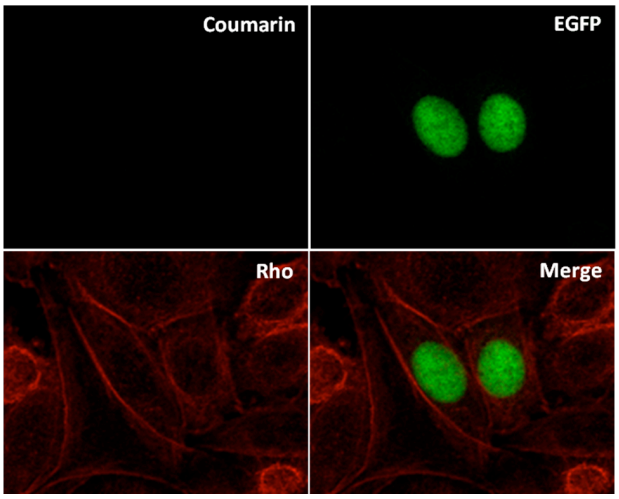
Supporting Figure S3. ESI-MS analysis of sfGFP-**2** expressed in *E. coli*. Observed MS: 28445.97 Da, expected MS: 28524.91 Da.



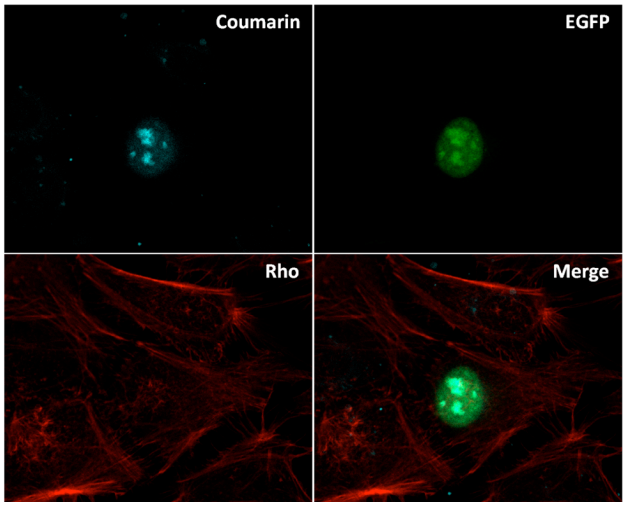
Supporting Figure S4. ESI-MS analysis of sfGFP-**3** expressed in *E. coli*. A) sfGFP-**3**: observed MS: 28460.60, expected MS: 28460.04 Da; B) irradiated sfGFP-**3**: observed MS: 28459.50 Da, expected MS: 28460.04 Da.



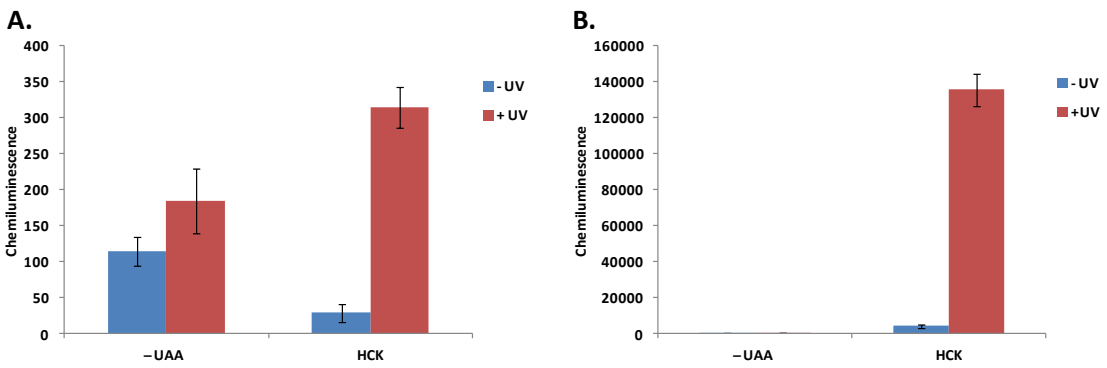
Supporting Figure S5. mCherry-EGFP-HA containing **1**, **2**, and **3** were expressed in HEK 293T cells and purified by anti-HA immunoprecipitation. MS/MS analysis confirmed the incorporation of **1**, **2**, and **3** at the specific site.



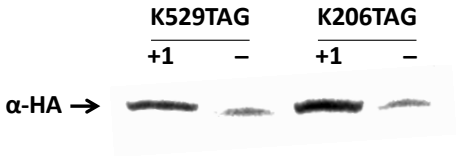
Supporting Figure S6. Fluorescent nuclear localization by confocal live cell imaging of CHO K1 cells transfected with pNLS-WT-EGFP.



Supporting Figure S7. Fluorescent nuclear localization by confocal microscopy images of HeLa cells co-transfected with pNLS-TAG-EGFP-HA and pBhcKRS-4PylT in 0.25 mM **1** with Rhodamin-phalloidine staining. Representative confocal micrographs showing the fluorescent nuclear localization of NLS-**1**-EGFP fusions.



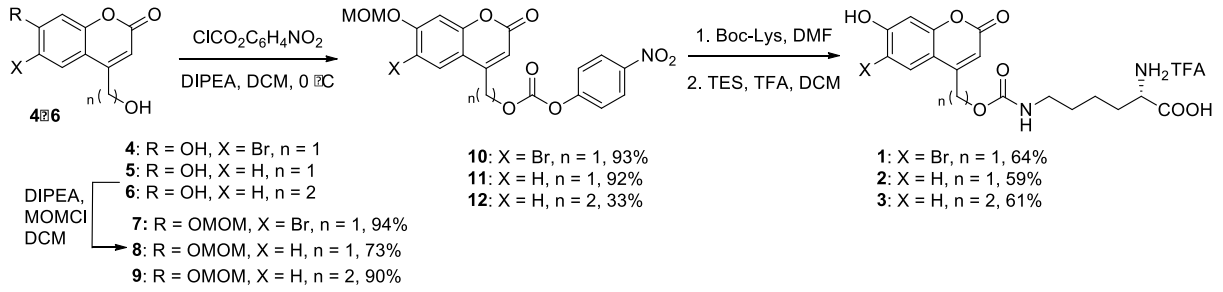
Supporting Figure S8. Bright-Glo luciferase assay of cells that were either kept in the dark or irradiated (365 nm, 4 min). **A)** No enzymatic activity was observed for the Fluc-K529→**1** mutant and only a low level of luminescence was observed after photolysis – compared to K206→**1**. **B)** Fluc-K206→**1** showed much higher luminescence signal an excellent light activation. Error bars represent standard deviations from three independent experiments.



Supporting Figure S9. Immunoblot of cells transfected with Fluc-K206→**1**-HA and Fluc-K529→**1**-HA with an anti-HA antibody and an HRP secondary antibody, followed by development using Pierce ECL Plus Substrate for 5 min and imaging on a ChemiDoc MP (Bio-Rad; auto-exposure).

Synthetic Protocols:

The coumarin lysines **1**-**3** were synthesized in three steps from the alcohols **4-6** (Scheme S1). Ethyl 4-chloroacetoacetate was reacted with either bromoresorcinol or resorcinol to prepare the corresponding coumarin **4** or **5** following literature procedures.[1 2] Similarly, the coumarin alcohol **6** was prepared from resorcinol and acetonedicarboxylic acid followed by reduction.[3] These alcohols **4-6** were then treated with MOMCl in the presence of DIPEA in DCM to protect the phenolic hydroxyl groups, delivering the bromocoumarin alcohol[4] **7** in 94**%** yield, and alcohols **8** and **9** in 73% and 90% yields respectively. The compounds **7**-**9** were activated using nitrophenyl chloroformate in the presence of DIPEA or Cs_2CO_3 in DCM to furnish the carbonates **10**[4] (92% yield)**, 11** (93% yield), and **12** (33% yield). Activation of **9** was performed in the presence of Cs_2CO_3 in DCM. Use of different bases (such as K_2CO_3 or DIPEA) or different solvents (THF or dioxane) did not improve the yield. Additionally, the poor solubility of the product in common organic solvents posed issues during purification using silica gel column chromatography. Attempted purification of the product by crystallization did not improve the yield either. The carbonates **10**-**12** were then reacted with Boc-lysine in DMF followed by treatment of the crude product with TFA in DCM (1:1) in the presence of TES as a cation scavenger to obtain the corresponding coumarin lysines **1**-**3**. Both, the MOM and the Boc protecting groups in **10**-**12** were removed in a single step resulting in the bromohydroxycoumarin lysine **1** in 64% yield, hydroxycoumarin lysine **2** in 59% yield, and the hydroxymethylene coumarin lysine **3** in 61% yield.



Supporting Scheme S1. Synthesis of the coumarin lysines **1**-**3**.

4-(Hydroxymethyl)-7-(methoxymethoxy)-2H-chromen-2-one (**8**). Diisopropylethyl amine (1.80 mL, 10.4 mmol) and MOMCl (175 µL, 2.29 mmol) were added to a solution of the 7-hydroxy coumarin **5** (400 mg, 2.08 mmol) in dry DCM (5 mL) stirred at 0 °C under an inert atmosphere. The reaction mixture was allowed to warm to room temperature and stirring was continued for 12 h. Next, the reaction mixture was diluted with DCM (20 mL), the organic layer was washed with a saturated aqueous solution of $NaHCO_3$ (3 × 20 mL) and brine (20 mL), and then dried over anhydrous Na_2SO_4. The solvent was removed under reduced pressure to obtain MOM-protected coumarin **8** (491 mg, yield 79%) as a white solid; mp 147-150 °C. ^{1}H NMR (300 MHz, CD_3OD): δ = 7.54 (d, *J* = 4.6 Hz, 1H), 6.97 (s, 1H), 6.38 (s, 1H), 5.23 (s, 2H), 4.78 (s, 2H), 3.42 (s, 3H). ^{13}C NMR (75 MHz, CD_3OD): δ = 164.3, 162.4, 158.7, 156.9, 126.7, 115.4, 113.8, 109.8, 105.1, 96.2, 61.4, 57.2. MS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{12}H_{13}O_5$: 237.0768; found 237.0760.

4-(2-Hydroxyethyl)-7-(methoxymethoxy)-2H-chromen-2-one (**9**). Diisopropylethyl amine (251 µL, 1.44 mmol) and MOMCl (46 µL, 0.58 mmol) were added to a solution of the 7-hydroxy coumarin alcohol **6** (100 mg, 0.48 mmol) in dry DCM (2 mL) stirred at 0 °C under an inert atmosphere. The reaction mixture was allowed to warm to room temperature and the stirring

was continued for 12 h. The reaction mixture was diluted with DCM (10 mL), the organic layer was washed with a saturated aqueous solution of $NaHCO_3$ solution (3 × 10 mL), brine (5 mL), dried over anhydrous Na_2SO_4, and filtered. The solvent was removed under reduced pressure to furnish the MOM-protected coumarin alcohol **9** (121 mg, yield 90%) as a low melting solid. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 7.55 (d, *J* = 4.3 Hz, 1 H), 6.96 (d, *J* = 4.3 Hz, 1 H), 6.21 (s, 1H), 5.21 (s, 2H), 3.97 (t, *J* = 6.3 Hz, 1 H), 3.46 (s, 3H), 2.99 (t, *J* = 6.3 Hz, 1 H). ^{13}C NMR (75 MHz, $CDCl_3$): δ = 161.5, 160.3, 155.4, 153.5, 125.7, 113.9, 113.5, 112.6, 104.2, 94.5, 60.9, 56.5, 34.9. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{13}H_{15}O_5$: 251.09; found 251.09.

(7-(Methoxymethoxy)-2-oxo-2H-chromen-4-yl)methyl (4-nitrophenyl) carbonate (**11**).
Diisopropylethyl amine (936 µL, 5.35 mmol) and *p*-nitrophenyl chloroformate (430 mg, 2.14 mmol) were added to a solution of the coumarin alcohol **8** (353 mg, 1.07 mmol) in dry DCM (5 mL) stirred at 0 °C under an inert atmosphere. The reaction mixture was warmed to room temperature and stirring was continued for 12 h. The reaction mixture was concentrated under reduced pressure and the residue was dissolved in EtOAc (20 mL). The organic layer was washed with a saturated aqueous solution of $NaHCO_3$ solution (3 × 20 mL) and brine (20 mL), and dried over anhydrous Na_2SO_4. The filtrate was concentrated and the crude product was precipitated in Et_2O:hexanes (1:1) mixture to obtain the coumarin carbonate **11** (395 mg, yield 92%) as a white solid; mp 138-140 °C. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 8.32-8.29 (m, 2H), 7.46-7.41 (m, 3H), 7.07-6.99 (m, 2H), 6.46 (s, 1H), 5.45 (s, 1H), 5.25 (s, 2H), 3.4 (s, 3H). ^{13}C NMR (75 MHz, $CDCl_3$): δ = 160.8, 155.3, 147.5, 125.7, 124.6, 124.5, 121.9, 113.9, 111.4, 104.5, 94.6, 65.7, 56.6. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{19}H_{16}NO_9$: 402.08; found 402.08.

2-(7-(Methoxymethoxy)-2-oxo-2H-chromen-4-yl)ethyl (4-nitrophenyl) carbonate (**12**).
Cs_2CO_3 (253 mg, 0.780 mmol), DMAP (63 mg, 0.52 mmol), and *p*-nitrophenyl chloroformate (156 mg, 0.780 mmol) were added to a solution of the coumarin alcohol **9** (353 mg, 1.07 mmol) in dry THF (5 mL) stirred at 0 °C under an inert atmosphere. The reaction mixture was allowed to warm to room temperature and the stirring was continued for 12 h. The reaction mixture was filtered to discard any solid precipitate and the filtrate was concentrated under reduced pressure. EtOAc (15 mL) was added to the residue and the organic layer was washed with a saturated aqueous solution of $NaHCO_3$ (3 × 10 mL) and brine (20 mL). The organic layer was dried over anhydrous Na_2SO_4, filtered, and the filtrate was concentrated and the obtained crude product was purified by column chromatography on silica gel using an EtOAc:hexanes (1:1) mixture to obtain the coumarin carbonate **12** (70 mg, 33% yield) as a crystalline white solid. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 8.28 (d, *J* = 4.6 Hz, 2H), 7.56 (d, *J* = 4.5 Hz, 1H), 7.35 (d, *J* = 4.6 Hz, 2H), 7.05-6.99 (m, 2H), 6.25 (s, 1H), 5.23 (s, 2H), 4.59 (t, *J* = 6.6 Hz, 2H), 3.48 (s, 3H), 3.21 (t, *J* = 6.6 Hz, 2H). ^{13}C NMR (75 MHz, $CDCl_3$): δ = 160.9, 160.6, 155.5, 152.6, 150.9, 125.6, 125.2, 122.0, 113.8, 113.2, 104.5, 94.5, 66.57, 56.6, 31.0. MS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{20}H_{18}NO_9$: 416.0903; found 416.0903.

(*S*)-2-Amino-6-((((6-bromo-7-hydroxy-2-oxo-2H-chromen-4-yl)methoxy)carbonyl)amino)hexanoic acid (**1**). Boc lysine (400 mg, 1.62 mmol) was added to a solution of the bromocoumarin carbonate **10** (600 mg, 1.25 mmol) in dry DMF (7 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 30 h. Next, the reaction mixture was poured into water (100 mL) and the product in the aqueous layer was extracted using EtOAc (3 × 20 mL). The combined organic layers were washed with water (3 × 20 mL) and brine (20 mL), dried over anhydrous Na_2SO_4 and filtered. The solvent was removed to obtain crude product as a solid (637 mg). The crude product was subjected to the next step synthesis without further purification. ^{1}H NMR (300 MHz, DMSO-d_6): δ = 7.97 (s, 1H), 7.58 (br, 1H), 7.15 (br, 1H), 6.28 (s, 1H), 5.43 (s, 2H), 5.27 (s, 2H), 3.80 (br, 1H), 3.42 (s, 3H), 3.20 (br, 2H), 1.65-1.42 (m, 2H), 1.36-1.34 (s, 13H). Triethylsilane (377 µL, 2.34 mmol) and the Bhc Boc-

lysine were added to a solution of TFA:DCM (1:1, 14 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 45 minutes. The reaction mixture was concentrated under reduced pressure and the obtained residue was dissolved in MeOH (5 mL) and then re-concentrated. The process was repeated three times to remove any residual TFA. The crude product was dissolved in MeOH (1 mL) and precipitated in Et_2O (150 mL) under vigorous stirring. The precipitate was collected, redissolved in MeOH, and precipitated. The processes was repeated twice to obtain the Bhc-lysine TFA salt **1** (450 mg, yield 64%) as a crystalline white solid. ^{1}H NMR (300 MHz, DMSO-d_6): δ = 7.82 (s, 1H), 7.03 (s, 1H), 6.14 (s, 1H), 5.22 (s, 2H), 3.57 (t, *J* = 6.0 Hz, 13H), 3.00-2.98 (m, 2H), 1.74-1.69 (br, 2H), 1.42 (br, 4H). ^{13}C NMR (75 MHz, DMSOd_6): δ = 171.1, 159.8, 157.9, 155.2, 151.3, 128.2, 110.1, 108.15, 106.4, 103.3, 60.8, 52.6, 30.0, 28.9, 21.9. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{17}H_{20}BrN_2O_7$: 443.04; found 443.11.

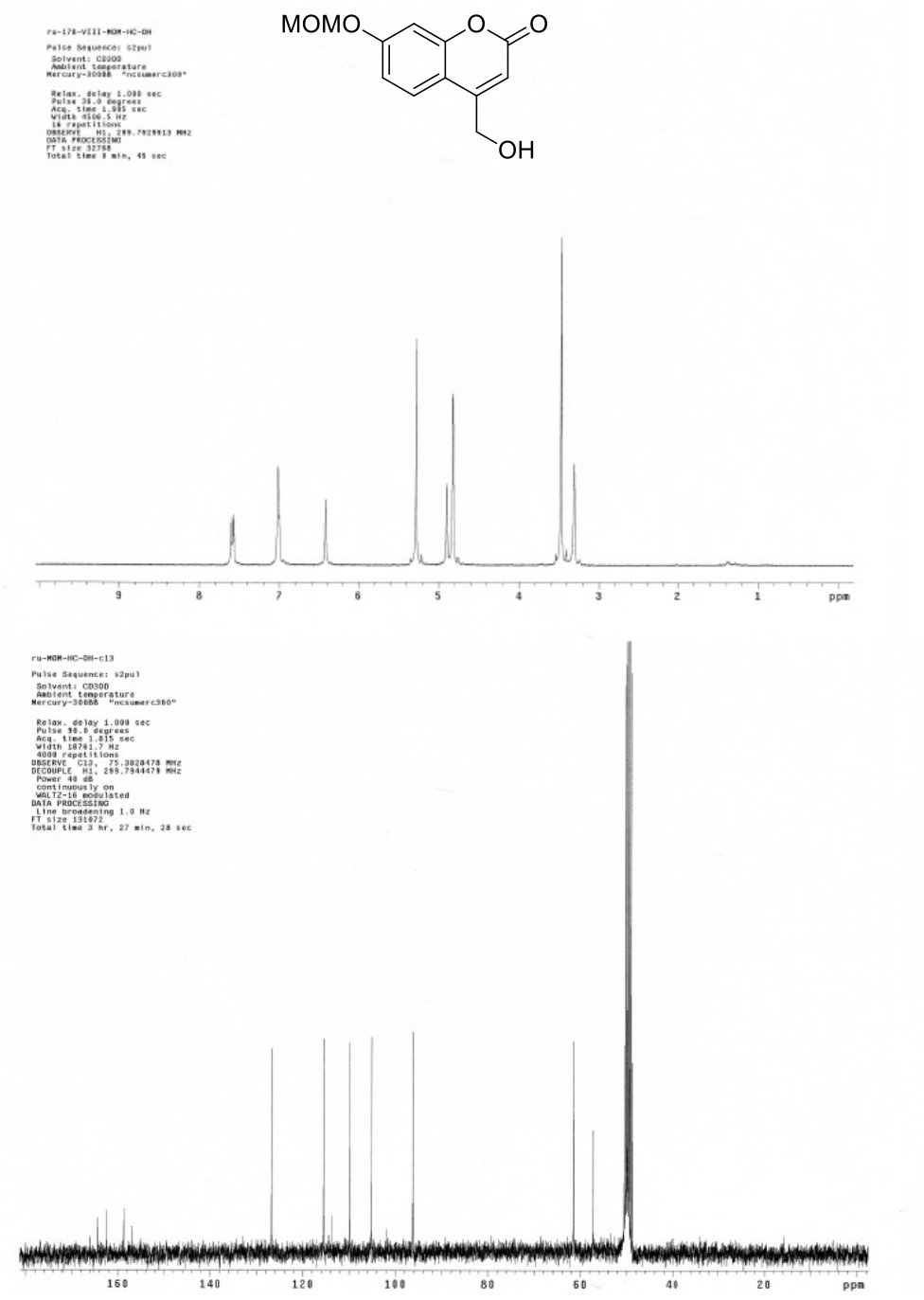
(*S*)-2-Amino-6-((((7-hydroxy-2-oxo-2H-chromen-4-yl)methoxy)carbonyl)amino)hexanoic acid (**2**). Boc lysine (311 mg, 1.26 mmol) was added to a solution of the coumarin carbonate **11** (390 mg, 0.970 mmol) in dry DMF (5 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 12 h, the reaction mixture was poured into water (50 mL) and the product in the aqueous layer was extracted using EtOAc (2 × 20 mL). The combined organic layers were washed with water (3 × 20 mL) and brine (20 mL), and dried over anhydrous Na_2SO_4. The filtrate was concentrated under reduced pressure and the solid product was dried under a high vacuum. The crude product was used to the next step without further purification. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 7.40 (d, *J* = 4.3 Hz, 1H), 7.00-6.96 (m, 2H), 6.33 (s, 1H), 5.36-5.31 (m, 1H), 5.29-5.22 (m, 4H), 4.28 (br, 1H), 3.47 (s, 3H), 3.25-3.23 (m, 2H), 1.82-1.70 (m, 2H), 1.57-1.24 (s, 13H). The crude compound was treated with triethylsilane (350 µL, 2.17 mmol) and a solution of TFA:DCM (1:1, 10 mL) at room temperature under an inert atmosphere. Stirring was continued for 40 minutes. The reaction mixture was concentrated under reduced pressure. The obtained residue was dissolved in MeOH (2 mL) and concentrated under reduced pressure. The process of adding MeOH followed by the removal of solvents was repeated three times to remove residual amount of TFA. Finally, the crude product was dissolved in MeOH (1 mL), precipitated in Et_2O (50 mL) under vigorous stirring. The precipitate was collected, and the processes of dissolution followed by precipitation was repeated twice to furnish coumarin lysine **2** (275 mg, yield 59%) as a crystalline white solid; mp 215-218 °C. ^{1}H NMR (300 MHz, DMSO- d_6): δ = 7.56-7.48 (m, 2H), 6.82-6.76 (m, 2H), 6.10 (s, 1H), 5.44 (s, 1H), 3.57 (t, *J* = 3.0 Hz, 1H), 3.01 (br, 2H), 1.73 (br, 2H), 1.41 (br, 4H). ^{13}C NMR (75 MHz, DMSO- d_6): δ = 172.0, 162.6, 160.9, 160.7, 156.0, 155.6, 152.7, 150.3, 126.3, 113.9, 109.9, 107.6, 103.1, 61.5, 53.5, 30.8, 29.6, 22.6. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{17}H_{21}N_2O_7$: 365.13; found 365.13.

(*S*)-2-Amino-6-((((7-hydroxy-2-oxo-2H-chromen-4-yl)methoxy)carbonyl)amino)hexanoic acid (**3**). Boc lysine (57 mg, 0.23 mmol) was added to a solution of the coumarin carbonate **12** (74 mg, 0.17 mmol) in dry DMF (1.5 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 30 h, the reaction mixture was poured into water (20 mL), and the aqueous layer was extracted using EtOAc (2 × 15 mL). The combined organic layers were washed with water (3 × 15 mL), brine (10 mL), then dried over anhydrous Na_2SO_4 and filtered. The filtrate was concentrated to dryness to obtain the crude product. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 7.56 (d, *J* = 4.2 Hz, 1H), 7.01-6.92 (m, 2H), 6.22 (s, 1H), 5.50 (br, 1H), 5.22 (s, 2H), 4.37-4.32 (m, 3H), 3.47 (s, 3H), 3.18 (br, 2H), 3.08-3.00 (m, 2H), 1.82-1.51 (m, 2H), 1.49-1.23 (s, 13H). Without further purification, the crude compound was subjected to the next step by treatment with triethylsilane (55 µL, 0.34 mmol) and a solution of TFA:DCM (1:1, 2 mL) at room temperature under an inert atmosphere. Stirring was continued for 40 minutes. The reaction mixture was concentrated under reduced pressure to remove both the solvent and TFA. The residue was redissolved in DCM (2 mL) followed by concentration under reduced pressure and

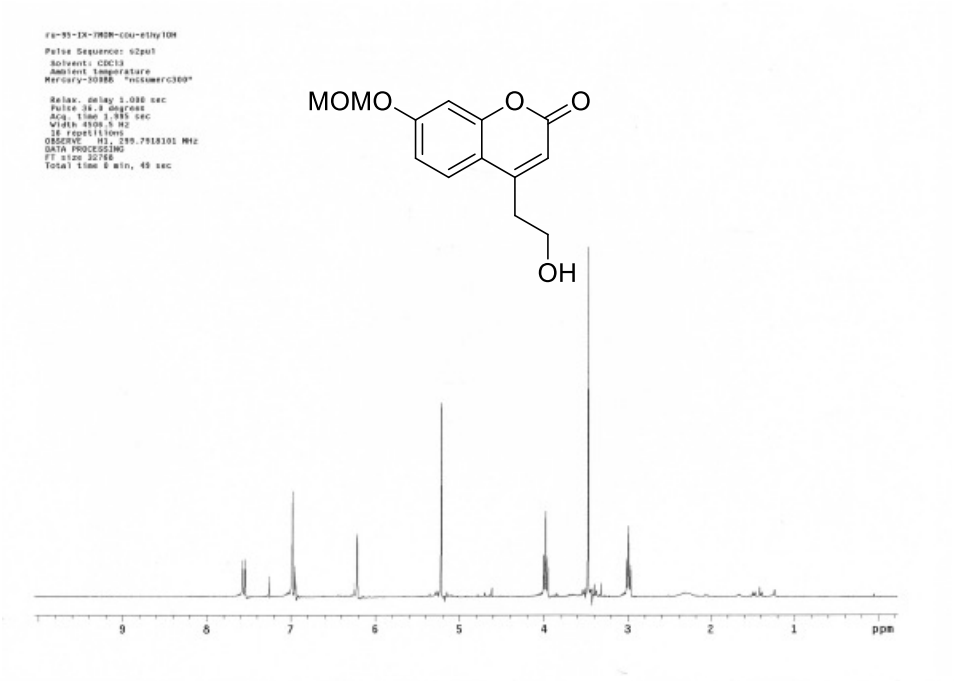
the process of adding DCM followed by the removal of solvents was repeated three times to remove any residual amounts of TFA. The crude product was dissolved in MeOH (300 µL) and precipitated in Et_2O (15 mL) under vigorous stirring. The precipitate was collected, re-dissolved in MeOH (300 µL), and precipitated again into Et_2O. The process was repeated twice more to furnish the coumarin lysine TFA salt **3** (53 mg, 61% yield) as a crystalline white solid. ^{1}H NMR (300 MHz, CD_3OD): δ = 7.67 (d, *J* = 4.3 Hz, 1H), 6.84-6.81 (m, 1H), 6.71 (s, 1H), 6.13 (s, 1H), 4.37-4.33 (m, 2H), 3.81-3.77 (m, 1H), 3.11-3.08 (m, 2H), 1.89-1.87 (m, 2H), 1.49 (br, 4H). ^{13}C NMR (75 MHz, CD_3OD): δ = 171.7, 162.8, 161.8, 155.7, 154.6, 126.1, 113.3, 111.8, 110.5, 102.5, 62.3, 53.5, 40.0, 31.4, 30.3, 29.2, 22.1. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{18}H_{22}N_2O_7$: 379.15; found 379.15.

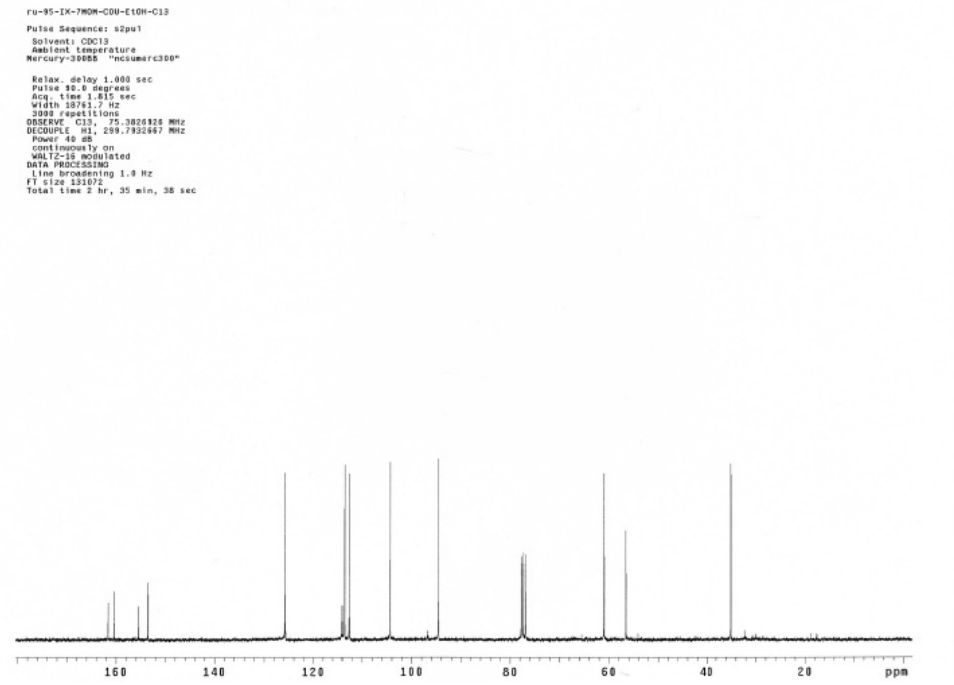
NMR Spectra of New Compounds:

4-(Hydroxymethyl)-7-(methoxymethoxy)-2H-chromen-2-one (**8**).

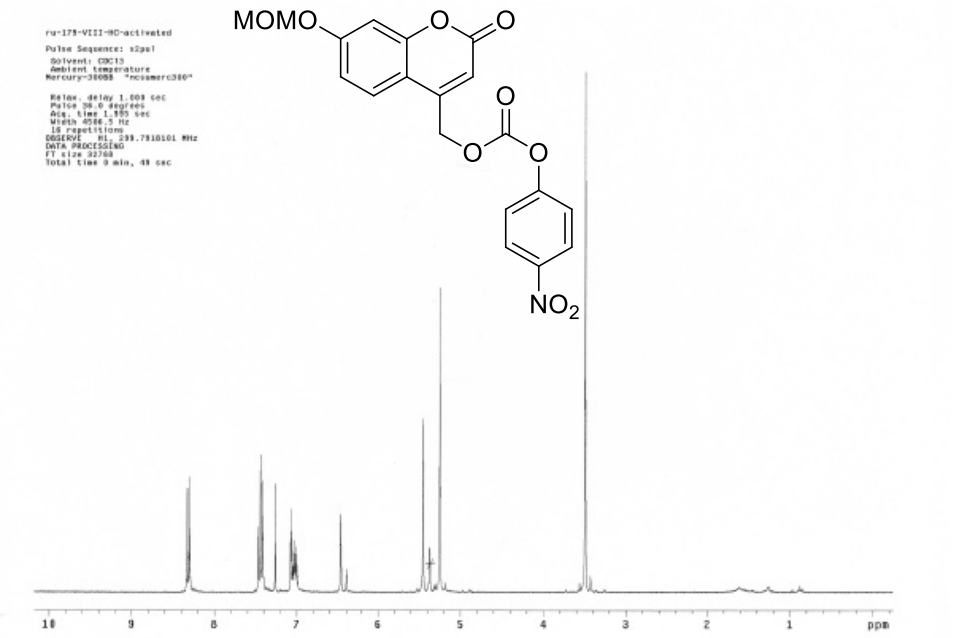


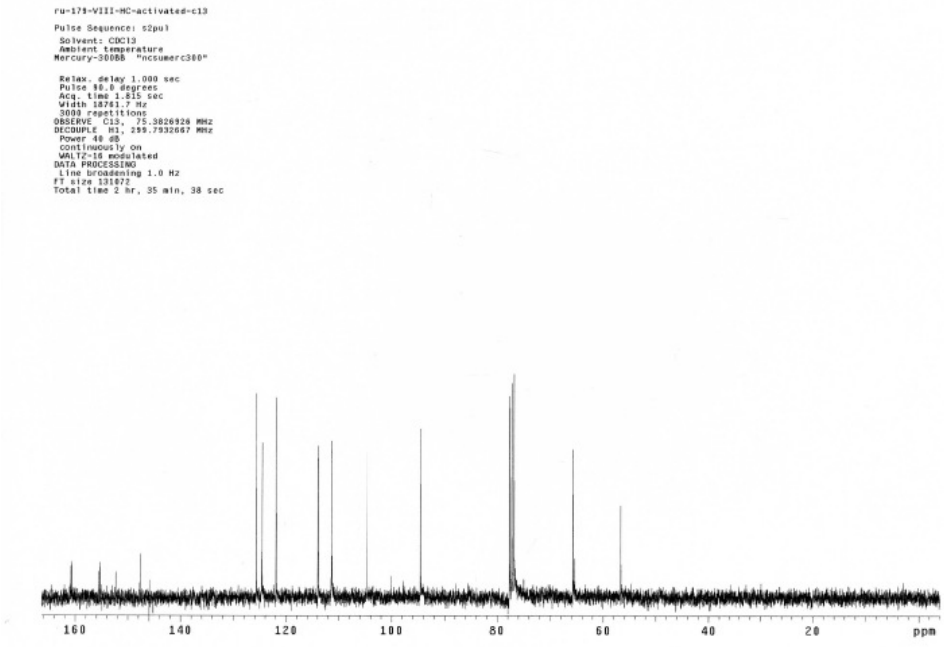
4-(2-Hydroxyethyl)-7-(methoxymethoxy)-2H-chromen-2-one (**9**).



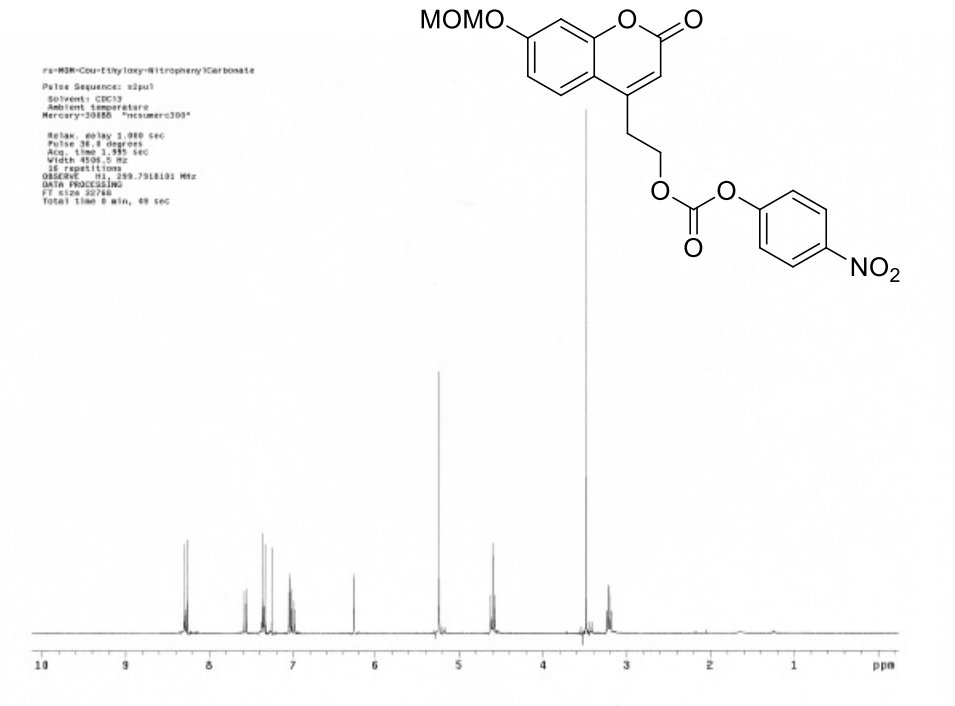


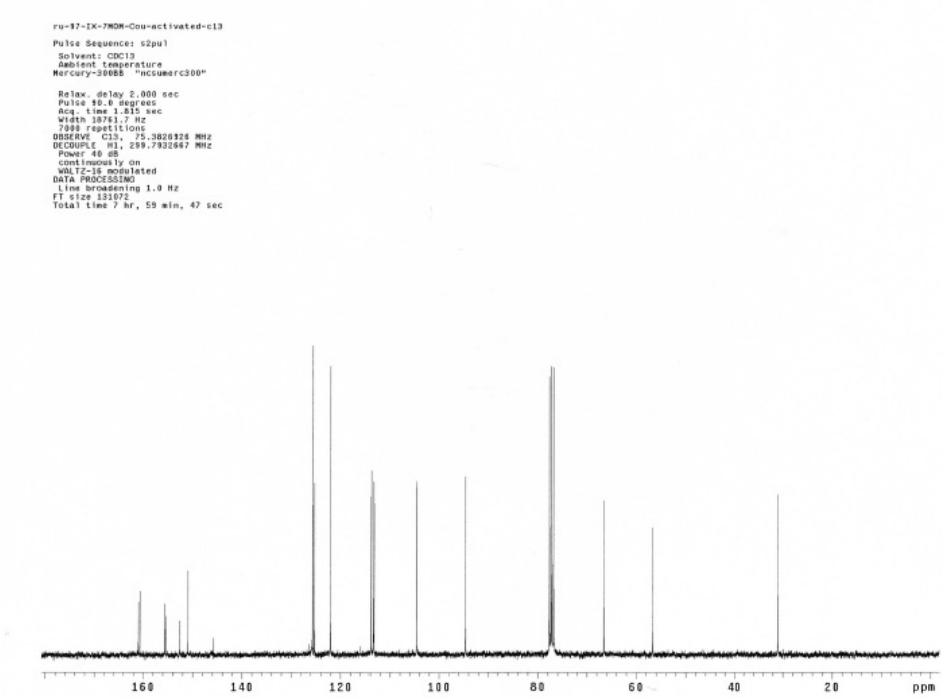
(7-(Methoxymethoxy)-2-oxo-2H-chromen-4-yl)methyl (4-nitrophenyl) carbonate (**11**)



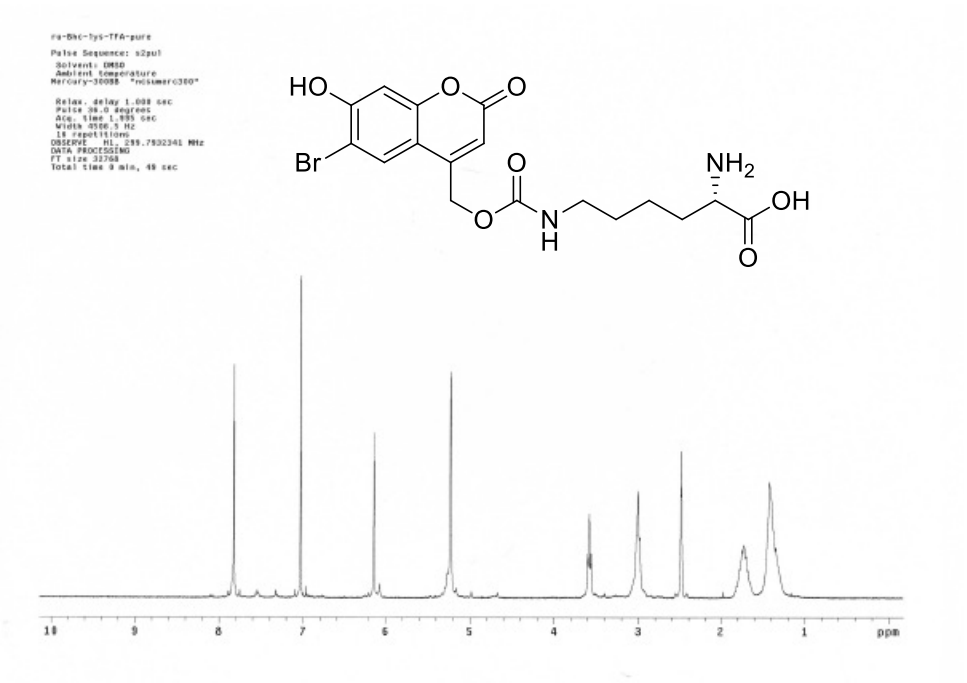


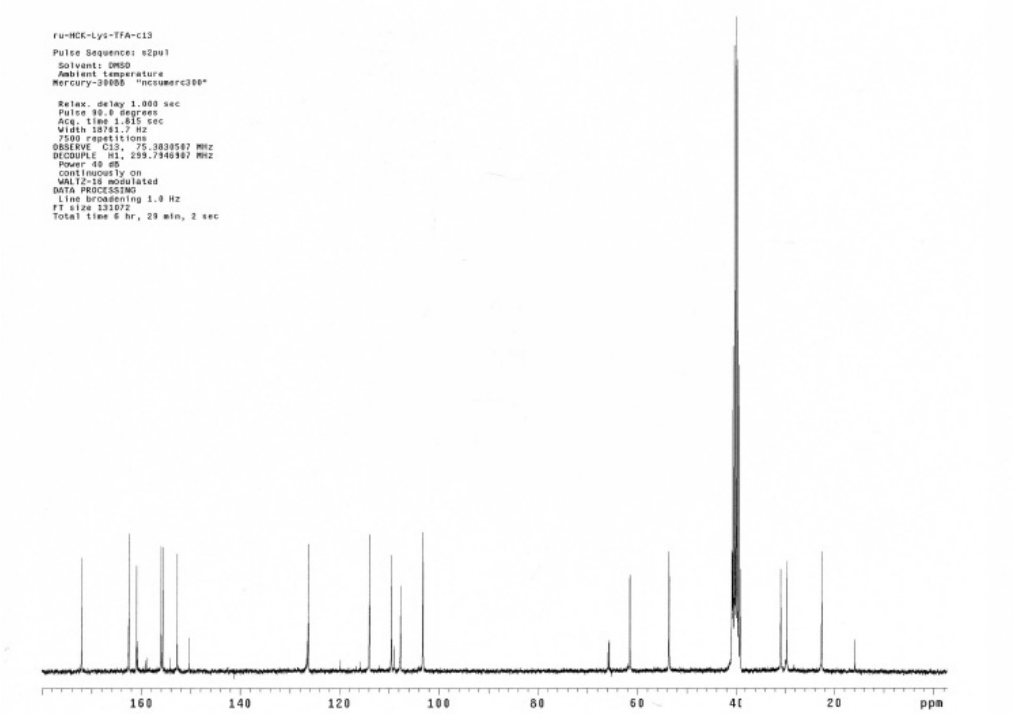
2-(7-(Methoxymethoxy)-2-oxo-2H-chromen-4-yl)ethyl (4-nitrophenyl) carbonate (**12**).



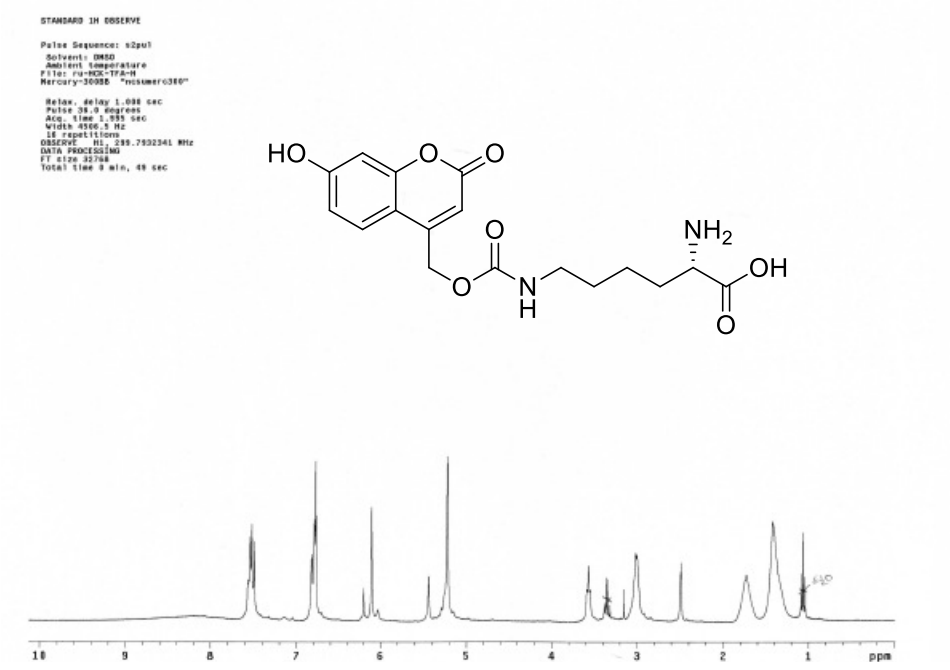


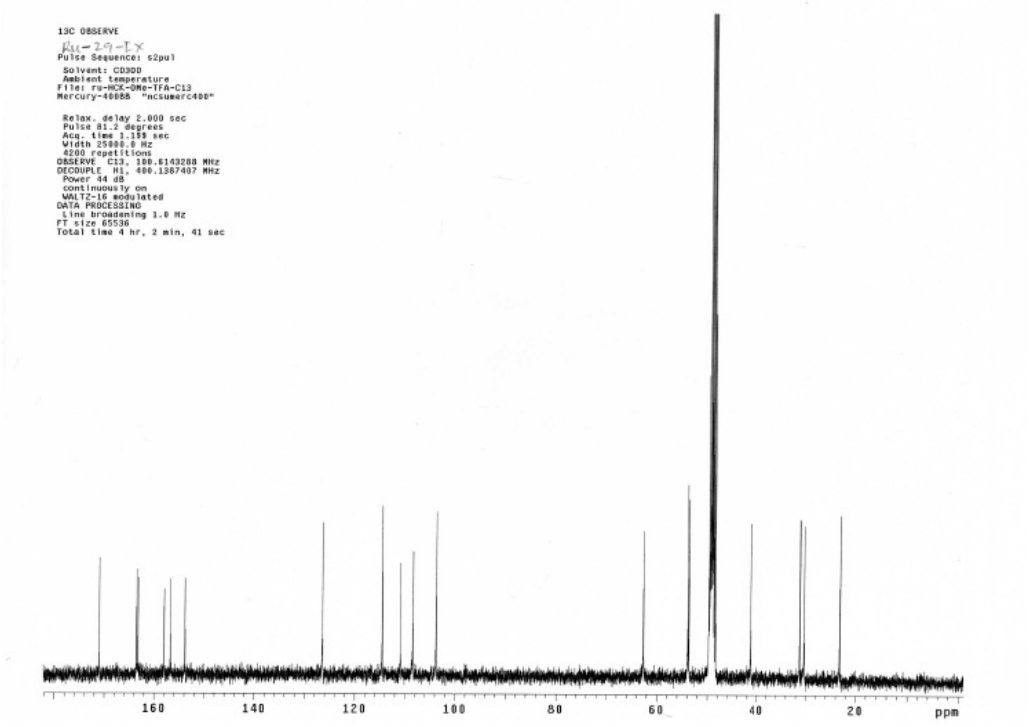
(*S*)-2-Amino-6-((((6-bromo-7-hydroxy-2-oxo-2H-chromen-4-yl)methoxy)carbonyl)amino)hexanoic acid (**2**).



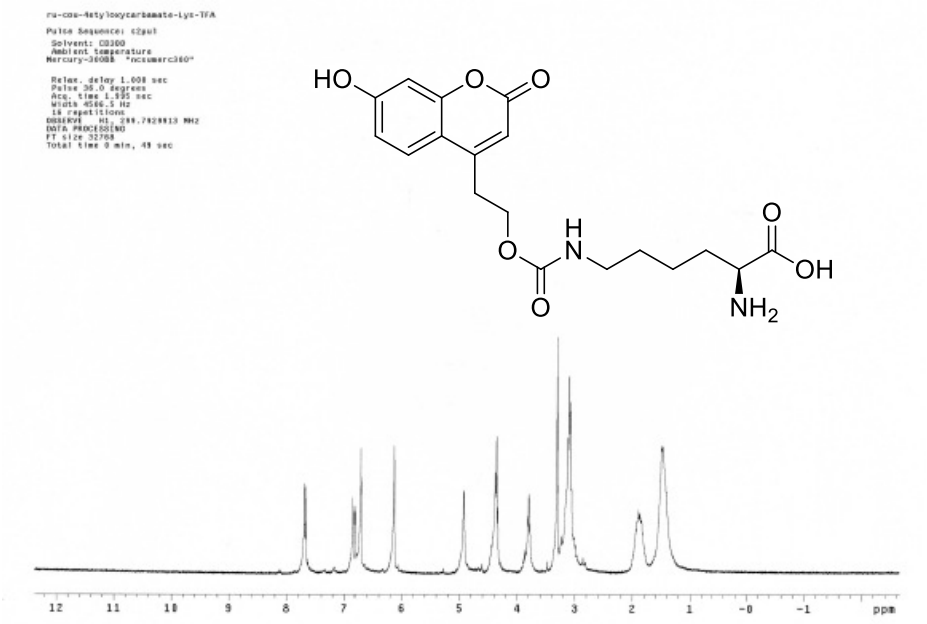


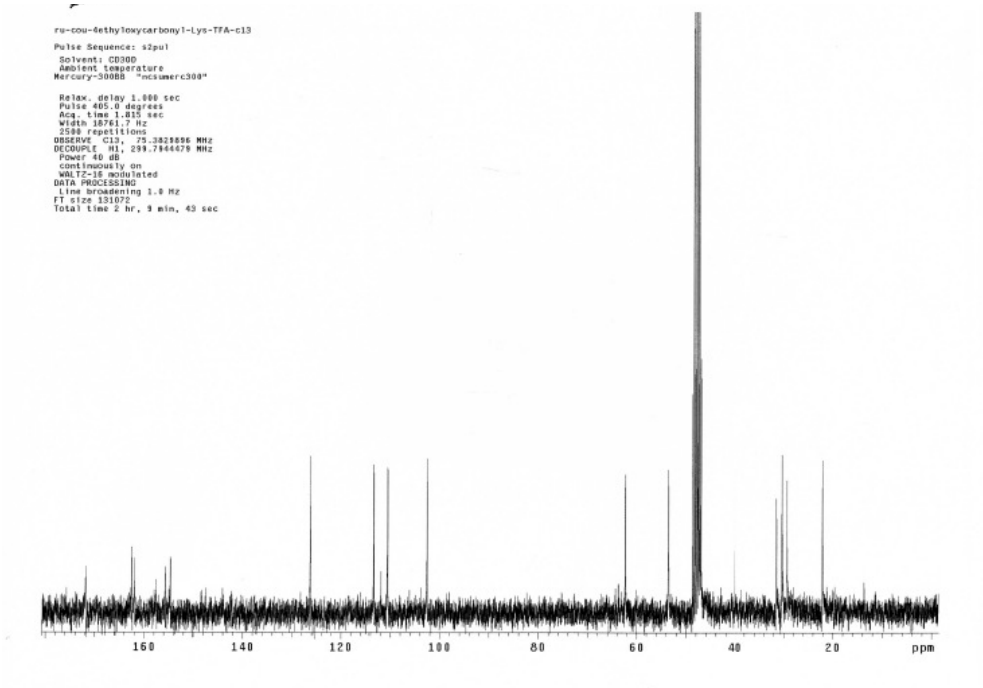
(*S*)-2-Amino-6-((((7-hydroxy-2-oxo-2H-chromen-4-yl)methoxy)carbonyl)amino)hexanoic acid (**1**).





(*S*)-2-Amino-6-((((7-hydroxy-2-oxo-2H-chromen-4-yl)methoxy)carbonyl)amino)hexanoic acid (**3**).





References

1. Bourbon, P.; Peng, Q.; Ferraudi, G.; Stauffacher, C.; Wiest, O.; Helquist, P., Synthesis, Photophysical, Photochemical, and Computational Studies of Coumarin-Labeled Nicotinamide Derivatives. *Journal of Organic Chemistry* **2012,** *77* (6), 2756-2762.
2. Campos-Toimil, M.; Orallo, F.; Santana, L.; Uriarte, E., Synthesis and vasorelaxant activity of new coumarin and furocoumarin derivatives. *Bioorganic & Medicinal Chemistry Letters* **2002,** *12* (5), 783-786.
3. Zhu, Q.; Uttamchandani, M.; Li, D.; Lesaicherre, M.; Yao, S., Enzymatic profiling system in a small-molecule microarray. *Organic Letters* **2003,** *5* (8), 1257-1260.
4. Montgomery, H.; Perdicakis, B.; Fishlock, D.; Lajoie, G.; Jervis, E.; Guillemette, J., Photo-control of nitric oxide synthase activity using a caged isoform specific inhibitor. *Bioorganic & Medicinal Chemistry* **2002,** *10* (6), 1919-1927.