Electronic Supplementary Material (ESI) for RSC Advances

Supporting Information for:

A photocleavable affinity tag for the enrichment of alkyne-modified biomolecules

Timo Koopmans[a], Frank Dekker[b], Nathaniel I. Martin[a,*]

[a]*Department of Medicinal Chemistry & Chemical Biology, Utrecht Institute for Pharmaceutical Sciences, University of Utrecht, David de Wied Building, Office: 5.64, Universiteitsweg 99, 3584 CG Utrecht, The Netherlands. Tel: +31 (0)6 1878 5274*

[b]*Department of Pharmaceutical Gene Modulation, Groningen Research Institute of Pharmacy, University of Groningen, Antonius Deusinglaan 1, 9713 AV Groningen, The Netherlands.*

*n.i.martin@uu.nl

Table of Contents

General Procedures

Reagents, solvents and solutions. All reagents employed were of American Chemical Society (ACS) grade or finer and were used without further purification unless otherwise stated. For anhydrous reactions, solvents were dried according to Perrin *et al.*[1] and Vogel.[2] Removal of solvent was performed under reduced pressure using a rotary evaporator, followed by evacuation (< 0.1 mm Hg) to constant sample weight. The known azido amine PEG spacer **2** was prepared by a combination of previously published procedures.[3,4] 3-Amino-3-(2-nitrophenyl)-propionic acid **3** was obtained from Acros and Boc-protected as reported by Bosques and Imperiali.[5] Fluorophosphonate probe **10** was prepared as described by Gillet and coworkers.[6]

Purification Techniques. All reactions and fractions from column chromatography were monitored by thin layer chromatography (TLC) using plates with a UV fluorescent indicator (normal SiO_2, Merck 60 F_{254}). One or more of the following methods were used for visualization: UV absorption by fluorescence quenching; iodine staining; phosphomolybdic acid:ceric sulfate:sulfuric acid:H_2O (10 g:1.25 g:12 mL:238 mL) spray; and 50% sulfuric acid spray. Flash chromatography was performed according to the method of Still *et al.*[7] using Merck type 60, 230-400 mesh silica gel.

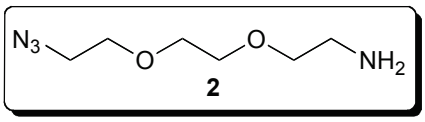
Instrumentation for Compound Characterization. High resolution mass spectrometry (HRMS) analysis was performed using an ESI-TOF LC/MS instrument. Optical rotations were recorded at 20 °C. ^{1}H NMR spectra were recorded at 300.1 MHz with chemical shifts reported in parts per million (ppm) downfield relative to tetramethylsilane (TMS). ^{1}H NMR data are reported in the following order: multiplicity (s, singlet; d, doublet; t, triplet; q, quartet; qn, quintet and m, multiplet), number of protons, coupling constant (*J*) in Hertz (Hz) and assignment. When appropriate, the multiplicity is preceded by br, indicating that the signal was broad. ^{13}C NMR spectra were recorded at 75.5 MHz with chemical shifts reported relative to $CDCl_3$ δ 77.0. ^{13}C NMR spectra were recorded using the attached proton test (APT) sequence. All literature compounds had ^{1}H NMR, and mass spectra consistent with the assigned structures.

References

1. D. D. Perrin, W. L. F. Armarego, D. R. Perrin. Purification of Laboratory Chemicals; 2nd ed.; Pergamon Press: New York, 1980.
2. A. Vogel. Vogel's Textbook of Practical Organic Chemistry; 4th ed.; Wiley and Sons Inc.: New York, 1978.
3. E. Klein, S. DeBonis, B. Thiede, D. A. Skoufias, F. Kozielski and L. Lebeau, *Bioorg Med Chem*, 2007, **15**, 6474-6488.
4. K. M. Bonger, R. J. van den Berg, L. H. Heitman, I. J. AP, J. Oosterom, C. M. Timmers, H. S. Overkleeft and G. A. van der Marel, *Bioorg Med Chem*, 2007, **15**, 4841-4856
5. C. J. Bosques and B. Imperiali, *J Am Chem Soc*, 2003, **125**, 7530-7531
6. L. C. Gillet, K. Namoto, A. Ruchti, S. Hoving, D. Boesch, B. Inverardi, D. Mueller, M. Coulot, P. Schindler, P. Schweigler, A. Bernardi and S. Gil-Parrado, *Mol Cell Proteomics*, 2008, **7**, 1241-1253.
7. W. C. Still, M. Kahn, A. J. Mitra. *J Org Chem*, 1978, **43**, 2923-2925.

Preparative Details and Analytical Data for Compounds 2-6

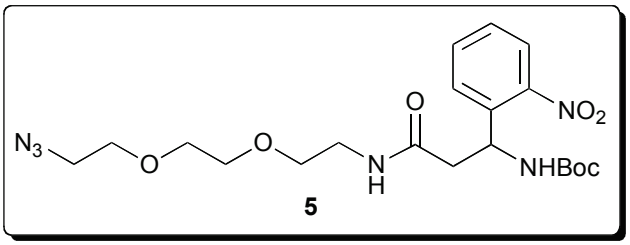
Compound 2: *2-(2-(2-azidoethoxy)ethoxy)ethanamine*



Triethylene glycol (7.51 g, 50 mmol) was dissolved in DCM (50 mL) after which *p*-toluenesulfonyl chloride (19.07 g, 100 mmol) was added and the mixture cooled to 0ºC. Powdered KOH (22.4 g, 400 mmol) was carefully added, not allowing the temperature above 10ºC and the mixture was stirred for 3 h at 0ºC. DCM (50 mL) and H_2O (100 mL) were then added and the organic layer separated and the water layer further extracted with DCM (2x 50 mL). The combined organic extracts were washed with H_2O (2x75 mL), dried over Na_2SO_4 and evaporated under vacuum to provide the intermediate ditosylate as a white solid (22.93 g, quant). The crude ditosylate (22.93 g, 50 mmol) was next dissolved in DMF (75 mL) and NaN_3 (13.0 g, 200 mmol) was added followed by TBAI (930 mg, 2.5 mmol, 5 mol%). The mixture was fitted to a condenser, heated to 80ºC and stirred overnight. The following morning DMF was evaporated under vacuum and the residue resuspended in Et_2O (100 mL). Insoluble material was removed by filtration over Hyflo after which Et_2O was removed under vacuum to yield the intermediate diazide as a yellow oil (10.03 g, quant). To a stirred solution of crude 1,2-bis(2-azidoethoxy)-ethane (10.03 g, 50 mmol) in Et_2O/1 N HCl/THF (5:5:1, 150 mL) a solution of triphenylphosphine (13.14 g, 50 mmol) in Et_2O (75 mL) was added dropwise over 3 h and the solution left to stir overnight. The following morning the solution was decanted from the precipitated triphenylphosphine oxide and the organic layer extracted with 4 N HCl (2x80 mL). The combined aqueous layers were then washed with Et_2O (4x100 mL). The pH of the aqueous layer was adjusted to 14 by direct addition of NaOH pellets. The aqueous layer was extracted with DCM (3x50 mL) and the combined organic extracts dried over Na_2SO_4. Following filtration and solvent evaporation under vacuum, the desired product was obtained as a pale yellow oil in >95% purity (7.16 g, 82%) providing analytical data in agreement with previously published values.[3,4]

Analytical data: ^{1}H NMR (300 MHz, $CDCl_3$) δ 3.67-3.55 (m, 6H), 3.50-3.45 (m, 2H), 3.37-3.30 (m, 2H), 2.84-2.81 (m, 2H), 1.38 (s, 2H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 73.8, 70.9, 70.5, 70.3, 50.9, 42.0; LRMS (ESI) Calcd for $C_6H_{15}N_4O_2$ $[M+H]^+$ 175.12, found 175.25.

Compound 5: tert-*butyl 3-(2-(2-(2-azidoethoxy)ethoxy)ethylamino)-1-(2-nitrophenyl)-3-oxopropylcarbamate*



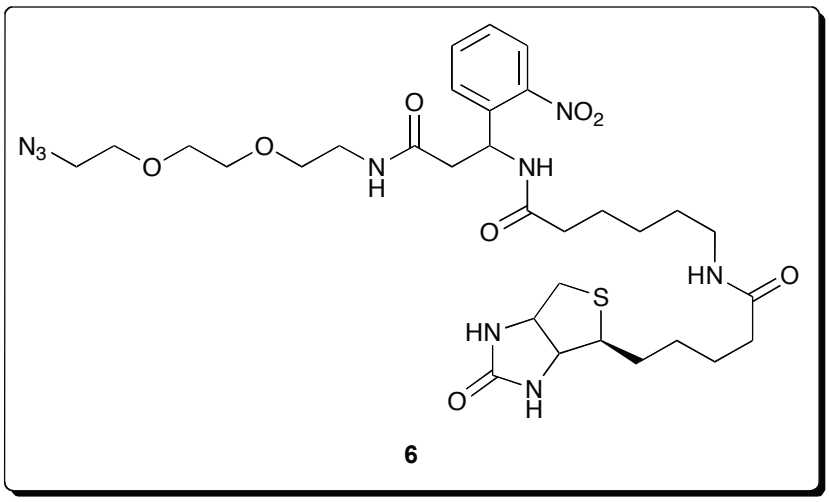
1-Amino(2-nitrophenyl)-3-propionic acid **3** (4.20 g, 20 mmol) was dissolved in 10% Et_3N in MeOH (30 mL). Boc_2O (8.73 g, 40 mmol) was added and the mixture stirred at 50ºC for 1.5 h at which point the solution became clear. Solvents were then evaporated under vacuum and the residue partitioned between 1 M HCl (50 mL) and EtOAc (100 mL). After separation, the aqueous layer was further extracted with EtOAc (2x75 mL) and the combined organic layers dried over Na_2SO_4, filtered, and evaporated under vacuum. The product was recrystallized from EtOAc/hexanes to afford 6.0 g (97%) of **4** giving analytical data in agreement with published values.[5]

Boc-protected intermediate **4** (1.0 g, 3.2 mmol) was dissolved in CH_2Cl_2 (30 mL) and HOBt (473 mg, 3.5 mmol) and BOP (1.57 g, 3.5 mmol) were added. After stirring for 5 minutes, azido amine **2** (610 mg, 3.5 mmol) and DIPEA (1.22 mL, 7.0 mmol) were added and the mixture

stirred overnight at room temperature. The following morning the reaction mixture was washed with 1M $KHSO_4$ (2x30 mL), saturated $NaHCO_3$ (2x30 mL) and brine (30 mL). The organic layer was dried over Na_2SO_4, filtered and evaporated. The residue was applied to a silica column and eluted with EtOAc/hexanes (2:1→4:1). Product containing fractions were pooled and evaporated to yield compound **5** as a light yellow, waxy solid (1.3 g, 87%).

Analytical data: R_f = 0.45 (4:1 EtOAc/hexanes); ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.94-7.92 (m, 1H), 7.66-7.54 (m, 2H), 7.40-7.35 (m, 1H), 6.74 (bs, 1H), 6.19 (bs, 1H), 5.52 (bs, 1H), 3.64-3.33 (m, 12H), 2.85-2.79 (m, 1H), 2.68-2.61 (m, 1H), 1.35-1.13 (m, 9H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 170.6, 155.2, 148.2, 138.1, 133.7, 128.9, 128.3, 125.1, 79.9, 70.6, 70.4, 70.3, 69.8, 50.8, 49.0, 41.0, 39.4, 28.5. HRMS (ESI) Calcd for $C_{20}H_{30}N_6O_7Na$ $[M+Na]^+$, 489.2074, found 489.2072.

Compound 6: ***N-(3-(3-(2-(2-azidoethoxy)ethoxy)propylamino)-1-(2-nitrophenyl)-3-oxopropyl)-6-(5-((3aS,4S,6aR)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamido)hexanamide***



6

Compound **5** (1.20 mmol, 560 mg) was deprotected by treatment with 15 mL of TFA/CH_2Cl_2 (1:1) for 1 hour at room temperature. Following solvent removal under vacuum (3 x co-evaporation with $CHCl_3$ to remove residual TFA) the residue was dissolved in DMF (10 mL) and cooled on ice. N-methylmorpholine (5.0 mmol, 546 μL) was added followed by NHS-LC-Biotin (1.0 mmol, 455 mg) as a solution in DMF (5 mL). The mixture was stirred for 1 hour at 0°C after which it was warmed to room temperature and left to stir for the evening. The following morning solvent was removed under vacuum and the residue applied directly to a silica column eluting with a gradient of 2-10% MeOH in CH_2Cl_2. Product-containing fractions were pooled and evaporated and the residue further triturated with CH_2Cl_2 followed by overnight cooling at 4ºC (to remove a trace baseline impurity that co-eluted with the product). The desired compound was thus obtained in pure form as translucent, pale yellow flakes (388 mg, 55%).

Analytical data: R_f = 0.33 (10% MeOH in CH_2Cl_2); ^{1}H NMR (300 MHz, DMSO) *diastereomers* δ 8.40 (d, 1H, *J* = 7.4 Hz), 7.88-7.83 (m, 2H), 7.73-7.65 (m, 2H), 7.60-7.57 (m, 1H), 7.48-7.43 (m, 1H), 6.40 (s, 1H), 6.34 (s, 1H), 5.53-5.47 (m, 1H), 4.30-4.26 (m, 1H), 4.12-4.08 (m, 1H), 3.59-3.47 (m, 4H), 3.38-3.33 (m, 4H), 3.18-2.92 (m, 5H), 2.82-2.77 (m, 1H), 2.66-2.48 (m, 3H), 2.19-2.14 (m, 1H), 2.04-1.99 (m, 3H), 1.65-1.08 (m, 14H); ^{13}C NMR (75 MHz, DMSO) δ 174.4, 171.8, 171.4, 168.8, 162.7, 148.3, 137.9, 133.2, 128.4, 128.1, 123.9, 69.58, 69.55, 69.2, 69.1, 61.0, 59.2, 55.4, 54.9, 50.0, 45.6, 38.5, 38.3, 38.2, 35.2, 35.1, 33.6, 28.9, 28.2, 28.0, 25.9, 25.3, 24.9, 24.2. HRMS (ESI) Calcd for $C_{31}H_{47}N_9O_8S$ $[M+H]^+$, 706.3347, found 706.3324.

Protein labeling procedures and enrichment protocols

Labeling of CALB with fluorophosphonate probe

High purity *Candida antarctica* lipase B (CALB) was obtained as a lyophilized powder from C-LEcta GmbH (Leipzig, Germany) and dissolved at a concentration of 18 mg/mL (corresponding to 75 μM) in 50 mM Tris-HCl at pH 8.0. A 1.0 mL aliquot of 75 μM CALB solution was treated with 1.5 mM fluorophosphonate probe **10** (by dilution of a 67x stock in DMSO). After incubation at room temperature for 1 hour, excess probe was removed by exchanging for fresh buffer using a 5000 Da molecular weight cut-off spin concentrator. The labeled CALB species **11** was characterized by ESI-TOF LC/MS indicating a high degree of label incorporation (figure S1).

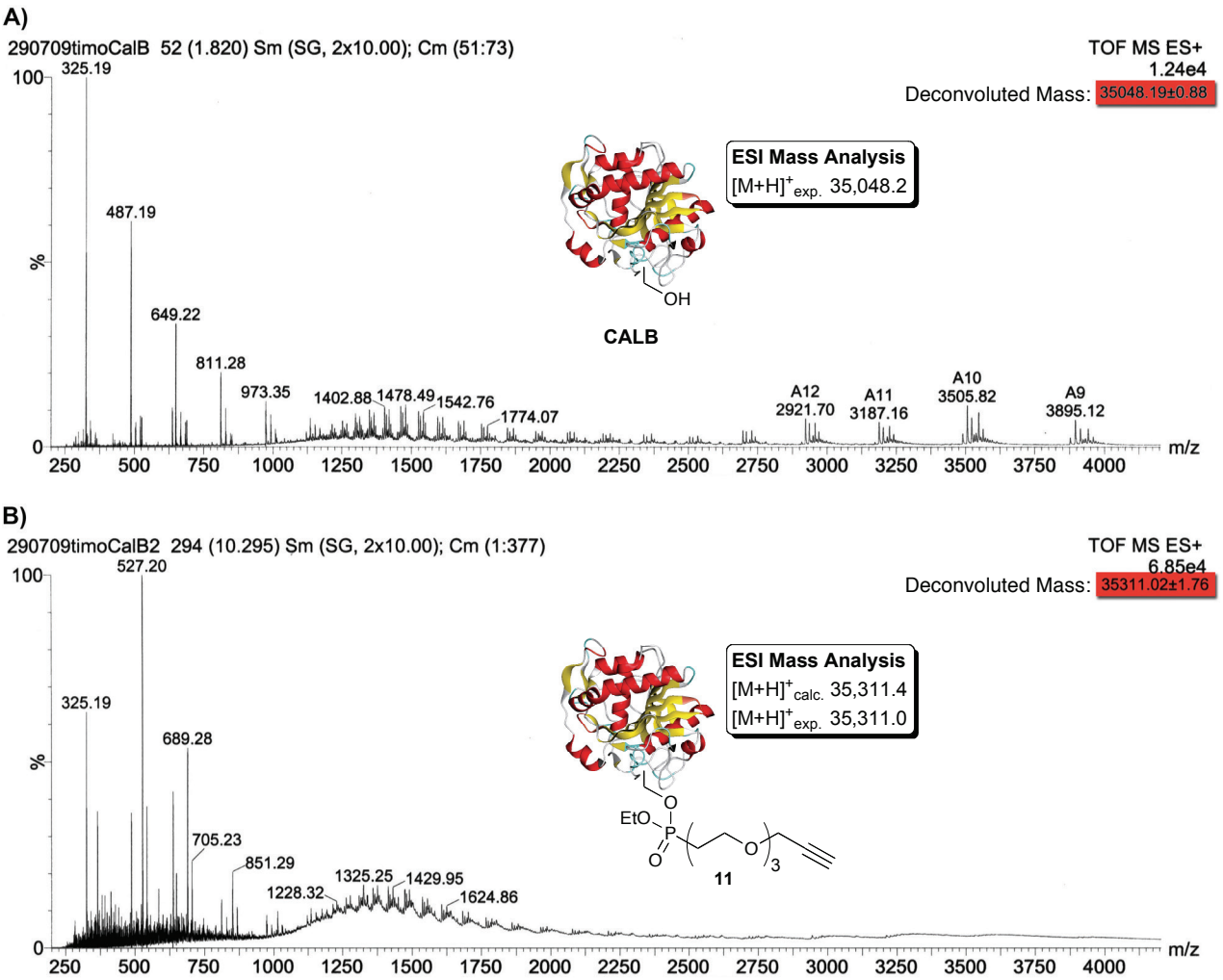


Figure S1. ESI-TOF-MS Characterization of **A)** CALB and **B)** fluorophosphonate probe labeled CALB **11** (samples analysed as solutions in 50:50 H_2O/acetonitrile containing 0.2% formic acid). Multiply charged species measured are noted on the spectrum (e.g. A10, A11, A12 corresponding to the $[M+10H]^{10+}$, $[M+11H]^{11+}$, and $[M+12H]^{12+}$ respectively) and were used in deconvolution of the experimental mass value.

CuAAC Ligation of photocleavable enrichment tag to alkyne-modified CALB

Alkyne-modified CALB species **11** was treated with photocleaveble enrichment tag **6** at a concentration of 5 mM (by dilution of a 50x stock in DMSO), followed by addition of 10 mM TCEP (from 10x stock in H_2O) and 2 mM TBTA ligand (from 16x stock in DMSO/*t*-BuOH 1:4) giving a final concentration of 5% *t*-BuOH. Samples were vortexed and 10 mM $CuSO_4$ (from 100x stock in H_2O) was added. Samples were vortexed again and reacted for 1 h at room temperature. The reaction mixture was split up in 500 µL aliquots and 125 µL ice-cold TCA/Acetone (1:1 w/v) was added to each aliquot. The mixtures were left on ice for 30 min. and then centrifuged 5 min. at 10,000x*g* at 4°C to pellet the protein. Pelleted protein was washed two times by adding 600 µL cold acetone, resuspending the pellet by sonication and repelleting. The protein was finally resuspended in a total volume of 1 mL 100 mM PBS. The click reaction product **12** was evaluated and deemed to be essentially quantitative by ESI-TOF LC/MS analysis (figure S2).

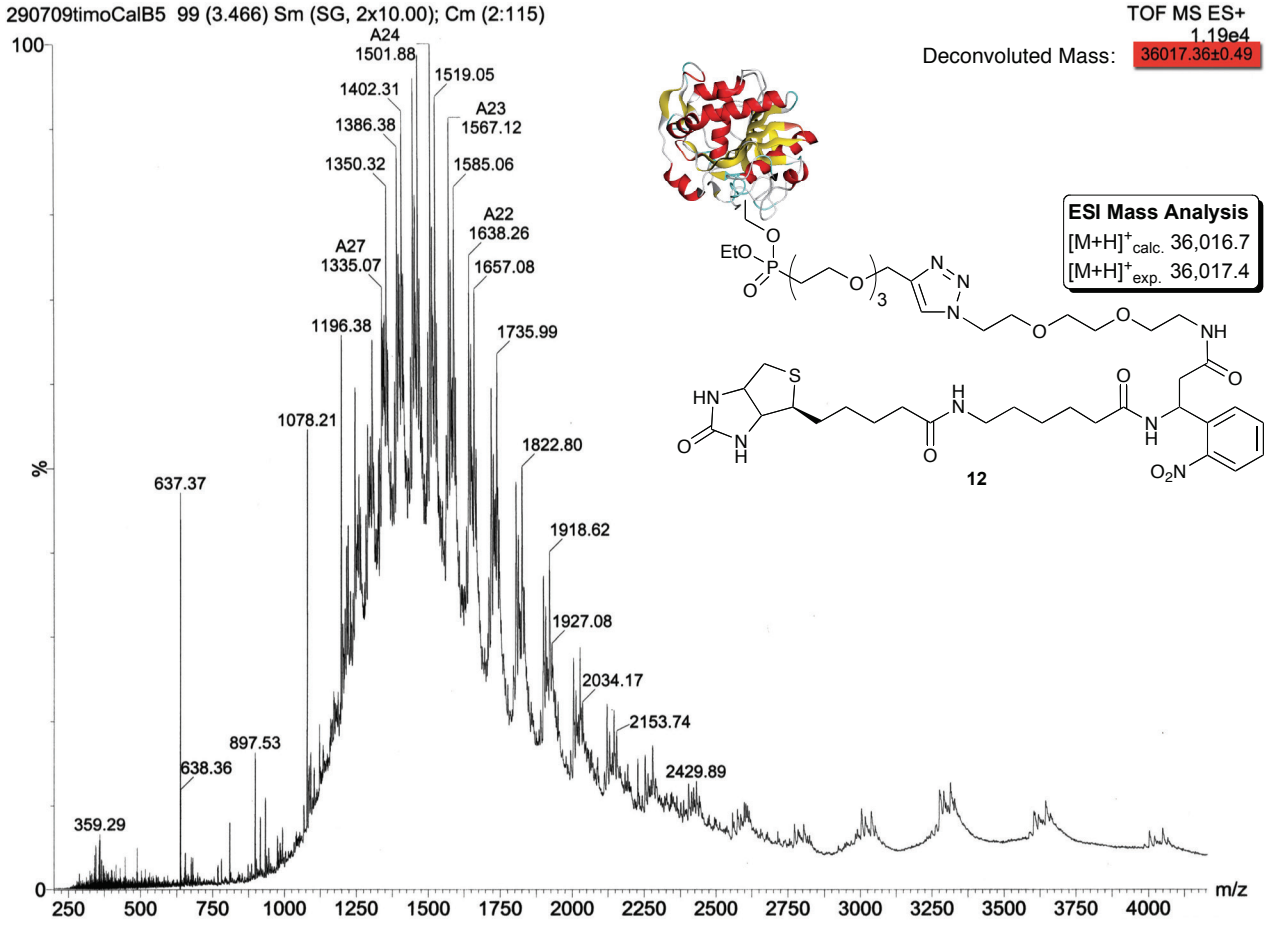


Figure S2. ESI-TOF-MS Characterization of CuAAC ligation product **12** (samples analysed as solutions in 50:50 H_2O/acetonitrile containing 0.2% formic acid). Multiply charged species measured are noted on the spectrum (e.g. A22, A23, A24 corresponding to the $[M+22H]^{22+}$, $[M+23H]^{23+}$, and $[M+24H]^{24+}$ respectively) and were used in deconvolution of the experimental mass value.

Binding of Biotin Tagged CALB to Streptavidin Resin and Photolytic Elution

A 100 µL aliquot of a 50% streptavidin-agarose slurry (Pierce) was pre-washed and equilibrated by 3x supsension and centrifugation with 400 µL PBS buffer. The resin was then treated with a 200 µL solution containing biotin labeled CALB **12** at an approximate concentration of 75 µM in PBS and incubated for 2 hours 4°C. The resin was then centrifuged for 1 minute at 5,000x*g* and the supernatant stored (for analysis as "post resin" Lane C, Figure 3 in main article). The resin was washed with 400 µL PBS buffer (4x) and divided into two aliquots for direct comparison of UV cleavage versus cleavage by SDS boil. One sample aliquot was directly irradiated by shining 366 nm UV-light* over the open top of the sample tube for 2 h on ice with regular agitation (short vortex every 20 minutes). The resin was then centrifuged at 5,000x*g*, and 22.5 µL of supernatant removed and mixed with 7.5 µL of standard 4x sample buffer (containing 10% SDS) to provide a 30 µL sample used for subsequent SDS-PAGE analysis (Lane E, Figure 3 in main article). For comparison, the other half of the resin was directly treated with 50 µL of 4x sample buffer (containing 10% SDS) and heated at 100°C for 10 minutes (Lane D, Figure 3 in main article).

Protein recovery from E. Coli lysate

A 100 µL aliquot of biotin labeled CALB **12** (approximate concentration 75 µM) in PBS was pre-mixed with a 100 µL aliquot of *E. coli* lysate (±1 mg/mL total protein). A 22.5 µL aliquot of this mixture was removed for analysis (Lane C, Figure S3) and the remained was then incubated with streptavidin-agarose resin (Pierce) for 2 h at 4°C. After incubation, the resin was centrifuged for 1 minute at 5,000x*g* and the supernatant removed and stored for analysis ("post resin" Lane D, Figure S3). The resin was washed with 400 µL PBS buffer (4x) and divided into two aliquots for direct comparison of UV cleavage versus cleavage by SDS boil. One sample aliquot was directly irradiated by shining 366 nm UV-light over the open top of the sample tube for 2 h on ice with regular agitation (short vortex every 20 minutes). The resin was then centrifuged at 5,000x*g*, and 22.5 µL of supernatant removed and mixed with 7.5 µL of standard 4x sample buffer (containing 10% SDS) to provide a 30 µL sample used for subsequent SDS-PAGE analysis (Lane F, Figure S3). For comparison, the other half of the resin was directly treated with 50 µL of 4x sample buffer (containing 10% SDS) and heated at 100°C for 10 minutes (Lane E, Figure S3).

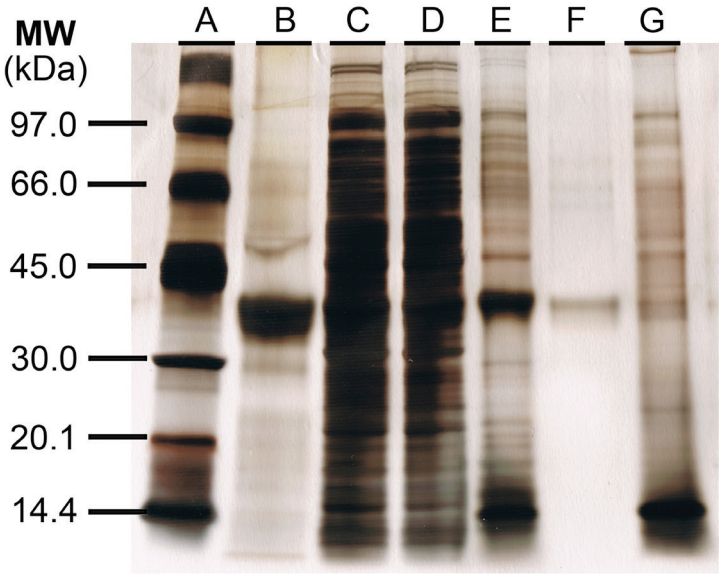
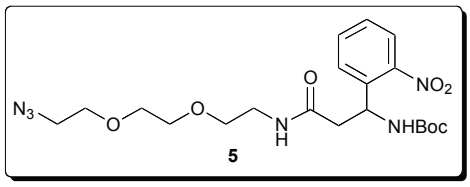


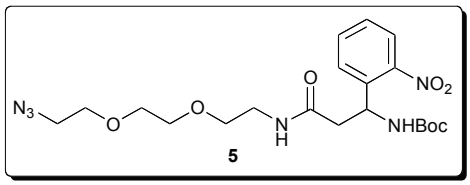
Figure S3. Enrichment of CALB from *E. coli* lysate using photocleavable enrichment tag **6.** SDS-PAGE analysis: **Lane A)** MW ladder; **B)** solution of **12**; **C)** solution of **12** mixed with *E. coli* lysate before resin binding **D)** solution of **12** mixed with *E. coli* lysate post resin binding; **E)** elution by SDS boil; **F)** elution by irradiation at 366 nm; **G)** control lane (fresh resin treated with non-biotinylated CALB species **11** followed by SDS boil).

*A variety of UV light sources were compared (conventional TLC lamp, trans-illuminator, and pen-light) each resulting in similar levels of photocleavage.

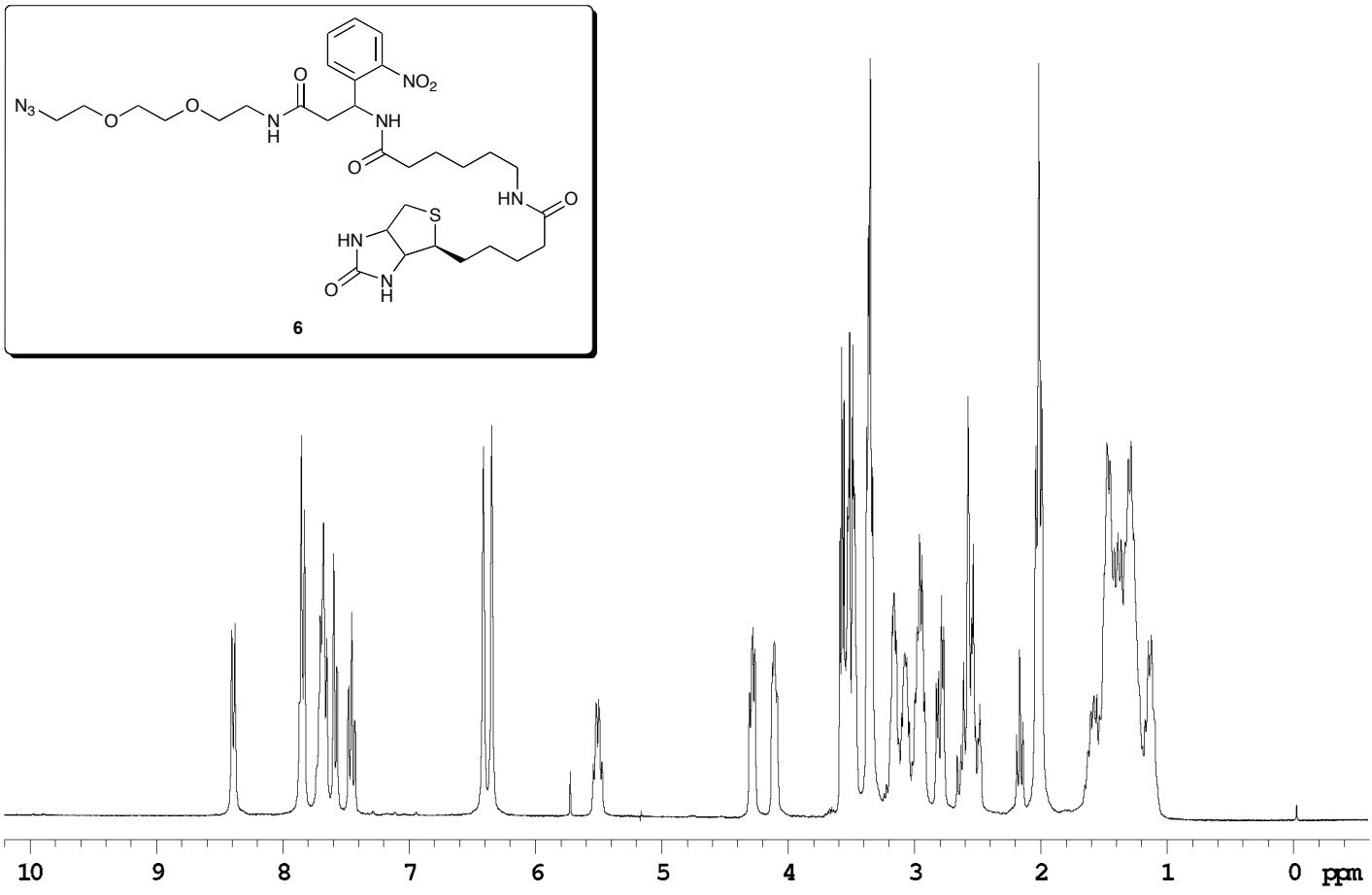
Compound 5: ^{1}H NMR (300 MHz, $CDCl_3$)



Compound 5: ^{13}APT NMR (75 MHz, $CDCl_3$)



Compound 6: ^{1}H NMR (300 MHz, DMSO-d$_6$)



Compound 6: ^{13}C APT NMR (75 MHz, DMSO-d$_6$)