

Chemical probes of surface layer biogenesis in *Clostridium difficile*

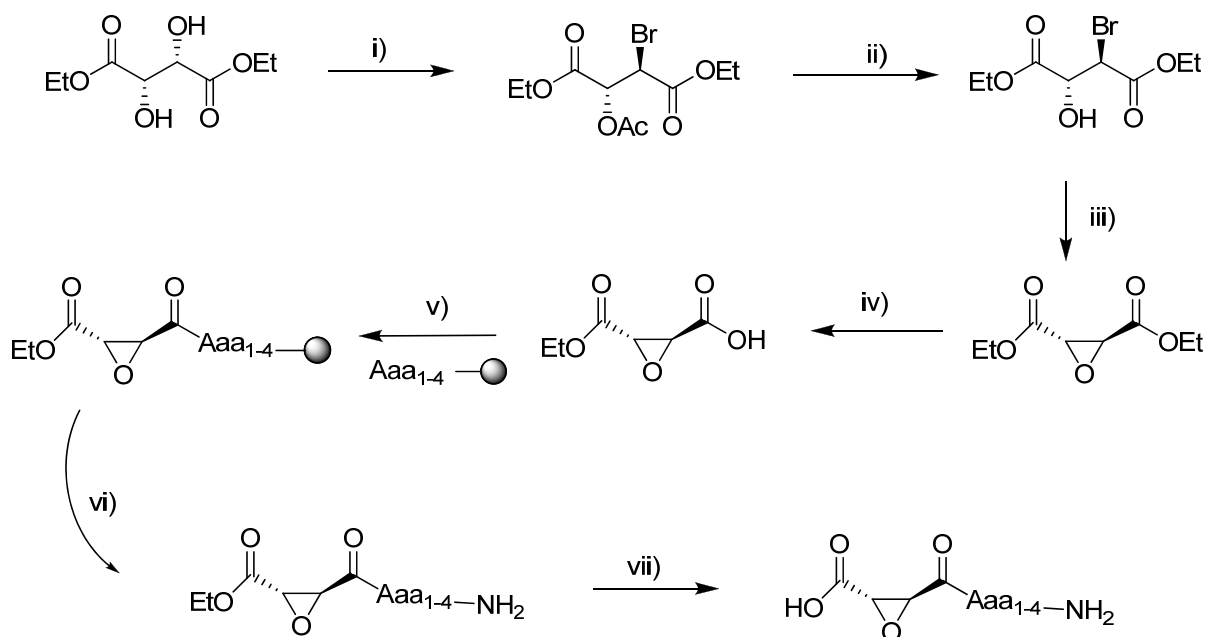
T. H. Tam Dang,^{1,†} Lucia de la Riva,^{1,2,†} Robert P. Fagan,² Elisabeth M. Storck,¹ William P. Heal,¹
Claire Janoir,³ Neil F. Fairweather,² and Edward W. Tate^{1*}

¹Department of Chemistry, Imperial College London, Exhibition Rd., London SW7 2AZ, UK; ² Centre for Molecular Microbiology and Infection, Department of Life Sciences, Imperial College London, Exhibition Rd., London SW7 2AZ; ³ EA 4043, Université Paris-Sud 11, Faculté de Pharmacie, 92296 Châtenay-Malabry, France. [†]These authors contributed equally to this work.

*Author for correspondence: e.tate@imperial.ac.uk

SUPPORTING INFORMATION

Contents	Page
Scheme S1: synthesis of inhibitors and ABPs	S1
Table S1: SlpA cleavage site alignments	S1
Figure S1: activity and labeling across clinical strains	S2
Figure S2: concentration series for 12 and 4	S2
Figure S3: effects of inhibition on cell lysis	S3
Figure S4: structure of clickable fluorescent reporter	S3
Figure S5: fluorescent labeling in ABP-treated cells	S4
Figure S6: Sequence of <i>SlpA</i> and <i>Cwp84</i> in <i>C. difficile</i> strain 630	S5
Figure S7: Anti-His-HRP blot against His-tagged Cwp84	S6
Table S2A: proteomic data for proteins identified in this study	S6
Table S2B: peptides matched for Cwp84	S7
Supplementary Methods	S8
Table S3: plasmids generated and used in this work	S9
Characterization data	S12
Table S4: data for synthetic inhibitors and ABPs used in this study	S12
Supplementary References	S19



Scheme S 1 : Synthesis of inhibitors followed the general scheme outlined previously by Bogyo *et al.* (1). Reagents and conditions: i) HBr-HOAc, 0°C to RT, 16 h; ii) HCl-EtOH, Δ, 5 h; iii) DBU, Et₂O, 2 h; iv) KOH, EtOH, 0 °C, 3 h; v) HCTU, DiPEA, 1 h; vi) TFA, TIS, H₂O, 2 h; vii) NaOH, 30 min. Yields vary from 1 to 87% Aaa indicates the peptide, synthesized by SPPS.

<i>C. difficile</i> Strain	LMW/HMW Cleavage site
630	...KRLET KS ▼ ANDTIASQ...
R7404	...KRLET KS ▼ ADIIADAD...
R8366	...KRVTT KS ▼ AAKASIAD...
CDKK959	...KRLTT KG ▼ ATGTLADE...
SE528	...KRLTT KG ▼ ATDIENTP...
CDKK167	...KRLQGFST YR ▼ ATNYN...
Y	...KRLQGFSS YG ▼ KFSTD...
R13541	...KRLNTAST YA ▼ SSNYK...

Table S1: Alignment of SlpA sequences at the LMW-HMW cleavage site across clinical strains of *C. difficile* tested in the present study (2, 3); 630 is the reference strain on which the majority of the work was performed, and for which the whole genome has been sequenced (4). The arrow indicates the site of cleavage of SlpA which is based on determination of the N-terminal residue of the HMW SLP by N-terminal sequencing (green) or by homology (black).

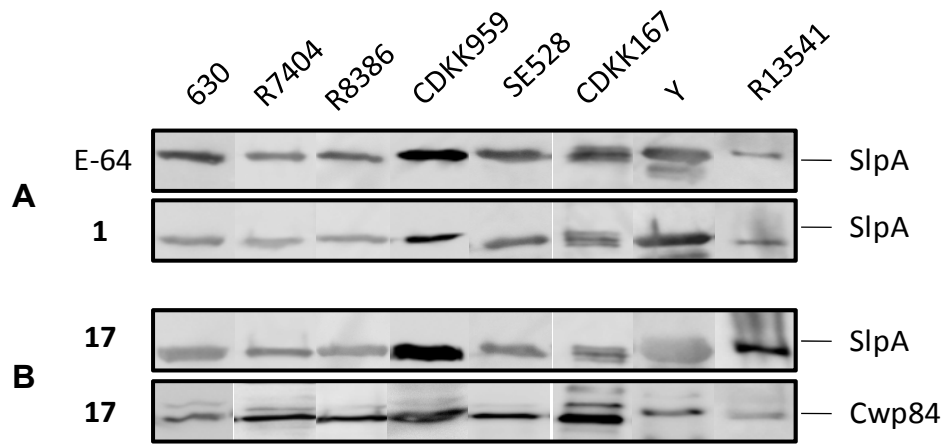


Figure S1: A: *in vivo* activity of E-64 and **1** across a panel of clinical *C. difficile* strains, visualized by Western blot with anti-LMW SLP antibody; B: *in vivo* activity and labeling of **17**.

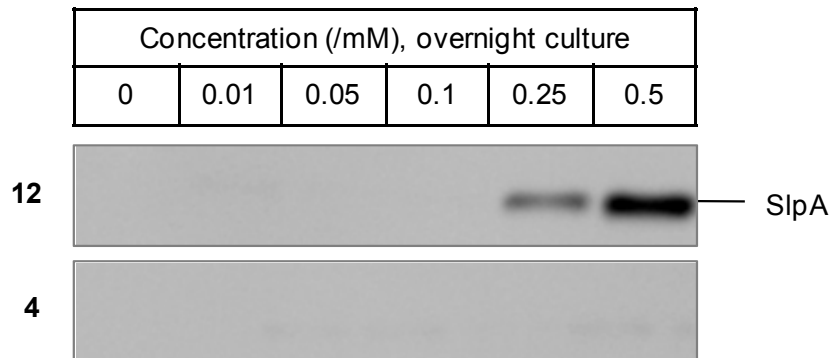


Figure S2: concentration series for **12** and **4**, visualized by Western blot using LMW SLP antibody.

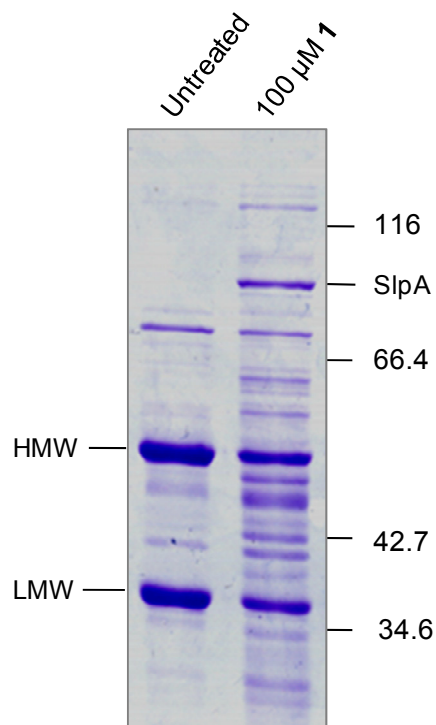


Figure S3: Coomassie blue-stained gels of protein released from uninhibited and inhibited cells after boiling in 0.7% SDS at 100°C for 5 min. Inhibited cells release a large quantity of intracellular protein, whilst untreated cells only release proteins from the S-layer.

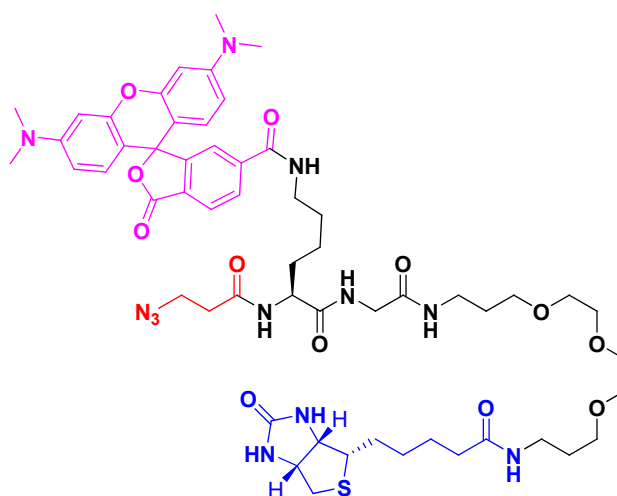


Figure S4: Structure of fluorescent reporter used for CuAAC ligation to ABPs.

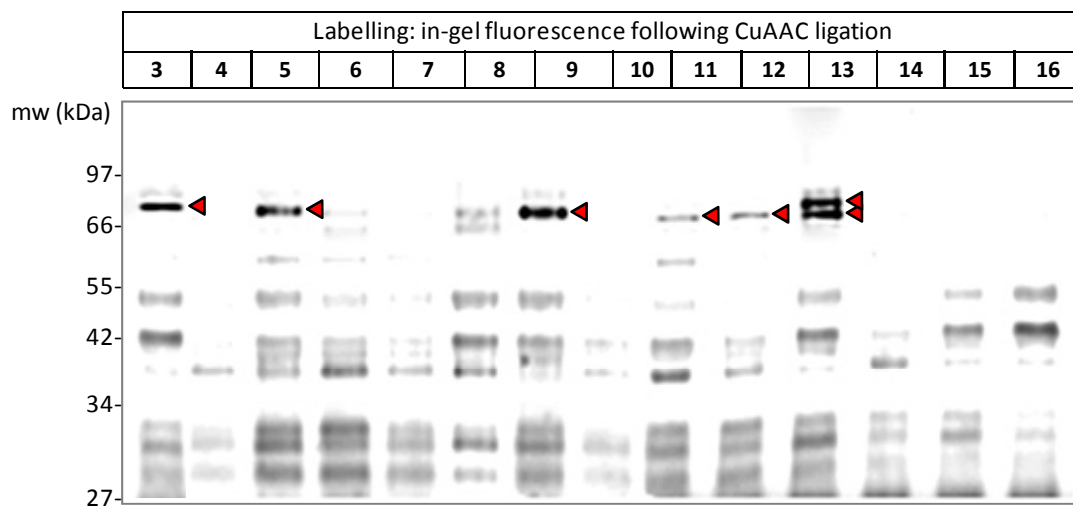
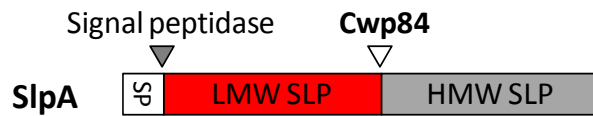


Figure S5: in-gel fluorescence image of the whole gel from Fig. 2C (main text). Activity-dependent labeling of bands shown in detail in Fig. 2C are marked with an arrow; bands below 66kDa are common across most experiments, and appear to be due to background CuAAC ligation.

A

MNKKNIAMSGLTVLASAAPVFAATTGTQGYTVVKNDWKKAVKQLQDGLKDNSIGKITVVSFNDGVVG
 EVAPKSANKKADRDAAAEKLYNLVNTQLDKLGDGDYVDFSVVDYNLENKIITNQADAEAIVTKLNSLNE
 KTLIDIATKDTFCMVSKTQDSEGKNVAATKALKVKDVATFGLKSGGSEDTGYVEMKAGAVEDKYGKV
 GDSTAGIAINLPSTGLEAYAGKGTIDFNKTLKVDVTGGSTPSAVAVSGFVTKDDTDLAKSGTINVRVI
 NAKESIDIDASSYTSANLAKRYVFDPEISEAYKAIVALQNDGIESNLVQLVNGKYQVIFYPEGKR
 LETKSANDTIASQDTPAKVVIKANKLKLKDYVDDLKTYNNTYSNVVTVAGEDRIETAIELSSKYNS
 DDKNAITDKAVNDIVLVGSTSIVDGLVASPLASEKTAPLLLTSKDKLDSSVKSEIKRVMNLKSDTGIN
 TSKKVYLAGGVNSISKDVENELKNMGLKVTRLSGEDRYETSLAIADEIGLDNDKAFVVGTTGLADAMS
 IAPVASQLKGDATPIVVVDGKAKEISDDAKSFLGTSVDVIIGGKNSVSKEIEESIDSATGKTPDRIS
 GDDRQATNAEVLKEDDYFTDGEVVNYFVAKDGGSTKEDQLVDALAAAPIAGRFKESPAPIILATDTLSS
 DQNVAVSKAVPKDGGTNLVQVGKGIASSVINKMKDLLDM

B

Cwp84:

MRKYKSKKLSKLLALLTVCFLIVSTIPVSAENHKTLDGVETAEYSESYLQYLEDVKNGDTAKYNGVIP
 FPHEMEGTTLRNKGRSSLPSAYKSSVAYNPMDLGLTTPAKNQGSLNTCWSFSGMSTLEAYLKLGKYGT
 YDLSEEHLRWATGGKYGWNLDDMSGSSNVTAIGYLTAWAGPKLEKDI PYNLKSEAQGATKPSNMDTA
 PTQFNVTDVVRLNKDKETVKNAIMQYGSVTSGYAHYSTYFNKDETAYNCTNKRAPLNHAVAVGWDDN
 YSKDNFASDVKPESENGAWLVKSSWGEFNSMKGFFWISYEDKTLTDTDNAMKSVSKPDSKDKMYQLE
 YAGLSKIMSNTAANVDFSRDSEKLD SVMFETDSVGAKEYEVYYPVVNGVVPQNNMSTKLASGTVSY
 SGYINVPTNSYSLPKGKGAIVVVIDNTANPNREKSTLAYETNIDAYLYEAKANLGESYILQNNKFED
 INTYSEFSPCNFVIKAITKTSSGQATSGESLTGADRYETAVKVSQKGTSSQNAVLVNGDAIVDALTA
 TPFTAIDSPILLTGKDNLDSKTKAELQRLGTRKVYLIGGENSLSKNVQTQLSNMGISVERISGSDRY
 KTSISLAQKLNLSIKSVSQVAVANGVNLADAI SVGAAAADNNMPIILTNEKSELQGADEF LNSSKITK
 SYIIGGTATLSSNLESKLSNPTRLAGSNRNETNAKIIDKFYPSDDLKYAFVVKDGSKSQGDLDGLAV
 GALGAKTDSPPVVLVGNKLDSESQNVLKSKKIETPIRVGGNGNESAFNELNLTLLGK

Figure S6: **A:** Sequence of *SlpA* in *C. difficile* strain 630 showing the signal peptide, LMW SLP and HMW SLP. The ‘HMW interaction domain’ (HID) is underlined; this region is sufficient to mediate binding to the LMW SLP following cleavage (5). The Ser321-Ala322 bond is cleaved *in vivo* to generate the LMW and HMW SLPs. **B:** Sequence of *Cwp84* in *C. difficile* 630. The active site cysteine C116 (identified by homology and confirmed by mutagenesis in this study) is highlighted in red, and the peptides identified by MALDI-ToF are highlighted in green.

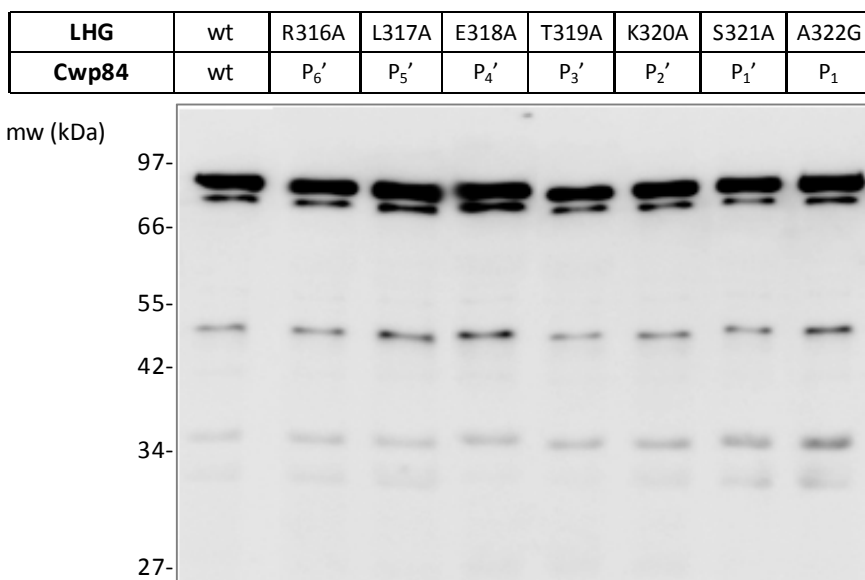


Figure S7: anti-His-HRP Western blot against His₆-tagged Cwp84 showing equal protease expression across co-expression experiments shown in Fig. 3 (main text).

Band No.	Protein Name	Approx. MW	Protein Score	Peptides matched	Coverage
1	SlpA	74kDa	205	12	25%
5	CD2797	222kDa	110	18	8.6%
6	Cwp84	87kDa	657	35	45%
7	Cwp66	66kDa	60	11	20.7%
8	CD2767	78kDa	110	14	26.1%

Table S2A: Band IDs for proteins identified in this study. Bands 2 (CD2791), 3 (HMW SLP) and 4 (LMW SLP) were identified previously (6), and by Western blot in the present study.

Calc. Mass	Obsrv. Mass	± da	± ppm	Start Seq.	End Seq.	Sequence
845.4839	845.4083	-0.0756	-89	567	573	TKAELQR
856.525	856.4453	-0.0797	-93	778	784	KIETPIR
862.4669	862.3854	-0.0815	-94	183	189	DIPYNLK
1172.504	1172.4114	-0.0926	-79	294	303	SSWGEFNSMK
1188.499	1188.4087	-0.0903	-76	294	303	SSWGEFNSMK +Oxidation (M)[9]
1226.6165	1226.5233	-0.0932	-76	352	362	VTAANVFDFSR
1226.6165	1226.5233	-0.0932	-76	352	362	VTAANVFDFSR
1274.6448	1274.5481	-0.0967	-76	704	715	LAGSNRNETNAK
1279.6892	1279.5887	-0.1005	-79	579	590	VYLIGGENSLSK
1291.5994	1291.501	-0.0984	-76	304	313	GFFWISYEDK
1302.6399	1302.5392	-0.1007	-77	336	346	MYQLEYAGLSK
1318.6348	1318.5366	-0.0982	-74	336	346	MYQLEYAGLSK +Oxidation (M)[1]
1407.7842	1407.675	-0.1092	-78	578	590	KVYLIGGENSLSK
1425.7625	1425.6594	-0.1031	-72	716	727	IIDKFYPSSDLK
1425.7625	1425.6594	-0.1031	-72	716	727	IIDKFYPSSDLK
1463.7489	1463.6265	-0.1224	-84	461	473	ANLGESYLQNNK
1524.7177	1524.6132	-0.1045	-69	664	677	SELQGADEFLNSSK
1539.7074	1539.6068	-0.1006	-65	133	145	GYGTYDLSEEHLR
1539.7074	1539.6068	-0.1006	-65	133	145	GYGTYDLSEEHLR
1552.8441	1552.7421	-0.102	-66	426	440	GAIVVVIDNTANPNR
1552.8441	1552.7421	-0.102	-66	426	440	GAIVVVIDNTANPNR
1584.8591	1584.7527	-0.1064	-67	738	754	SQGLDIDGLAVGALGAK
1675.8431	1675.7415	-0.1016	-61	591	605	NVQTQLSNM GISVER
1675.8431	1675.7415	-0.1016	-61	591	605	NVQTQLSNM GISVER
1691.8381	1691.7332	-0.1049	-62	591	605	NVQTQLSNM GISVER +Oxidation (M)[9]
1737.9606	1737.8569	-0.1037	-60	424	440	GKGAIVVVIDNTANPNR
1737.9606	1737.8569	-0.1037	-60	424	440	GKGAIVVVIDNTANPNR
1764.8837	1764.777	-0.1067	-60	92	108	SSVAYNPMDLGLTTPAK
1780.8785	1780.7812	-0.0973	-55	92	108	SSVAYNPMDLGLTTPAK +Oxidation (M)[8]
1780.8865	1780.7812	-0.1053	-59	131	145	LKGYGTYDLSEEHLR
1809.9817	1809.8806	-0.1011	-56	426	442	GAIVVVIDNTANPNREK
1828.9651	1828.859	-0.1061	-58	755	771	TDSPVVLVGNKLDSEKQ
1866.9443	1866.824	-0.1203	-64	664	680	SELQGADEFLNSSKITK
1933.9613	1933.854	-0.1073	-55	785	803	VGGNGNESAFNELNTLLGK
1957.9059	1957.8025	-0.1034	-53	363	380	DSEKLD SVMFETDSVGAK
1969.9767	1969.8677	-0.109	-55	258	275	APLNHAVAIVGWDDNYSK
1972.0345	1971.8685	-0.166	-84	734	754	DGSKSQGLDIDGLAVGALGAK
1976.9535	1976.8582	-0.0953	-48	63	79	YNGVIPFPHEM EGTTLR +Oxidation (M)[11]
1995.0981	1994.9933	-0.1048	-53	424	442	GKGAIVVVIDNTANPNREK
2126.0776	2125.9807	-0.0969	-46	257	275	RAPLNHAVAIVGWDDNYSK
2128.0122	2127.9282	-0.084	-39	443	460	STLAYETNIDAYLYEAK
2210.9951	2210.9956	0.0005	0	474	491	FEDINTYSEFSPCNFVIK +Carboxy methyl (C)[13]
2316.095	2316.0005	-0.0945	-41	496	518	TSSGQATSGESLTGADRYETAVK
2316.095	2316.0005	-0.0945	-41	496	518	TSSGQATSGESLTGADRYETAVK
2418.2188	2418.1165	-0.1023	-42	401	423	LASGTVSYSGYINVPTNSYSLPK
2764.3206	2764.2556	-0.065	-24	190	215	SEAQGATKPSNMDTAPTQFNVDVVR
2764.3206	2764.2556	-0.065	-24	190	215	SEAQGATKPSNMDTAPTQFNVDVVR

Table S2B: Full list of peptides matched in proteomic ID of Cwp84.

Supplementary Methods

SDS-PAGE and western blotting: Proteins were resolved on SDS-PAGE gels as described previously (5). For immunoblotting, proteins were transferred to Immobilon-P PVDF membrane and incubated with primary rabbit anti-LMW SLP antibody (1 in 200,000) and detected using HRP-labeled secondary antibody (1 in 2000). Signals were detected using SuperSignal[®] West Pico Chemiluminescent Substrate (Thermo Scientific) followed by image capture using a LAS-3000 Image Reader. For NeutrAvidin–HRP immunoblotting, after transfer the membranes were blocked with 5% BSA, 1.0% Tween 20 in PBS and were probed using the appropriate antibody–HRP conjugate (NeutrAvidin[™]) (1:5000). The membranes were then washed with PBS, 0.1% Tween 20 (PBS-T) for 3 × 5 min before detection as described above.

Bioorthogonal labeling of ABP alkyne-tagged proteins: It was found that prior dialysis of the samples into a low salt phosphate buffer was essential for efficient labelling; CuAAC ligation does not appear to tolerate the concentrated glycine buffer used to strip the S-layer from cells. Following dialysis through a low molecular weight cut-off spin concentrator, CuAAC ligation was performed essentially as described previously (7) using the capture reagent shown in Fig. S4 followed by direct visualization of labelled proteins by in-gel fluorescence.

Affinity purification of ABP tagged proteins: *C. difficile* cells were labeled overnight with ABPs, and the cells were harvested and glycine extracts prepared. Streptavidin beads were added and incubated for 2 hours at 4 °C. The samples were centrifuged and the pellet containing the beads and bound proteins were washed with PBS-T, SDS-PAGE loading buffer was added and the biotin-avidin bond was broken by heating. The supernatant containing unbound proteins was also retained for gel analysis.

Plasmid construction and mutagenesis: Residues 1-362 of SlpA were cloned as a fusion with glutathione-S-transferase in MCS2 of pACYCDuet-1 (Novagen) to form pLRP003. Residues 30-803 of Cwp84 were cloned in the MCS1 of the same plasmid to form pLRP007. Plasmids and derivative of pLRP007 made by inverse PCR are listed in Table S3. Expression of LHG derivatives was performed using Overnight Express medium (Novagen).

Plasmid	Description
pLRP003	pACYCDuet-1 expressing LHG
pLRP007	pACYCDuet-1 expressing Cwp84 and LHG
pLRP012	pLRP007 Cys116Ala in Cwp84
pLRP023	pLRP007 Arg316Ala in LHG
pLRP022	pLRP007 Leu317Ala in LHG
pLRP021	pLRP007 Glu318Ala in LHG
pLRP026	pLRP007 Thr319Ala in LHG
pLRP020	pLRP007 Lys320Ala in LHG
pLRP019	pLRP007 Ser321Ala in LHG
pLRP018	pLRP007 Ala322Gly in LHG

Table S3: Plasmids used in this work.

Protein identification: The bands of interest were excised, destained in 400 μ L of 50% (vol/vol) acetonitrile in 0.1 M ammonium bicarbonate (pH 8.4), and dried by using a SpeedVac drying apparatus. The gel slices were rehydrated in 20 μ L of trypsin working solution (Promega sequencing grade modified trypsin, prepared according to the manufacturer's instructions) and incubated at 37°C overnight. The supernatant was removed, and digestion was terminated by the addition of 50 μ L of 0.1% (vol/vol) trifluoroacetic acid (10 min, 37°C). The supernatant was removed, and the peptides were further extracted with 200 μ L of 60% (vol/vol) acetonitrile-0.1% (vol/vol) trifluoroacetic acid (15 min, 37°C). The supernatant was again removed and pooled with the previous supernatant. Both extraction steps were then repeated, and the supernatants were pooled. The volume of the combined supernatants was subsequently concentrated by using a SpeedVac drying apparatus. These samples were then premixed with 1 μ L of matrix (α -cyano-cinnamic acid) before loading onto a metal MALDI target plate. Matrix assisted laser desorption/ionisation-mass spectrometry (MALDI- MS) was performed using a PerSeptive Biosystems Voyager DE STR mass spectrometer (Foster City, CA, USA) in the linear mode with delayed extraction. Peptide m/z signals were identified and MS peak masses were compared with the peptide fingerprint database using Mascot (Matrix Science).

Chemistry materials and methods

All general laboratory chemicals obtained from chemical suppliers (Novabiochem UK or Aldrich Chemical Co.) were used without further purification.

Flash chromatography was performed on Screening Devices silica gel 60 (0.04 – 0.063 mm). TLC-analysis was conducted on DC-alufolien (Merck, Kieselgel60, F254) with detection by UV-absorption (254 nm).

¹H and ¹³C NMR spectra were recorded on a Brüker AV- 400 (400/100 MHz) spectrometer. Chemical shifts (δ) are given in ppm relative to the chloroform or D₂O residual solvent peak or tetramethylsilane as internal standard. Coupling constants are given in Hz.

Optical rotations were measured at 25°C on a Perkin-Elmer 241 polarimeter with sodium lamp, and are given in 10⁻¹ deg cm² g⁻¹; concentration (*c*) is in g per 100 mL.

Infrared spectra were obtained on PerkinElmer Spectrum 100 FT-IR spectrophotometer, from a thin film deposited onto a sodium chloride plate from dichloromethane. The IR spectra were then analyzed with Spectrum Express, Version 1.01.00 (PerkinElmer).

Mass spectra and accurate mass data were obtained on an AUTOSPEC P673 spectrometer by J. Barton at the departmental Mass Spectrometry Service (Imperial College London) by electrospray ionization, or chemical ionization techniques.

Purification of crude peptides was performed on a Gilson semi-preparative RP-HPLC system (Anachem Ltd., Luton, UK) equipped with 306 pumps and a Gilson 155 UV/Vis detector. Analytical RP-HPLC was performed on a Gilson analytical HPLC system (Anachem Ltd., Luton, UK) equipped with a Gilson 151 UV/Vis detector and Gilson 234 auto injector. For both HPLC systems, the peptide bond absorption was detected at 223 nm. The following elution methods were used: **Method A**: 2% to 98% MeCN in H₂O over 40 min; **Method B**: 2% MeCN in H₂O for 10 min then 2% to 98% MeCN in H₂O over 40 min; all solvents used were supplemented with 0.1% TFA and were degassed with helium. All peptides and reaction mixtures were analyzed using Method A unless otherwise specified.

LC-MS analysis was performed on Waters HPLC system (Waters 2767 autosampler for samples injection and collection; Waters 515 HPLC pump to deliver the mobile phase to the source; XBridge C18 column (Waters, 4.6 mm D × 100 mm L for analytical and 19mmD × 100 mmL for preparative); Waters 3100 mass spectrometer with ESI and Waters 2998 Photodiode Array (detection at 200-600 nm)). The elution method was 5% to 95% MeCN in H₂O over 15 min. All solvents used were supplemented with 0.1% formic acid and were degassed with helium.

Automated solid phase synthesis was performed using an Advanced ChemTech Apex 396 multiple peptide synthesizer (Advanced ChemTech Europe, Cambridge, UK). Rink-amide resin and Biotin-PEG-NovaTag resin were purchased from Novabiochem UK. The resin substitution ratios were as follows: Rink-amide resin: 0.71 mmol/g; Biotin-PEG-NovaTag resin: 0.48 mmol/g. *N*- α -9-Fluorenylmethoxycarbonyl (*N*- α -Fmoc) protected amino acids were obtained from Novabiochem UK

with the following side chain protecting groups: Arg(Pbf), Lys(Boc), Tyr(tBu) , Thr(tBu), Glu(tBu), Ser(tBu). Peptide synthesis grade dimethylformamide (DMF) was purchased from National Diagnostics, UK. Benzotriazole-1-yloxy-tris-pyrrolidino-phosphonium hexafluorophosphate (PyBop), *N*-hydroxybenzotriazole (HOBt), diisopropylethylamine (DIPEA) and 2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate (HBTU) were purchased from Novabiochem UK and Sigma Aldrich. Dry tetrahydrofuran (THF) and DMF, chloroform, dimethylsulfoxide (DMSO), dichloromethane (DCM), methanol, piperidine, thioanisole, *t*butylmethylether (TBME), trifluoroacetic acid (TFA), ethanedithiol (EDT), iodoacetonitrile, 3-mercaptoethyl propionate, sodium thiophenolate, monobasic and dibasic sodium phosphate, trimethylsilylisopropane (TIPS) were reagent grade from Sigma Aldrich.

Solid Phase Peptide Synthesis (SPPS)

Peptide synthesis was carried out in peptide synthesis grade DMF. Resin (25 μ mol per well) was swelled in DMF for 60 min before coupling each cycle, *N*- α -amino Fmoc group was treated with 20% (v/v) piperidine in DMF for 15 min, with two further repetitions. A fivefold excess of amino acid over resin reactive groups (125 μ mol, 250 μ L of 0.5 M solution) was used for each coupling. *In situ* activation and coupling was carried out for 45 min with a mixture of HBTU/HOBt (125 μ mol, 250 μ L of 0.5 M solution) and DIPEA (125 μ mole, 250 μ L of 0.5 M solution). The peptide was washed with DMF (3 \times 1 mL) between each deprotection and coupling step. A typical cycle consisted of deprotection, DMF wash, coupling and a further DMF wash. The *N*- α -Fmoc protection at the final residue was removed at the end of the synthesis under the usual conditions. After synthesis was complete a pre-activated solution of (2*S*,3*S*)-3-(ethoxycarbonyl)oxirane-2-carboxylic acid (12.01 mg, 3 eq), HCTU (51.7 mg, 3 eq) and DiPEA (19.3 mg, 6 eq) in DMF (1 mL) was added to each well and agitated for one hour. The peptidyl resin was removed from the synthesizer, washed several times (3 \times 2 mL DMF, 3 \times 2 mL DCM, 3 \times 2 mL methanol, 3 \times 2 mL diethylether) and dried *in vacuo*.

Deprotection and Cleavage from resin

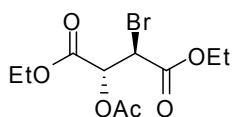
Cleavage and deprotection of peptides was achieved by adding 1.5 mL of deprotection mixture of TFA : H₂O : TIPS (97:2.5:1) to dry peptidyl resin (25 μ mol). The mixture was agitated on an orbital shaker for up to 3 hr and then filtered. The resin was washed twice with a small volume of TFA and the combined washings and filtrate were precipitated with 10 mL ice cold *tert*-butylmethylether (tBME). The mixture was centrifuged at 5000 rpm for 15 min at 0 °C, the supernatant discarded and the remaining peptide washed with a fresh aliquot of tBME. The process was repeated three times to ensure complete removal of all organic impurities. The crude peptide was dried in a desiccators over silica gel to yield an off-white solid.

Characterization data

Inhibitor	Calc. Mass	ESI-MS/[M+H ⁺]	Retention time R _t	Yield (%)
1	401.2143	401.2138	10.72 ^a	22.4
2	429.2456	429.2456	13.79 ^a	22.4
3	524.2833	524.2811	6.29	43.8
4	326.1352	326.1339	10.99	7.4
5	402.1665	402.1679	7.25	1
6	368.1822	368.1830	6.77	3.5
7	368.1822	368.1808	9.26	7.9
8	354.1660	354.9669	9.01	3.4
9	418.1599	418.1614	6.98	8.4
10	342.1301	342.1293	14.30 ^a	7.0
11	392.1568	392.1570	4.82	2.0
12	383.1931	383.1920	4.76	4.2
13	411.1992	411.2005	5.40	1
14	384.1407	384.1394	5.44	4.2
15	312.1196	312.1206	4.76	6.8
16	352.1509	352.1494	5.81	4.6
17	858.4773	858.4786	11.65 ^a	46
18	745.3959	745.3939	12.36 ^a	16

Table S4: data for synthetic inhibitors and ABPs used in this study. R_t of LC-MS using Acetonitrile - H₂O with 0.1% of formic acid; ^a indicates the R_t by RP-HPLC using Acetonitrile - H₂O with 0.1% of TFA, gradient of 2-98% over 20 min. Yields are following purification by HPLC, and take into account the entire synthetic sequence starting from resin loading.

(2R, 3R)-Diethyl-2-acetoxy-3-bromosuccinate

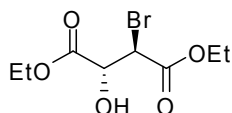


To neat (-) diethyl-D-tartrate (10 mL, 58.44 mmol) cooled in ice was added hydrobromic acid 33% in solution in acetic acid (5.7 M) (40 mL) drop wise over 30 minutes. The reaction mixture was allowed to warm to room temperature and stirred 16 h. The reaction mixture was poured into 50 g crushed ice and extracted four times with diethyl ether. The combined organic layers were washed three times with water and once with brine solution. After drying over magnesium sulphate all volatiles were removed under reduced pressure to afford the product (2R, 3R)-diethyl-2-acetoxy-3-bromosuccinate

(18.61 g, 59.88 mmol, 88% yield) as a colorless oil. The product contained acetic acid (ca. 50%) and was used for the next synthesis step without further purification.

δ_{H} /ppm (400 MHz; CDCl_3): 5.56 (1H, d, $J = 5.4$, CHOAc), 4.76 (1H, d, $J = 5.4$, CHBr), 4.28–4.16 (4H, m, $2 \times \text{CH}_2$), 2.07 (3H, s, COCH₃), 1.33–1.24 (6H, m, $2 \times \text{CH}_3$). δ_{C} /ppm (100 MHz; CDCl_3): 169.44, 166.25, 165.63 (CO), 72.66 (COAc), 62.81, 62.53 (CH₂), 43.86 (CBr), 20.70 (COCH₃), 13.94, 13.86 (CH₃).

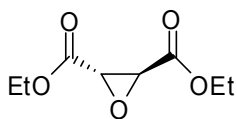
(2R,3R)-Diethyl 2-bromo-3-hydroxysuccinate



A solution of acetyl chloride (2.08 mL, 29.22 mmol) was added to a solution of (2R, 3R)-diethyl-2-acetoxy-3-bromosuccinate (18.61 g, 59.88 mmol) in ethanol (60 mL). The reaction mixture was heated to a gentle reflux for 6 hours, 1 hour longer than literature procedure, due to difficulty of visualizing on TLC. After 6 hours all volatiles were removed by rotary evaporation. The product (2R,3R)-diethyl-2-bromo-3-hydroxysuccinate was recovered as pale yellow oil (15.30 g, 56.86 mmol, 97% yield).

δ_{H} /ppm (400 MHz; CDCl_3): 7.48 (1H, s, OH), 4.68 (1H, d, $J = 4.5$, HOCH), 4.65 (1H, d, $J = 4.5$, BrCH), 4.36–4.16 (4H, m, $2 \times \text{CH}_2$), 1.29 (6H, t, $J = 7.2$, $2 \times \text{CH}_3$). δ_{C} /ppm (100 MHz; CDCl_3): 170.32 (COCO), 166.72 (COCBr), 72.52 (COH), 62.84, 62.53 (CH₂), 47.55 (CBr), 14.00, 13.88 (CH₃).

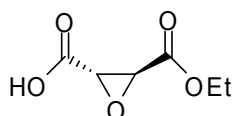
(2S,3S)-Diethyl -2,3- epoxysuccinate



To an ice cooled solution of (2R,3R)-diethyl 2-bromo-3-hydroxysuccinate (15.03 g, 56.86 mmol) in diethyl ether (80 mL) stirring, a solution of 1,8-diazabicyclo-undec-7-ene (DBU) (11 mL, 73.9 mmol) in diethyl ether (40 mL) drop-wise over 2 hours. Once the addition was completed the reaction mixture was stirred for a further 30 minutes. Water (10 mL) was added to the mixture. The separated organic layer was washed three times with 1 M potassium hydrogen sulphate and twice with water. After drying over magnesium sulphate all volatiles were removed under reduced pressure to afford a product (8.31 g, 44.15 mmol, 77% yield) as yellow oil.

δ_{H} /ppm (400 MHz; CDCl_3): 4.35 – 4.13 (4H, m, CH₂), 3.65 (2H, s, CHOCH), 1.31 (6H, t, $J = 7.2$, CH₃). ν_{max} (neat)/cm⁻¹ 2983, 2940, 2875 (CH), 1748 (CO), 1447, 1371, 1329, 1197, 1029, 901

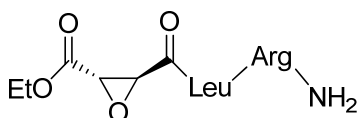
(2S,3S)-3-(Ethoxycarbonyl)oxirane-2-carboxylic acid



To an ice cooled solution of (2S,3S)-diethyl oxirane-2,3-dicarboxylate (1.00 g, 5.31 mmol) in ethanol (13 mL) was added drop wise a solution of potassium hydroxide (0.35 g, 6.26 mmol) in ethanol (7 mL) over 15 minutes. After this addition the reaction mixture was allowed to stir in ice for 3 hours. The cooling bath was then removed and the reaction mixture was stirred for further 2 hours at room temperature. Afterwards the ethanol was removed under reduced pressure leaving an off white solid. This salt was then taken up with 15 mL of water and washed three times with dichloride methane. The aqueous phase was acidified by addition of concentrated hydrochloric acid to pH 3 and extracted four times with ethyl acetate. The combined organic phase was dried over magnesium sulphate and concentrated under reduced pressure. After freeze drying the pure product was obtained as a pale yellow oil (0.67 g, 4.19 mmol, 79% yield).

δ_{H} /ppm (400 MHz; C): 4.27 (2H, q, $J = 7.3$, CH_2), 3.64 (1H, d, $J = 1.6$, CH), 3.61 (1H, d, $J = 1.7$, CH), 1.32 (3H, t, $J = 7.1$, CH_3). δ_{C} /ppm (100 MHz; CDCl_3): 168.87, 167.05 (CO), 62.12 (CH_2), 51.98, 51.90 (CH), 13.95 (CH_3). $[\alpha]_{\text{D}}^{20} +94.44$ ($c=1.008$ in diethylether). ν_{max} (neat)/ cm^{-1} : 2988 (CH), 1744 (CO), 1448, 1374, 1206, 1095, 1025, 899.

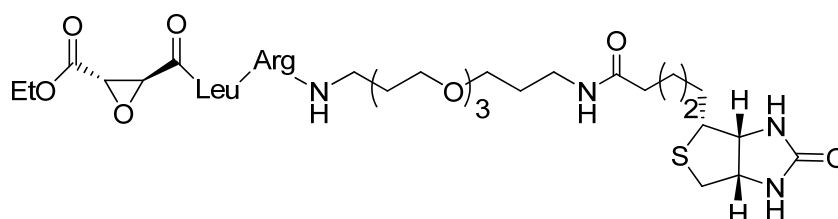
(2S,3S)-ethyl 3-((4R,7R)-10-guanidino-7-hydrazinyl-2-methyl-5,6-dioxodecan-4-ylcarbamoyl)oxirane-2-carboxylate (2)



Prepared by standard SPPS. Product was obtained as white solid.

δ_{H} /ppm (400 MHz; D_2O): 4.334.30 (1H, m, NHCHArgCO), 4.274.22 (1H, m, NHCHLeuCO), 4.21 (2H, q, $J = 7.15$, CH_3CH_2), 3.71 (1H, d, $J = 1.86$, CHOCHCOOEt), 3.64 (1H, d, $J = 1.86$, CHOCHCOOEt), 3.13 (2H, t, $J = 6.92$, $\text{CH}_2\text{NHCNHNH}_2$), 1.84-1.51 (7H, m, $\text{CH}_2\text{CH}(\text{CH}_3)_2$, CH_2), 1.21 (3H, t, $J = 7.17$, CH_3CH_2), 0.86, 0.81 (6H, d, $J = 5.80$ and 5.75 , $\text{CH}(\text{CH}_3)_2$). **ESI-MS:** $m/z = 429.2463$ ($[\text{M} + \text{H}]^+$), 469.3159 ($[\text{M} + \text{Na}]^+$). **R_t:** 10.88 (Method A)

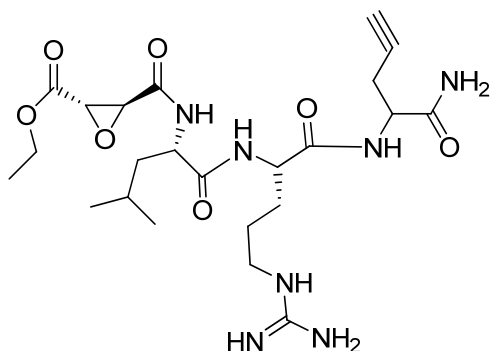
(2S,3S)-ethyl 3-((22R,25R)-22-(3-guanidinopropyl)-27-methyl-5,23,24-trioxo-1-((3aR,4R,6aS)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)-10,13,16-trioxa-6,20,21-triazaoctacosan-25-ylcarbamoyl)oxirane-2-carboxylate (17)



Prepared by standard SPPS. Product was obtained as white solid.

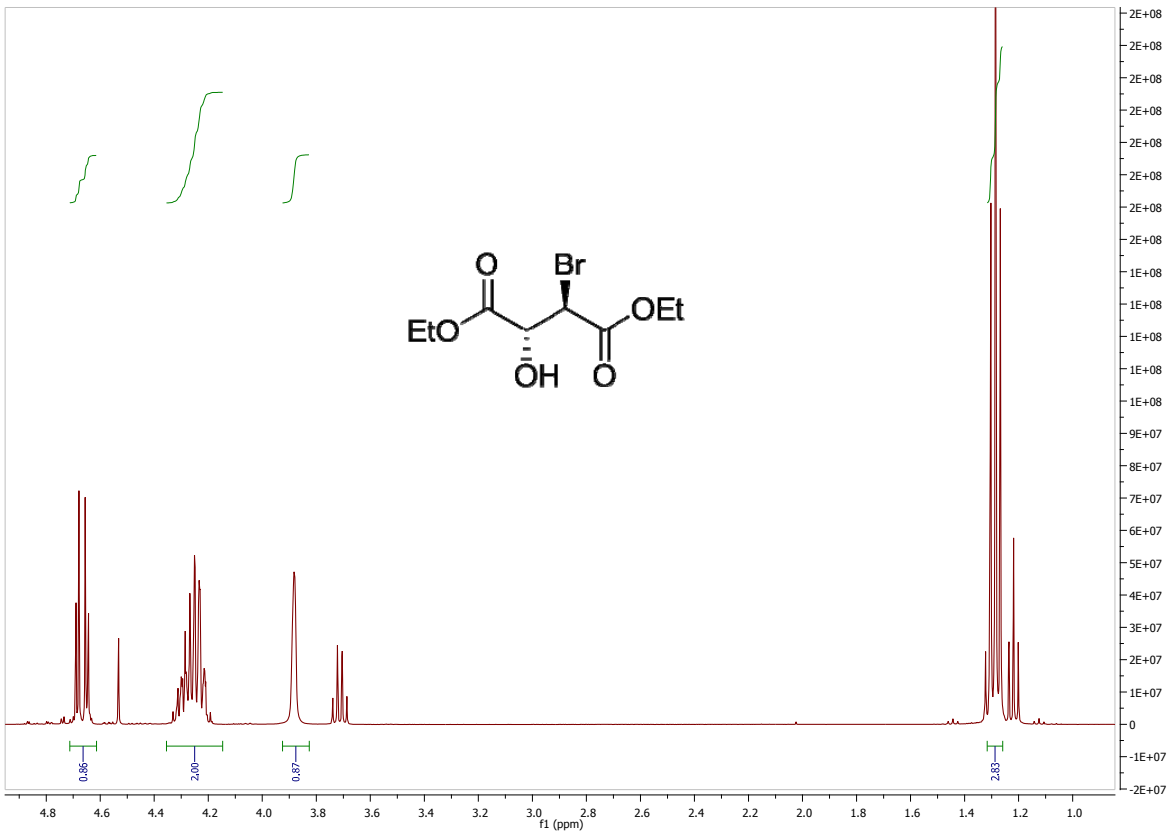
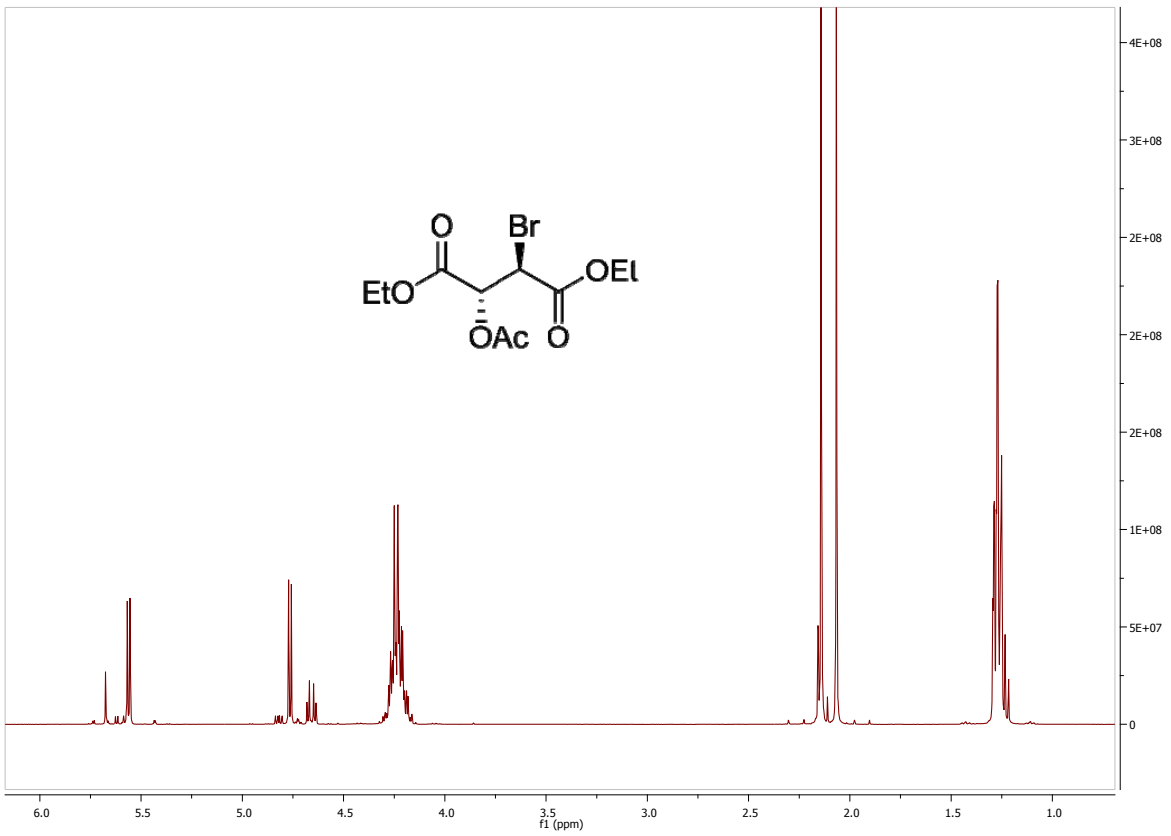
δ_{H} /ppm (400 MHz; D_2O): 4.51 (1H, dd, $J = 5.00, 7.88$, SCH_2CH), 4.32 (1H, dd, $J = 4.35, 7.74$, SCHCH), 4.29 (1H, t, $J = 8.76$, NHCHArgCO), 4.22-4.15 (3H, m, NHCHLeuCO , OCH_2CH_3), 3.69 (1H, s, CHOCHCOOEt), 3.63 (1H, s, CHOCHCOOEt), 3.59-3.55 (8H, m, $\text{O}(\text{CH}_2)_2\text{O}(\text{CH}_2)_2\text{O}$), 3.48, 3.45 (4H, t, $J = 6.36$, $\text{NH}(\text{CH}_2)_2\text{CH}_2\text{O}$), 3.26-3.09 (7H, m, $3 \times \text{CH}_2\text{NH}$, SCHCH), 2.89 (1H, dd, $J = 4.88, 13.13$, SCH_2CH equatorial, same site with CHCH), 2.68 (1H, d, $J = 13.04$, SCH_2CH axial), 2.15 (2H, t, $J = 7.19$, NHCOCH_2), 1.76-1.43 (15H, m, $(\text{CH}_2)_3\text{CHS}$, $\text{CH}_2\text{CH}(\text{CH}_3)_2$, $\text{PEG-CH}_2\text{CH}_2\text{NHCO}$, $\text{CH}(\text{CH}_3)_2$, $(\text{CH}_2)_2\text{CH}_2\text{NHCHNH}_2$), 1.35-1.26 (2H, m, $\text{CONHCH}_2\text{CH}_2\text{CH}_2\text{O}$), 1.20 (3H, t, $J = 7.08$, OCH_2CH_3), 0.84, 0.80 (6H, d, $J = 5.38$ and 5.34 , $\text{CH}(\text{CH}_3)_2$). **ESI-MS:** $m/z = 858.4786$ ($[\text{M} + \text{H}]^+$).

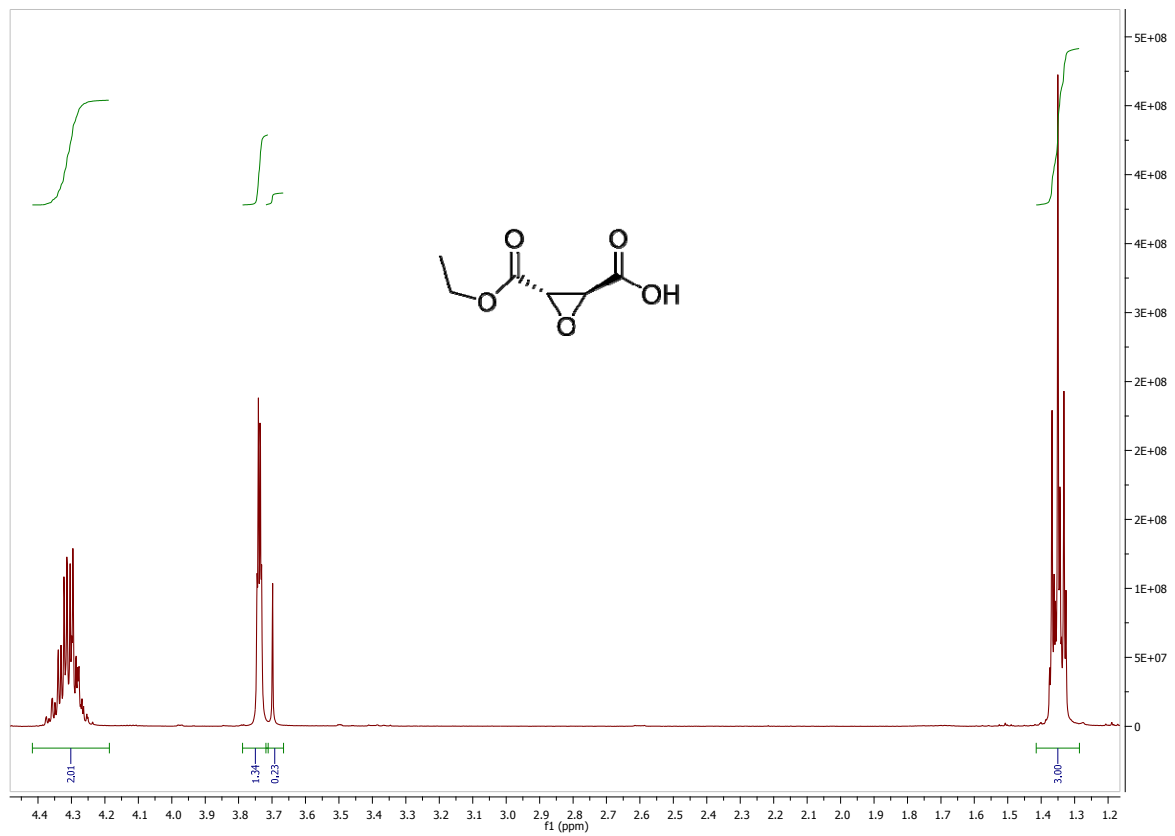
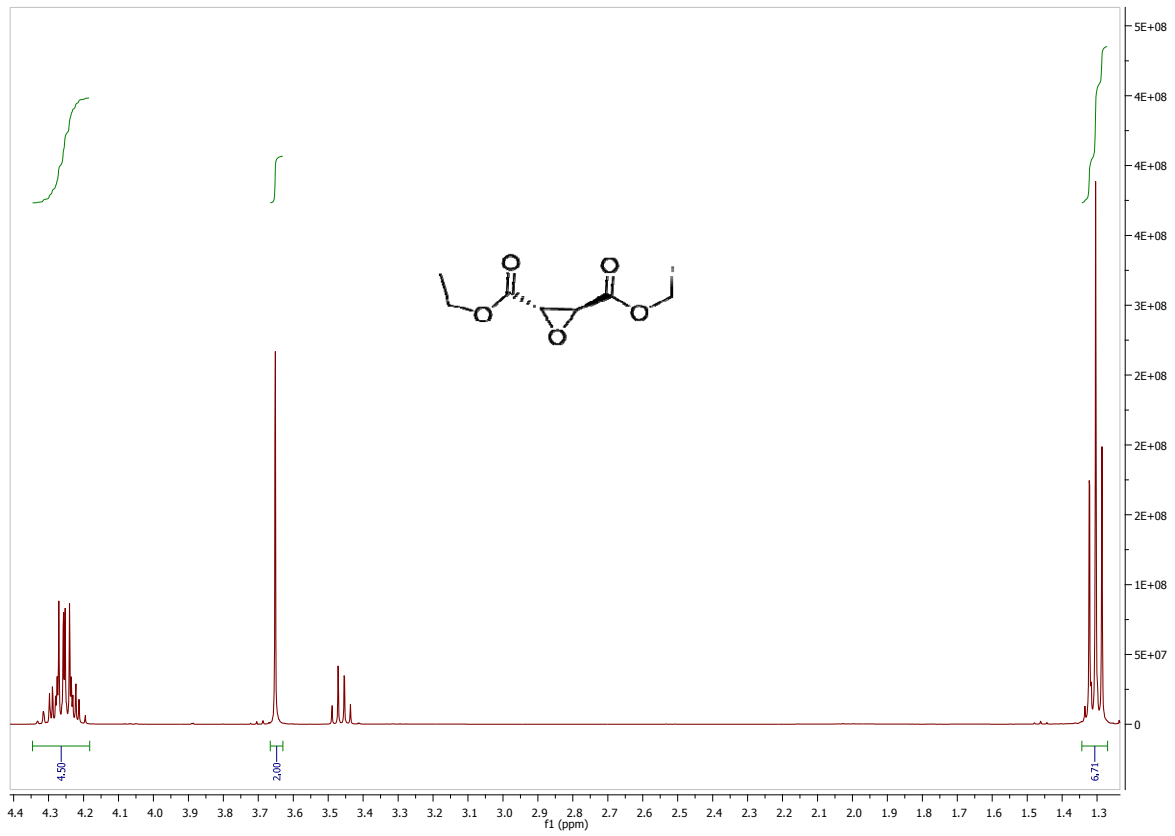
EtEP-Leu-Arg-Pra-NH₂(3)

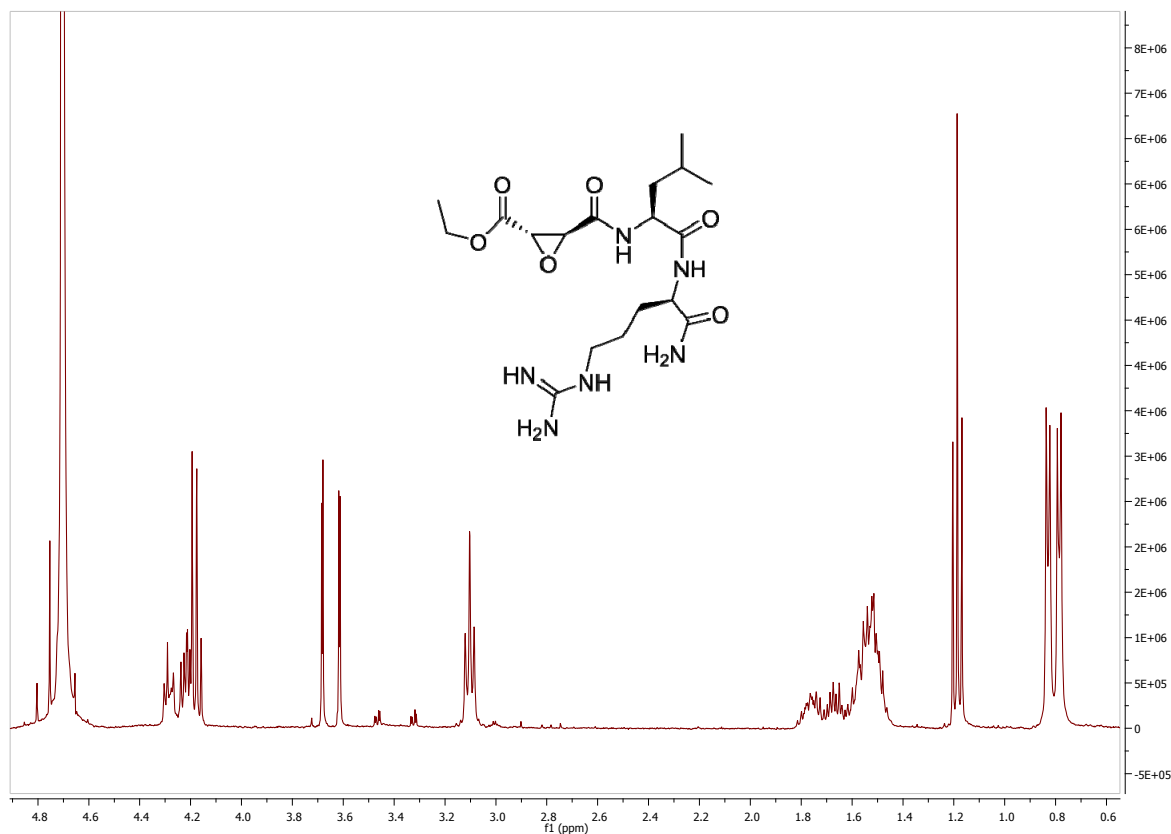


Prepared by standard SPPS. Product was obtained as white solid.

δ_{H} /ppm (400 MHz; D_2O): 4.41 (1H, t, $J = 6.54$, NHCHCO), 4.31 (2H, m, NHCHCO), 4.21 (2H, q, $J = 7.15, 7.18$, CH_2OCO), 3.71 (1H, d, $J = 1.78$, CHOCHCOOEt), 3.65 (1H, d, $J = 1.81$, CHOCHCOOEt), 3.13 (2H, t, $J = 6.91$, $\text{CH}_2\text{NHCNHNH}_2$), 2.66-2.64 (2H, m, CH_2CCH), 2.37 (1H, t, $J = 2.55$, CCH), 1.83-1.66, 1.62-1.50 (7H, m, $\text{CH}_2\text{CH}(\text{CH}_3)_2$, CH_2), 1.21 (3H, t, $J = 7.17$, CH_3CH_2), 0.85, 0.81 (6H, d, $J = 5.75, 5.72$, $\text{CH}(\text{CH}_3)_2$). **ESI-MS:** $m/z = 524.2832$ ($[\text{M} + \text{H}]^+$). **R_t:** 14.05 (Method C)







References for Supporting Information

1. Bogyo, M., Verhelst, S., Bellingard-Dubouchaud, V., Toba, S., and Greenbaum, D. (2000) Selective targeting of lysosomal cysteine proteases with radiolabeled electrophilic substrate analogs, *Chem. Biol.* **7**, 27-38.
2. Calabi, E., and Fairweather, N. (2002) Patterns of sequence conservation in the S-layer proteins and related sequences in *Clostridium difficile*, *J. Bacteriol.* **184**, 3886-3897.
3. Eidhin, D., Ryan, A., Doyle, R., Walsh, J. B., and Kelleher, D. (2006) Sequence and phylogenetic analysis of the gene for surface layer protein, *slpA*, from 14 PCR ribotypes of *Clostridium difficile*, *J Med Microbiol* **55**, 69-83.
4. Sebahia, M., Wren, B. W., Mullany, P., Fairweather, N. F., Minton, N., Stabler, R., Thomson, N. R., Roberts, A. P., Cerdeno-Tarraga, A. M., Wang, H., Holden, M. T., Wright, A., Churcher, C., Quail, M. A., Baker, S., Bason, N., Brooks, K., Chillingworth, T., Cronin, A., Davis, P., Dowd, L., Fraser, A., Feltwell, T., Hance, Z., Holroyd, S., Jagels, K., Moule, S., Mungall, K., Price, C., Rabinowitsch, E., Sharp, S., Simmonds, M., Stevens, K., Unwin, L., Whithead, S., Dupuy, B., Dougan, G., Barrell, B., and Parkhill, J. (2006) The multidrug-resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome, *Nat. Genet.* **38**, 779-786.
5. Fagan, R. P., Albesa-Jove, D., Qazi, O., Svergun, D. I., Brown, K. A., and Fairweather, N. F. (2009) Structural insights into the molecular organization of the S-layer from *Clostridium difficile*, *Mol. Microbiol.* **71**, 1308-1322.
6. Wright, A., Wait, R., Begum, S., Crossett, B., Nagy, J., Brown, K., and Fairweather, N. (2005) Proteomic analysis of cell surface proteins from *Clostridium difficile*, *Proteomics* **5**, 2443-2452.
7. Heal, W. P., Wickramasinghe, S. R., Leatherbarrow, R. J., and Tate, E. W. (2008) N-Myristoyl transferase-mediated protein labelling in vivo, *Org Biomol Chem* **6**, 2308-2315.