Supporting Information

Direct Acylation of Carrier-Proteins with Functionalized β-lactones

Jon W. Amoroso, Lawrence S. Borketey, Gitanjeli Prasad, and Nathan A. Schnarr*
Department of Chemistry, University of Massachusetts, 710 N. Pleasant Street, Amherst, Massachusetts 01003

schnarr@chem.umass.edu

Table of Contents

1. General

NMR spectra were recorded on a Bruker Advance 400, chemical shift values are reported in ppm on the δ scale relative to TMS (δ = 0.00) as an internal standard. High resolution FAB-MS were performed on a JEOL MStation JMS700. All reactions were carried out in flame-dried glassware with Ar inlet and Teflon-coated stir bar. Dichloromethane was distilled from calcium hydride. Other reagents were used as received from commercial sources.

2. Synthesis

5-hexynal

20 mL of anhydrous dichloromethane (DCM) were cooled in a dry-ice/acetone bath and oxalyl chloride (13.0 mmol) was added by syringe and stirred for several minutes. DMSO (13.0 mmol) with 10% v/v DCM was added dropwise by syringe. After gas evolution ceased, the mixture was stirred for

five minutes then 5-hexyn-1-ol (5.20 mmol) was added by syringe and stirred for five minutes. Triethylamine (26.01 mmol) was added by syringe and the mixture stirred for 15 minutes and warmed to RT. The product was washed with 0.5 M HCl and ½ saturated $NaHCO_3$, and dried over $MgSO_4$. The product was purified by flash chromatography over silica gel using 9:1 hexanes:EtOAc mobile phase. 5-hexynal: 0.4200 g, 4.37 mmol of product, 84%, colorless oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ: 9.811 (s, 1H), 2.619 (t, J = 7.05 Hz, 2H), 2.278 (dt, J = 6.70 Hz, 2.45 Hz, 2H), 1.999 (t, J = 2.38 Hz, 1H), 1.858 (quint J = 7.00 Hz, 2H)

^{13}C NMR (100 MHz, $CDCl_3$) δ: 201.719, 83.188, 69.386, 42.530, 20.809, 17.772

General procedure 1: aluminum-catalyzed generation of β-lactones from aldehydes and acid chlorides[1].

A suspension of anhydrous $AlCl_3$ (0.1 eq) and $AgSbF_6$ (0.3 eq) in DCM cooled to -60°C in an ethanol/ethylene glycol/dry ice bath, was added by syringe: N,N'-diisopropylethylamine (DIEA) (1.5 eq), acid chloride (1.5 eq) and aldehyde (1 eq). The mixture was stirred for five hours at -60°C then filtered through a silica plug. Then product was purified by flash chromatography over silica gel with a hexanes:ethyl acetate mobile phase.

General procedure 2: Nucleophile-catalyzed production of β-lactones from aldehydes and acid chlorides[2]

For some reactions DABCO was substituted for the quinine catalyst, owing to the similar nature of the nucleophillic substituent.

To a solution of trimethylsilyl quinine (TMSq[3]) (0.1 eq) or (DABCO) (0.2 eq) and $LiClO_4$ (1.0 eq) in 30 mL of DCM and 10 mL of Et_2O cooled to -78°C in a dry ice-acetone bath, was added via syringe DIEA (3.0 eq), and aldehyde (1 eq). Acid chloride (1 - 2 eq diluted in DCM) was added dropwise over *ca*. 3 hours. The mixture was stirred for 8 hrs at -78°C then diluted with an equal volume of ether and allowed to come to room temperature. The mixture was filtered through a silica plug then purified by flash chromatography over silica gel with a hexanes:ethyl acetate mobile phase.

4-(pent-4-yn-1-yl)oxetan-2-one (**1**): General Procedure 1

$AlCl_3$ (0.0427 mmol), $AgSbF_6$ (0.1281 mmol), acetyl chloride (0.5976 mmol), DIEA (0.6656 mmol), 5-hexynal (0.4300 mmol). **3**: 0.2389 mmol (56%), pale yellow oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ: 4.567 – 4.539 (m, 1H), 3.529 (dd, J = 16.76 Hz, 1.76 Hz, 1H), 3.129 (dd, J = 16.29 Hz, 1.67, Hz, 1H), 2.286 (dt, J = 6.61 Hz, 1.46 Hz, 2H), 2.001 (t, J = 2.4, 1H), 1.720 – 1.644 (m, 4H).

^{13}C NMR (100 MHz, $CDCl_3$) δ: 168.101, 83.236, 70.785, 69.355, 43.029, 33.590, 23.867, 17.982.

Hi-res FAB MS calculated m/z for $C_8H_{11}O_2$ $[M+H]^+$ = 139.075356. Observed m/z = 139.0766.

4-propyloxetan-2-one (**3**): General Procedure 1

$AlCl_3$ (0.99 mmol), $AgSbF_6$ (2.9 mmol), acetyl chloride (11 mmol), DIEA (11), butanal (9.8 mmol). **2**: 8.134 mmol (83%), colorless oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ: 4.530 (dt, J = 2.96 Hz, 5.76 Hz, 1H), 3.522 (dd, J = 16.31 Hz, 5.78 Hz, 1H), 3.067 (dd, J = 16.29 Hz, 4.28 Hz), 1.1.871 – 1.838 (m, 1H), 1.833 – 1.745 (m, 1H), 1.56 – 1.36 (m, 1H), 0.986 (t, J = 7.40 Hz)

^{13}C NMR (100 MHz, $CDCl_3$) δ: 166.477, 69.238, 41.036, 34.826, 16.429, 11.802

4-(4-nitrophenyl)oxetan-2-one (**4**): General Procedure 2

DABCO (0.6 mmol), DIEA (9 mmol), $LiClO_4$ (3 mmol), *p*-nitrobenzaldehyde (3.0 mmol), acetyl chloride (3.0 mmol). **N**: 1.35 mmol (45%), pale yellow oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ: 8.30 (d, J = 8.88, 2H), 7.60 (d, J = 8.46 Hz, 2H), 5.63 (dd, J = 6.3 Hz, J = 4.5 Hz, 1H), 4.04 (dd, J = 16.4, J= 6.3, 1H), 3.44 (dd, J = 16.5, J = 4.5, 1H)

^{13}C NMR (100 MHz, $CDCl_3$) δ: 166.51, 148.42, 144.40, 126.50, 124.36, 69.51, 46.90

3-methyl-4-propyloxetan-2-one (**5**): General Procedure 2

TMSq (0.3 mmol), DIEA (7.5 mmol), $LiClO_4$ (3.0 mmol), butyraldehyde (3.0 mmol), propionyl chloride (3.0 mmol). **N**: 1.68 mmol (56%), colorless oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ: 4.60 – 4.55 (m, 1H), 3.79 – 3.17 (m, 1H), 1.80 – 1.69 (m, 1H), 1.68 – 1.61 (m, 1H), 1.59 – 1.49 (m, 1H), 1.47 – 1.38 (m, 1H), 1.28 (d, J = 7.8 Hz, 1H), 1.00 (t, J = 7.4 Hz, 1H)

^{13}C NMR (100 MHz, $CDCl_3$) δ: 172.86, 75.56, 47.22, 32.00, 18.79, 13.79, 8.06

Sulforhodamine-B Azide (**SrB-A**)

3-azido-1-aminopropane[4] (0.22 mmol) and triethylamine (0.5 mmol) was dissolved in 1 mL of 5:1 DCM:DMF and cooled to 0°C in an ice-water bath. Sulforhodamine B sulfonyl chloride [mixture of

ortho- and *para*- sulfonyl chloride isomers] (0.20 mmol) was added portion-wise over *ca* 30 min and stirred overnight at room temperature. The product was purified by flash chromatography over silica gel with 90:5:5 DCM : acetonitrile : methanol mobile phase. The *ortho* isomer was distinguished by its reversible color change in in pH 9.0 buffer[5]. *Para* isomer (**SrB-*p*A)**: 17%, *ortho* isomer (**SrB-*o*A**) 14%. Srb-*o*A was found to perform best in the click reaction and was utilized for labeling experiments.
NMR spectra are of poor quality due to long relaxation times. Product is pure by LC-MS.
Hi-res FAB MS calculated m/z for $C_{30}H_{37}N_6O_6S_2$ $[M+H]^+$ = 641.221049. Observed m/z = 641.2216.

3. Plasmid Construction

pCAD01 was prepared from an *E. coli* optimized synthetic construct (DNA2.0, Menlo Park, CA). The synthetic gene was excised from the shipping vector via flanking *NdeI* and *NotI* restriction sites and ligated into pET21b.

4. Protein Expression and Purification

SpnB was expressed from pCAD01. KSAT3 and KSAT6 were expressed from pAYC02 and pAYC11[6] respectively. ACP2 was expressed from pNW6[7], ACP3 was expressed from pVYA05[8]. Apo SpnB and ACPs were expressed in *E. coli* BL-21 cells and Holo SpnB and ACPs were expressed in *E. coli* BAP-1[9]. pCAD01, pAYC02, and pAYC11 contain ampicillin resistance vectors, pNW6 and pVYA05 contain kanamycin resistance vectors.

General Procedure for Protein Expression and Isolation

Bacteria were grown in 1 L shake cultures of LB-antibiotic media at 37°C in a New Brunswick Scientific Excella E24 Incubator Shaker until the OD600 was between 0.6 and 0.8. Overexpression was induced with 200 μL of 1 M IPTG and carried out at 18°C for 18 hours, after this point all work was carried out at 4°C. Cells were pelleted by spinning at 3000 RPM for 10 minutes in a Sorvall RC6 Plus with a FiberLite F21S-8x50 rotor and resuspended in 50 mL of lysis buffer (20 mM Tris-HCl, 150 mM NaCl, 1 mM Na_2EDTA, 1 mM EGTA, 1% Triton, 2.5 mM sodium pyrophosphate, 1 mM β-glycerophosphate, 1 mM sodium Na_3VO_4, 1 μg/mL leupeptin, pH 7.5). Cells were lysed using a Misonics ultrasonic converter without microtip, amplitude 30 for five 30 second intervals with a 60 second cool down period between each. Lysed cells were spun at 10,000 rpm and with a FiberLite F10S-

6x500 Y rotor for 60 minutes. The lysate was equilibrated with 3 mL of PerfectPro Ni-NTA bead slurry for 60 minutes by stirring with a PTFE-coated stir bar at 60 RPM for 60 minutes. The lysate was then poured into a 15 mL column and the supernatant eluted. The column was then washed with two 15 mL portions of wash buffer (50 mM phosphate, 300 mM NaCl, 20 mM imidazole, pH 8.0), and eluted with 3 mL of elution buffer (50 mM phosphate, 300 mM NaCl, 250 mM imidazole, pH 8.0). The purified protein was loaded into an Amicon Ultra centrifugal concentrator and diluted to 15 mL with storage buffer (100 mM Tris, 1 mM EDTA, 1 mM dithioerythritol, 10% glycerol pH 8) and spun at 3000 rpm an Eppendorf Centrifuge 5810 R with swinging-bucket rotor. Dilution and filtration was repeated a total of three times. Protein concentration determined by Bradford assay, average concentration was approximately 500 μM. Proteins were flash frozen and stored at -80°C until use.

5. Loading Proteins with β-lactones

Labeling reactions for gel assays were performed at 20 μL total volume, reactions for LC-MS were performed at 50 μL total volume. Final concentrations reported in procedure.

Loading ACPs

Phosphate buffer (100 mM, pH 7), tris(2-carboxyethyl)phosphine (2.5 mM), and protein (0.025 mM) were reacted at ambient temperature for 15 min. Then β-Lactone was added to the appropriate concentration (1, 5, 10, 20, 50, or 75x with respect to protein for saturation experiments, 10 or 50x with respect to protein for loading experiments) and the mixture reacted at ambient temperature for 60 min.

Loading KSATs

Phosphate buffer (100 mM, pH 7), tris(2-carboxyethyl)phosphine (2.5 mM), and KSAT (0.025 mM) were reacted at ambient temperature for 15 min, then β-Lactone (2.5 mM) was added (1, 5, 10, 20, 50, or 75x with respect to protein for saturation experiments, 10 or 50x with respect to protein for loading experiments) and the mixture incubated at ambient temperature for 60 min.

Loading SpnB

Phosphate buffer (100 mM, pH 7), tris(2-carboxyethyl)phosphine (2.5 mM), and SpnB (0.015 mM) were reacted at ambient temperature for 15 min, then β-Lactone was added (10, 50, or 75x with respect to protein) and the mixture incubated at ambient temperature for 60 min.

Acyl Transfer

Phosphate buffer (100 mM, pH 7), tris(2-carboxyethyl)phosphine (2.5 mM), and ACP (0.025 mM) were reacted at ambient temperature for 15 min. Then β-Lactone (10 equivalents) was added and the mixture reacted at ambient temperature for 60 min. Then KSAT was added and the mixture incubated at ambient temperature for 60 min

6. Chromophore Attachment

The reaction was carried out at 25 μL, final concentrations reported. **SrB-oA**, (2x alkyne concentration), sodium ascorbate (1 mM), and copper (II) sulfate (1 mM) were added to samples which had been labeled with **1**. The reaction was performed at ambient temperature for 60 minutes.

7. Gel Assay

Labeled samples were diluted to 35 μL with gel-loading buffer. ACPs and KSATs were separated by 12.5% SDS-PAGE with 5% stacking gel (100 V, 50 mA, 135 min). SpnB was separated by 4 – 20% gradient HEPES-PAGE (100 V, 50 mA, 90 min). Gels were developed in 10% acetic acid to visualize **Srb-oA**. **Srb-oA** labeled proteins were imaged on a BioDoc-It Imaging System with UV-transilluminator. Total protein was stained using GelCode Blue.

8. Proteolysis

Promega Sequencing Grade Modified Trypsin was added to prepared samples so that the final trypsin:ACP ratio was 1:50 (w/w). The mixture was incubated at 30°C for 18 hours. Digestion was quenched by addition of an equal volume of 10% formic acid. Digests were flash frozen in $N_{2(l)}$ and stored at -20°C until analysis.

9. LC-MS

Separation was performed with a Waters 1525 system. The gradient employed was A = water + 0.1% formic acid, B = acetonitrile + 0.1% formic acid, 5-95% B over 60 min with a Vydac 218TP C18 5u column (4.6 x 150 mm). Mass spectra were acquired with a Waters Micromass ZQ mass detector in EI+ mode: Capillary voltage = 3.50 kV, cone voltage = 30 V, extractor = 3 V, RF lens = 0.0 V, source temp = 100°C, desolvation temp = 200°C, desolvation gas = 300 L/hr, desolvation gas = 0.0 L/hr The system was operated and spectra were processed using the Waters Empower software suite.

10. Supporting References

1. Nelson, S. G.; Wan, Zonghui.; Peelen, T. J.; Spencer, K. L. *Tetrahedron Letters*. **1999**, *40*, 6535

2. Zhu, C.; Shen. X.; Nelson. S. G. *Journal of the American Chemical Society*. **2004**, *126*, 5352 - 5353

3. Calter. M. A. *Journal of Organic Chemistry*. **1996**, *61*, 8006 - 8007

4. Carboni, B.; Benalil, A.; Vaultier. M. *Journal of Organic Chemistry*, **1993**, 58*,* 3736 – 3741

5. Corrie, J. E. T.; Davis. C. T.; Eccelston. J. F. *Bioconjugate Chemistry*, **2001**, *12*, 186 – 194

6. Chen. A. Y.; Schnarr. N. A.; Kim. C.; Cane. D. E.; Khosla. C. *Journal of the American Chemical Society*. **2006**, *128*, 3067 – 3074

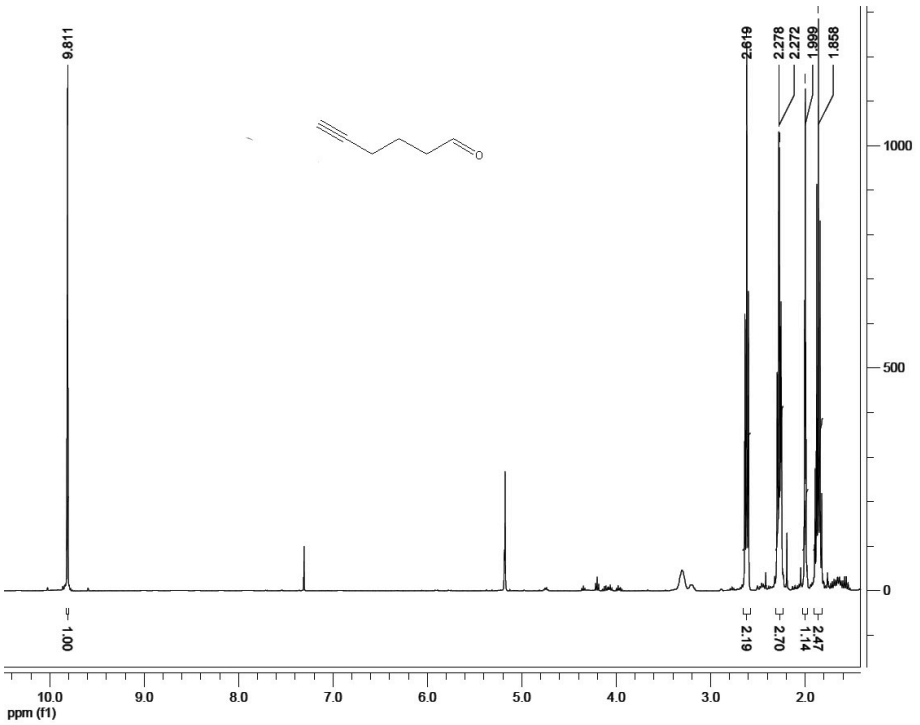
7. Wu. N.; Cane. D. E.; Khosla. C. *Biochemistry*, **2002** *41,* 5056 – 5066

8. Kim. C.; Alekseyev. V. Y.; Chen. A. Y.; Tang. Y.; Cane. D. E.; Khosla. C. *Biochemistry*, **2004** *43*, 13892 – 13898

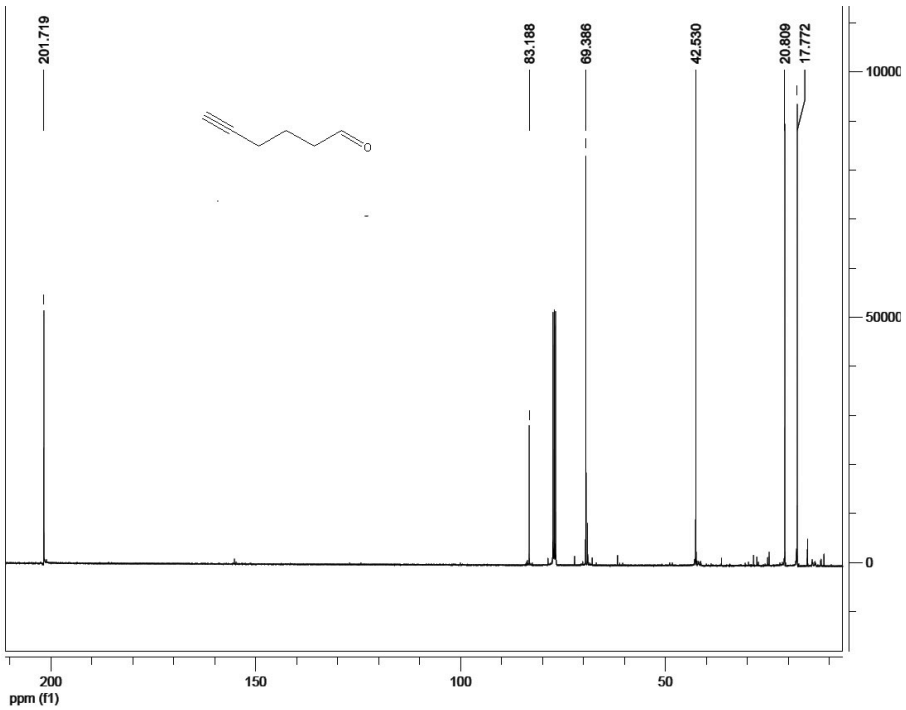
9. Pfeifer, B. A.; Admiraal. S. J.; Gramajo. H.; Cane. D. E.; Khosla. C. *Science.* **2001**, 291, 1790 - 1792

11. Chromatorgams and Spectra

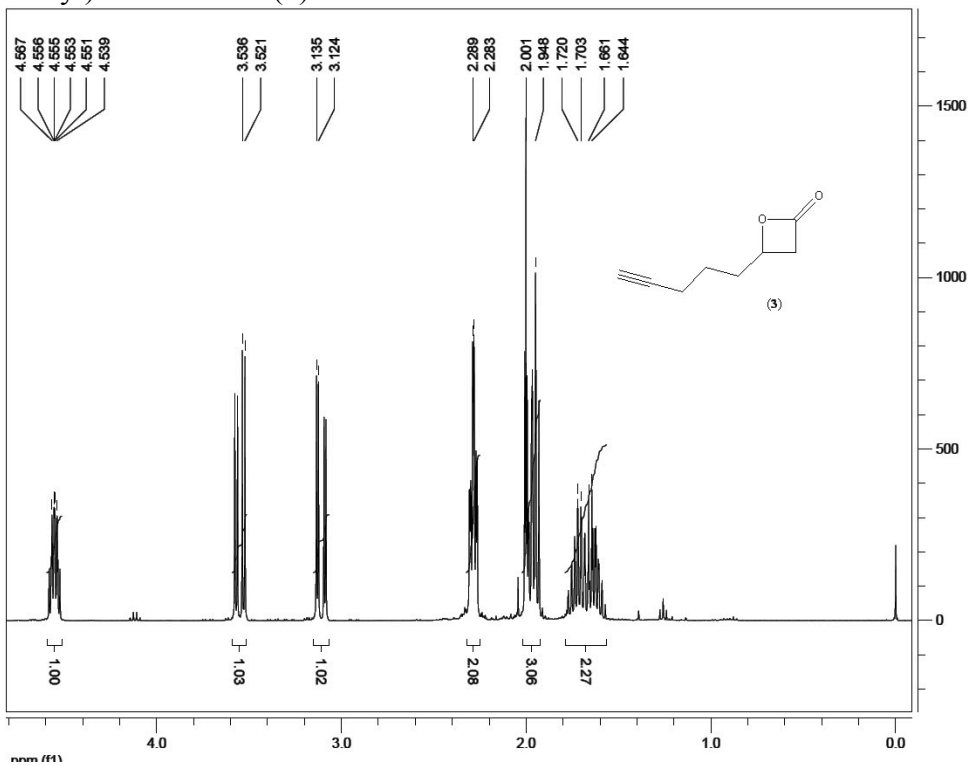
S1: 5-hexynal ^{1}H NMR



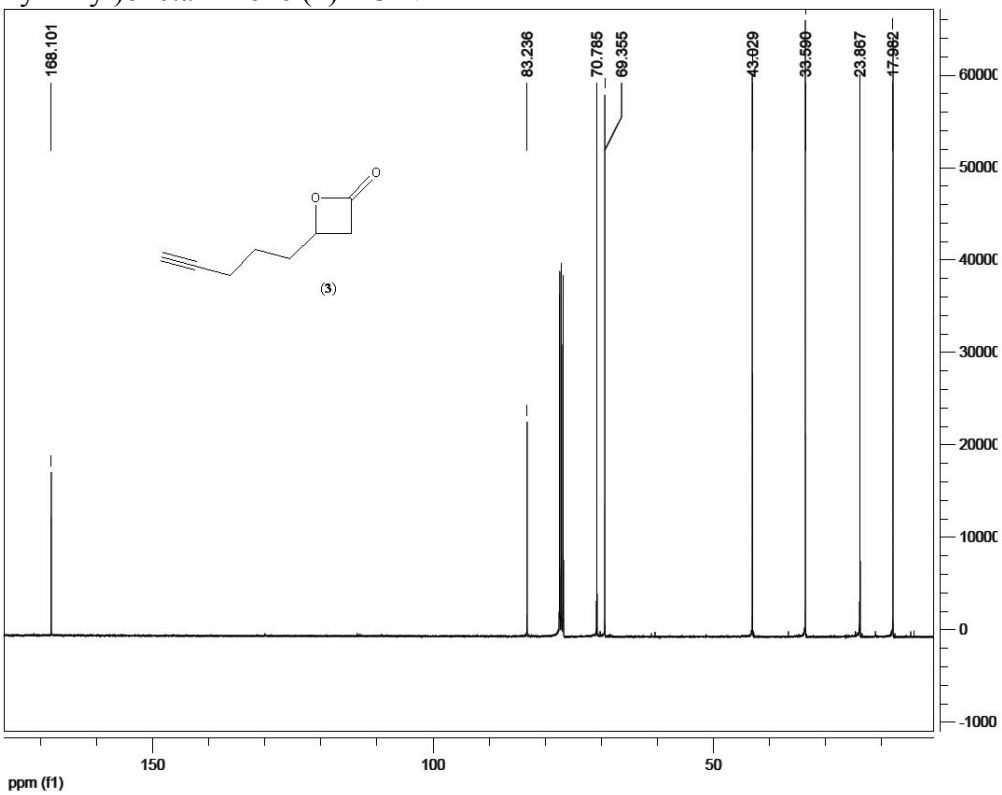
S2: 5-hexynal ^{13}C NMR



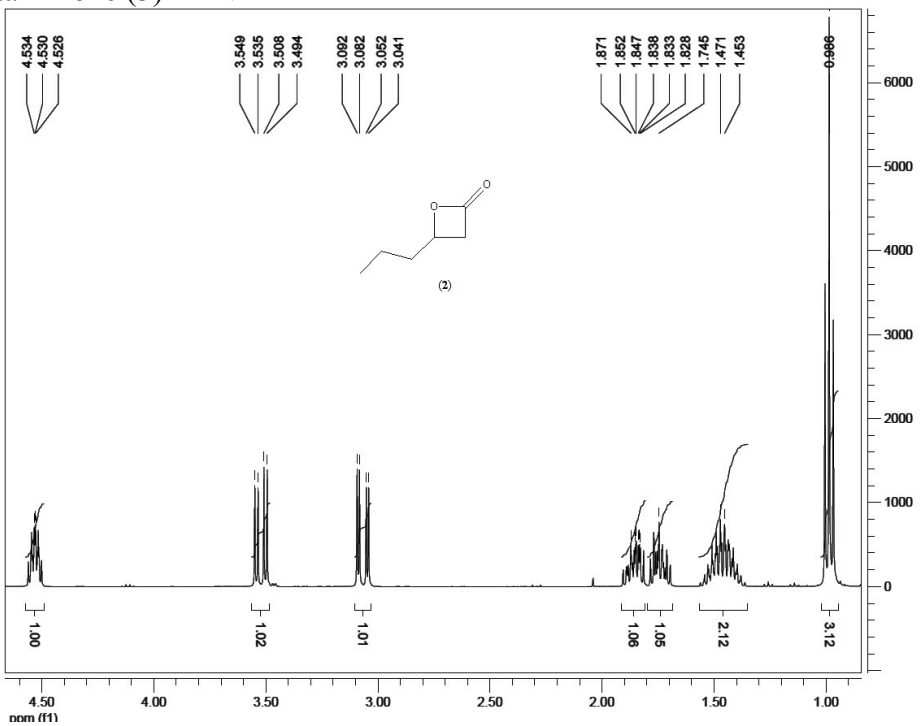
S3: 4-(pent-4-yn-1-yl)oxetan-2-one (**1**) ^{1}H NMR



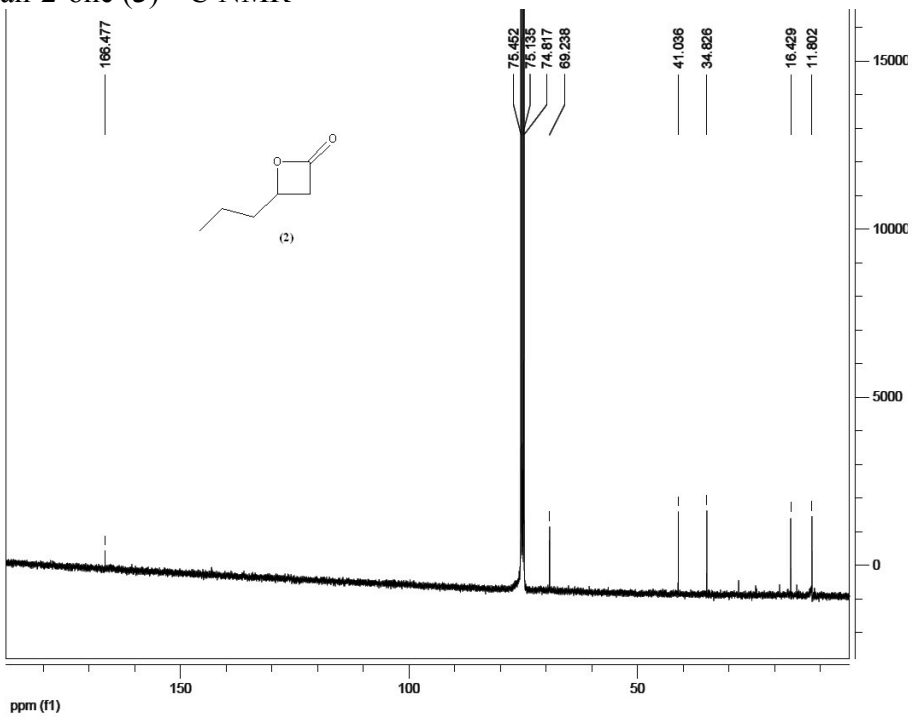
S4: 4-(pent-4-yn-1-yl)oxetan-2-one (**1**) ^{13}C NMR



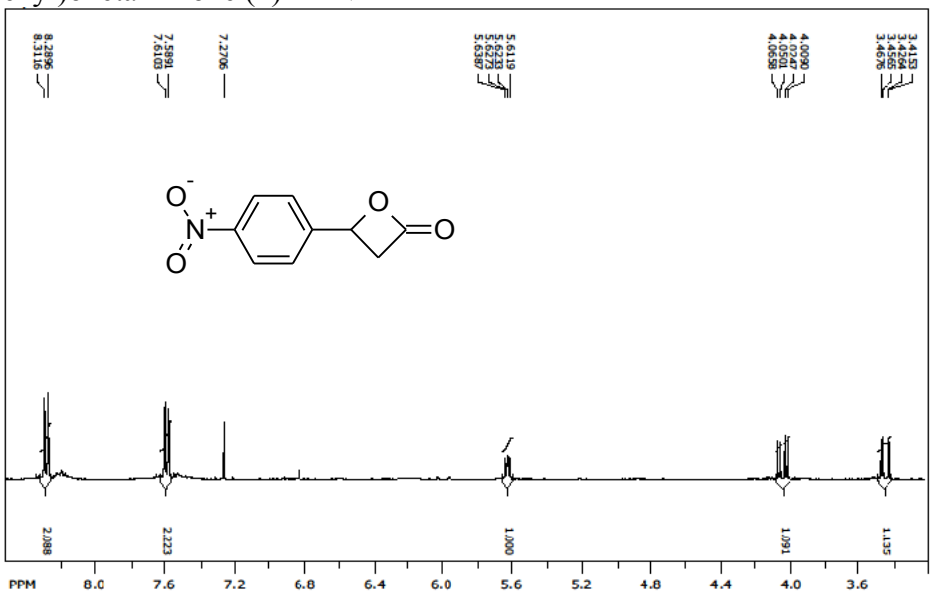
S5: 4-propyloxetan-2-one (**3**) ^{1}H NMR



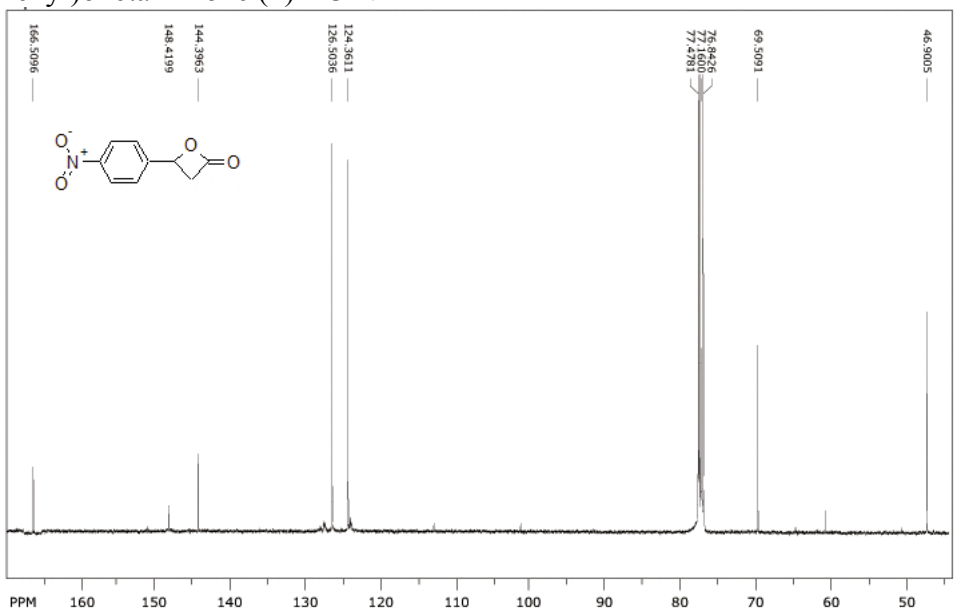
S6: 4-propyloxetan-2-one (**3**) ^{13}C NMR



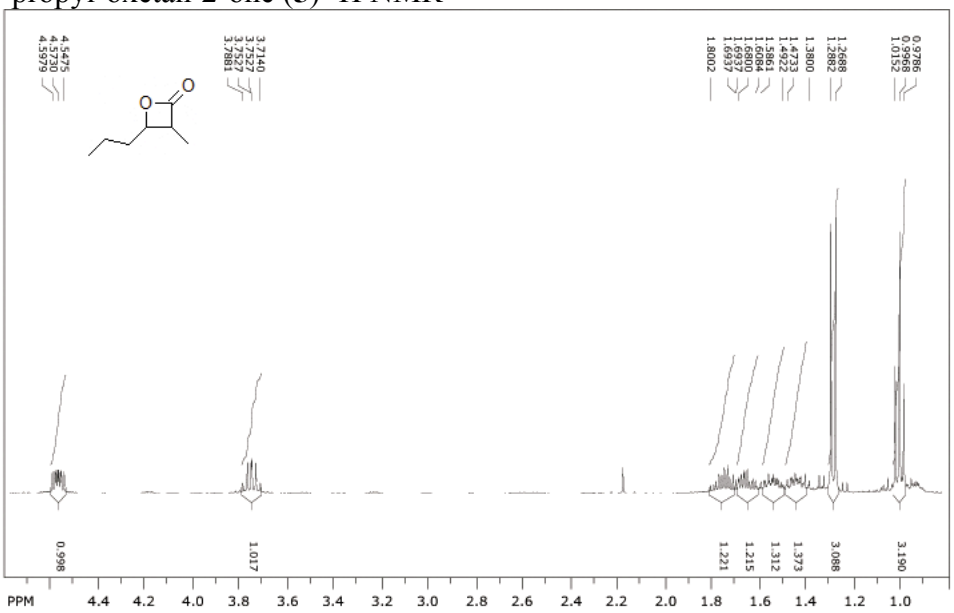
S7: 4-(4-nitrophenyl)oxetan-2-one (**4**) ^{1}H NMR



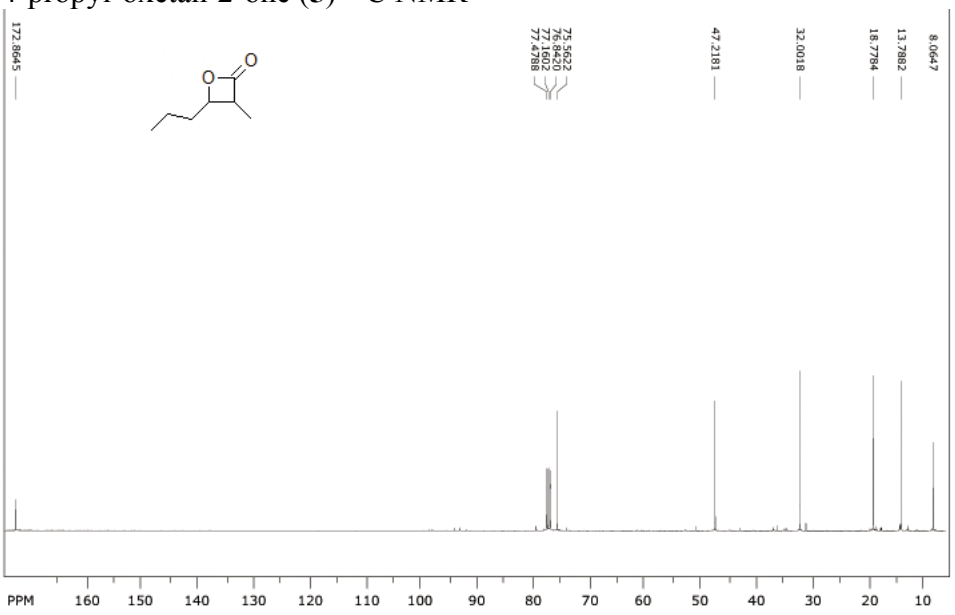
S8: 4-(4-nitrophenyl)oxetan-2-one (**4**) ^{13}C NMR



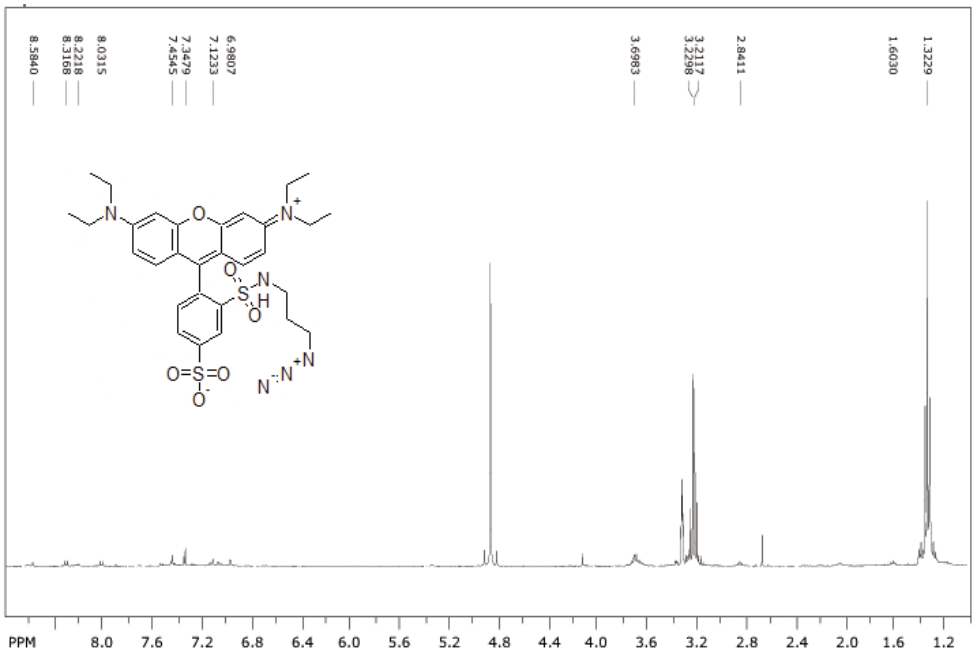
S9: 3-methyl-4-propyl-oxetan-2-one (**5**) ^{1}H NMR



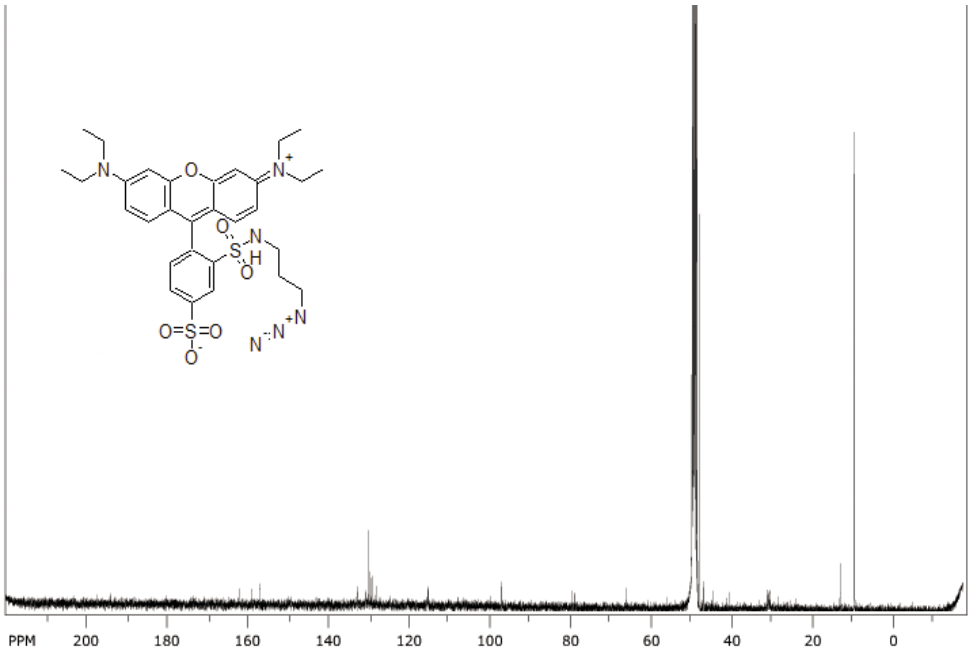
S10: 3-methyl-4-propyl-oxetan-2-one (**5**) ^{13}C NMR



S11: Sulforhodamine B azide ^{1}H NMR



S12: Sulforhodamine B azide ^{13}C NMR



Representative data for determination of lactone loading.

Tryptic digests and the exact masses of the resultant peptides were calculated using the Expasy Peptide Mass tool < http://ca.expasy.org/tools/peptide-mass.html>. The masses of adducts were calculated by adding the exact mass of the peptide to the exact mass of the lactone probe.

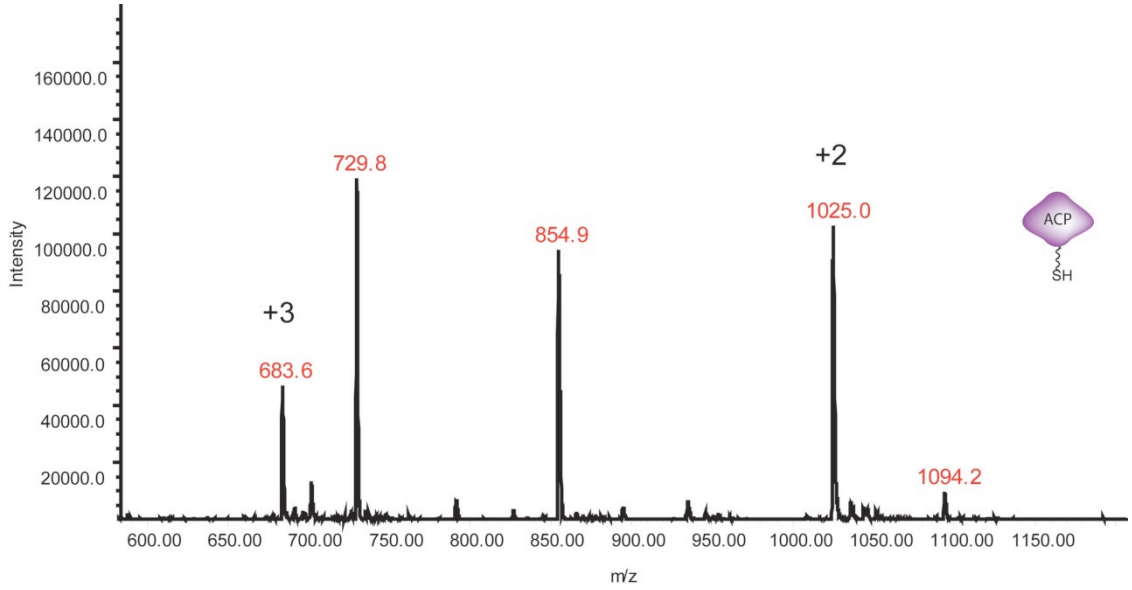
To determine the extent of lactone loading, the calculated mass of the modified and unmodified peptides were extracted from the TIC, then the mass of the modified and unmodified peptides were were compared using the formula:

% loading = $(I_m / (I_m + I_u)) \times 100$

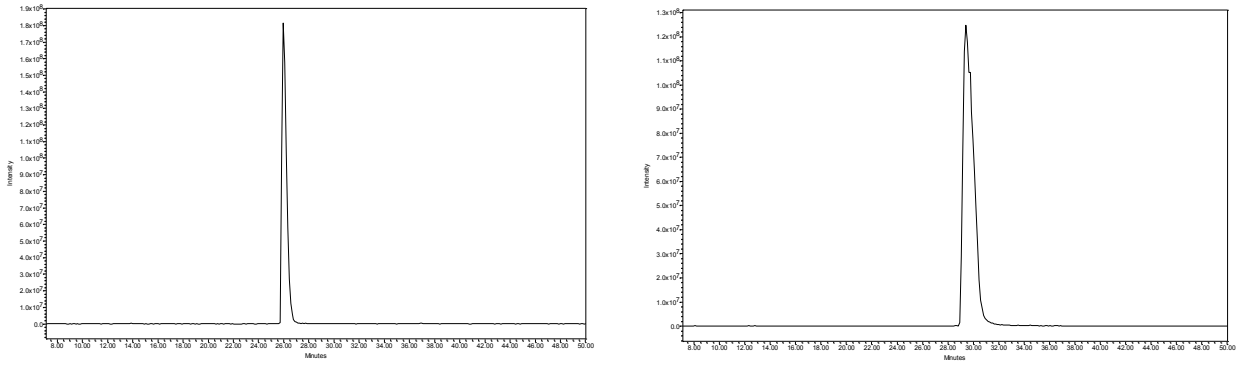
I_m = intensity of the modified peptide

I_u = intensity of unmodified peptide

S13: LC-MS data for trypsinized DEBS ACP3 in the absence of β-lactone. The spectrum of the Ppant-containing peptide fragment is shown. m/z values of +2, and +3 are labeled.

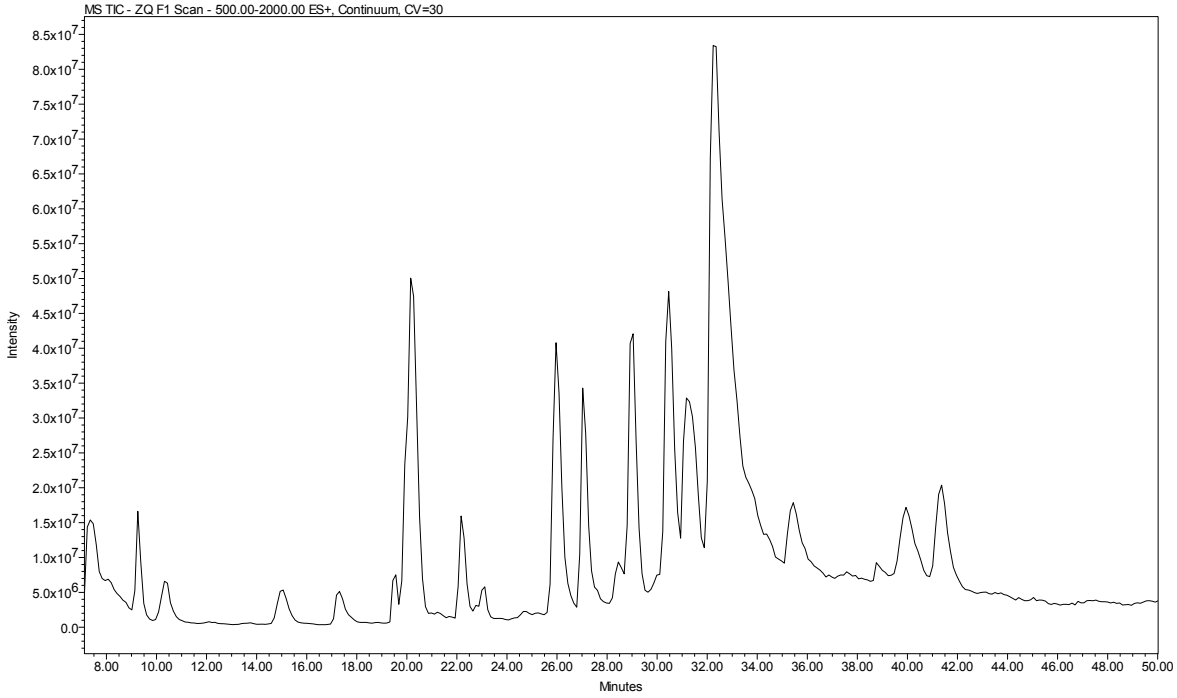


S14: Apo-ACP2 (Left) and ACP3 (Right) with 50 equivalents of compound **1**. Chromatograms are extracted at m/z = 589.3 (ACP2) and 854.4 (ACP3) each corresponding to unlabeled peptide. The presence and magnitude of these peaks indicates little to no reaction with compound **1**.

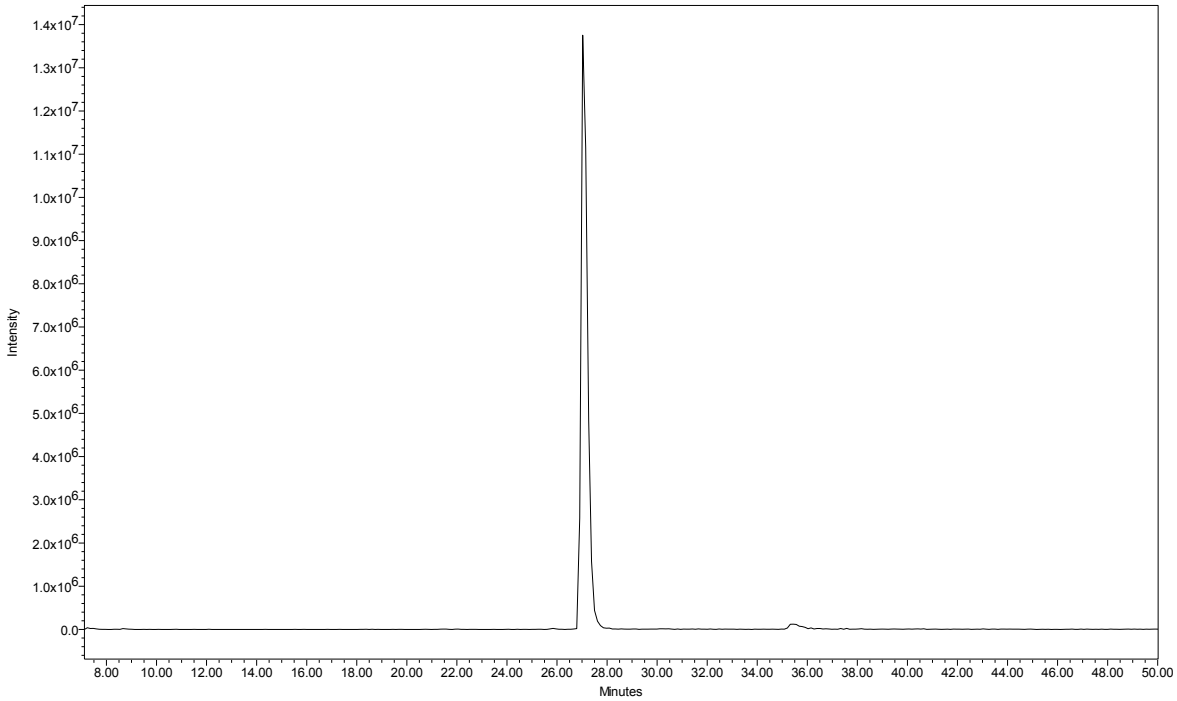


Figures S15-S19 show representative LC-MS data from which % Loading values are calculated

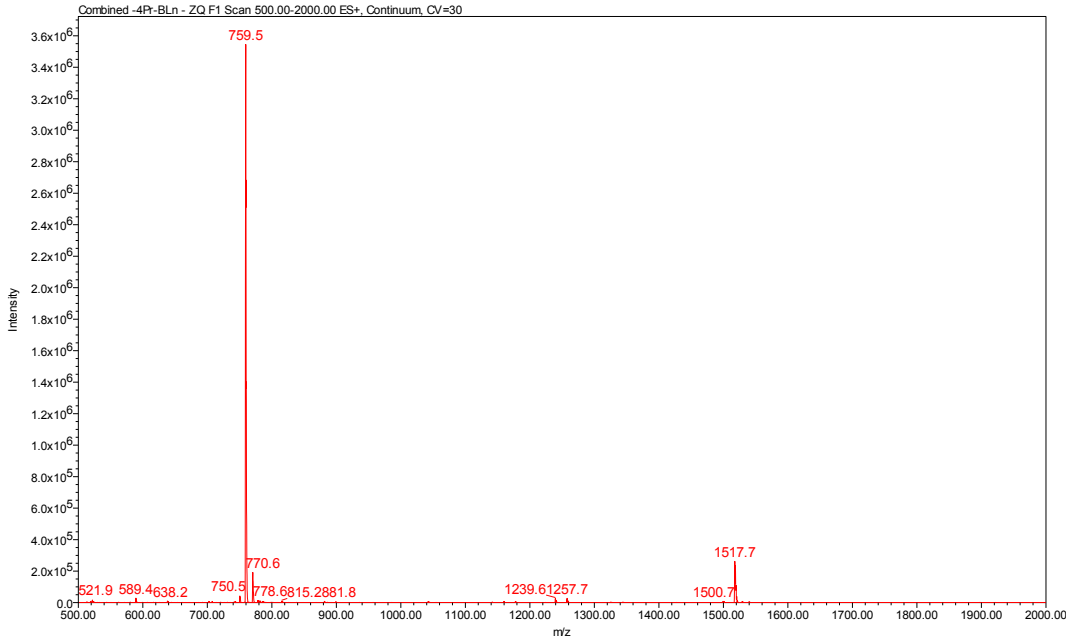
S15: *Holo* ACP2 + 10x 4-propyl-β-lactone TIC



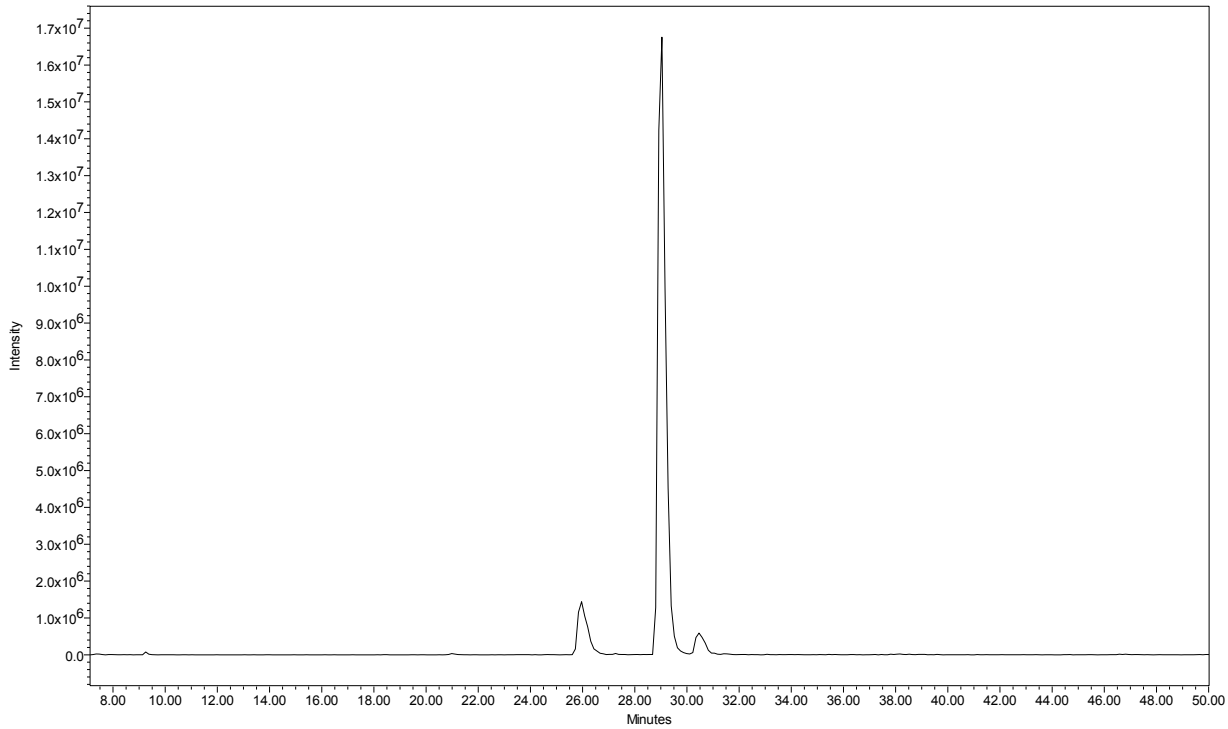
S16: *Holo* ACP2 + 10x 4-propyl-β-lactone TIC extracted at m/z = 759.34 $[M+2H]^{2+}$ (unmodified holo peptide)



S17: *Holo* ACP2 + 10x 4-propyl-β-lactone mass spectrum 26.6 – 27.1 min



S18: *Holo* ACP2 + 10x 4-propyl-β-lactone TIC extracted at m/z = 816.6 $[M+2H]^{2+}$ (holo peptide adduct)



S19: *Holo* ACP2 + 10x 4-propyl-β-lactone mass spectrum 29.2 – 29.5 min

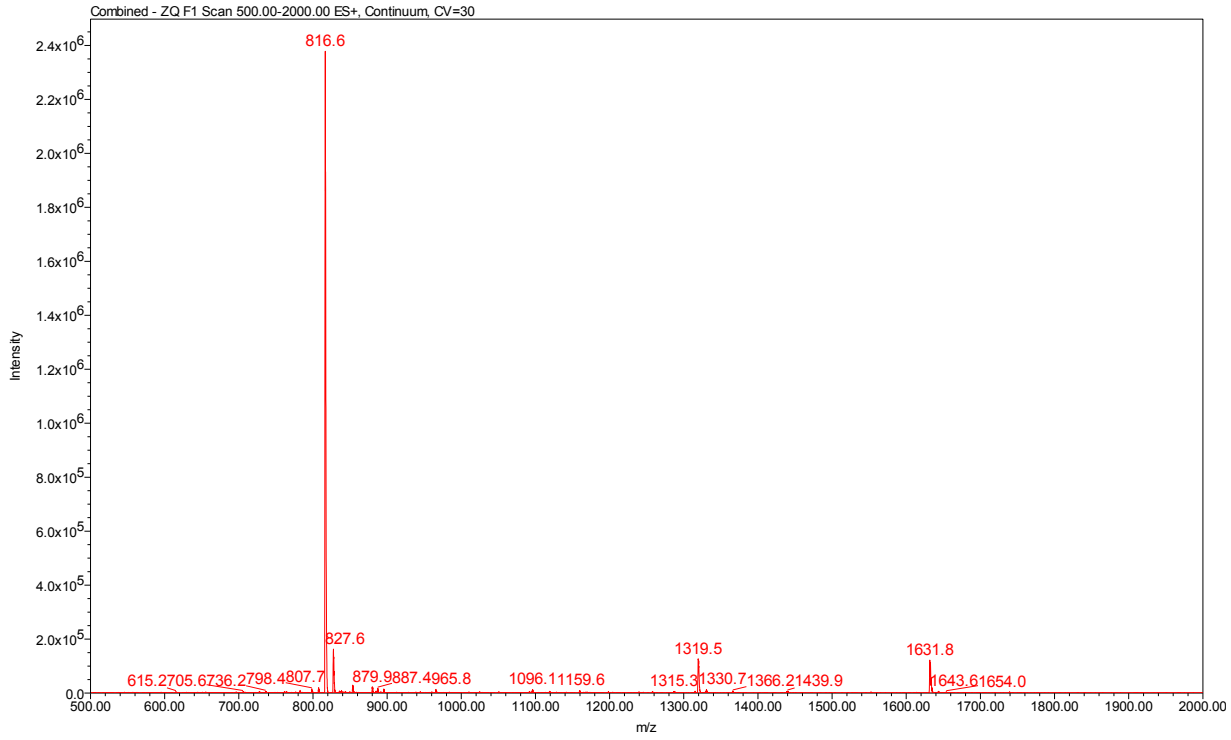


Table S1: Intensity of unmodified and modified peptides

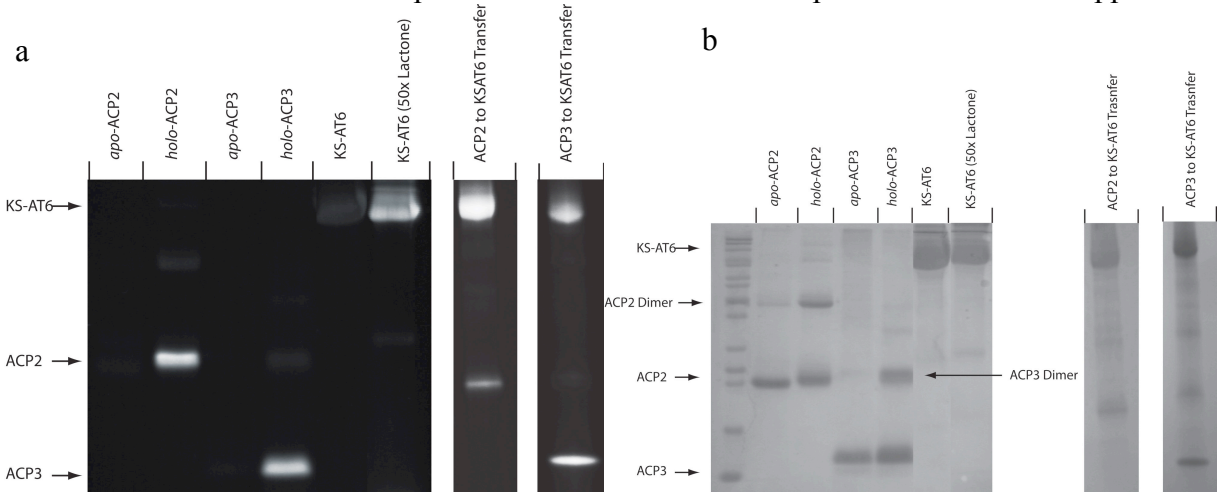
	m/z	Intensity
unmodified	759.5	1744512
adduct	816.6	1984160

(1984160 / (1744512 + 1984160)) x 100 = 53%

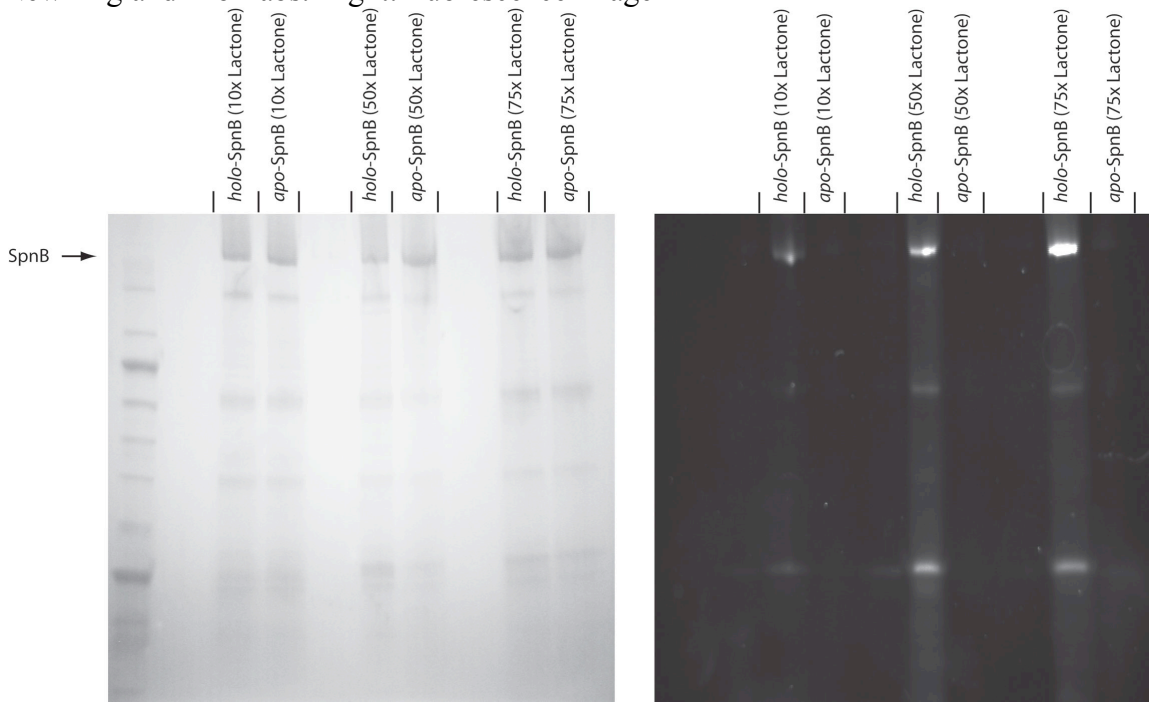
Table S2: Saturation Data. Average percent loading and standard deviations of saturation curve

Lactone Equivalents	ACP2 (%)	ACP3 (%)	KSAT6 (%)
5	16 ± 4	16 ± 4	6 ± 2
10	33 ± 0.02	55 ± 6	7 ± 3
20	77 ± 2	88 ± 2	16 ± 1
50	93 ± 4	97 ± 1	64 ± 4
75	98 ± 0.5	96 ± 0.1	67 ± 2

S20: Same gel as **Fig. 3** in text. Left: fluorescence image. Right stained with Fisher GelCode Blue. Far left lane contains Fisher EZ-Run *Rec* protein ladder. An unrelated experiment has been cropped out.



S21: Same gel as **Fig. 4** in text. Left: stained with Fisher GelCode Blue. Far left lane = 10 – 250 kDa ladder, New England Bio Labs. Right: fluorescence image



S22: LC-MS data for **SRB-oA**; a = 254 nm, b = 280 nm, c = TIC, d = TIC extracted at m/z = 641.1 $[m+H]^+$