Data Access Statement: The coordinates of the CysF X-ray crystal structures were deposited to PDB 8RA0. The AlphaFold structure of BhCysFE can be accessed at https://www.alphafold.ebi.ac.uk/entry/A0A562R406 (AF-A0A562R406-F1-model_v4). Structures used for modeling and docking studies can be accessed from PDB 5BSM, 5BSR, 5WM3, 5IE3, 4FUT, 4GXR and 4GXQ. All proteins characterized in this study can be accessed from UniProt using the accession codes presented in Supplementary Tables 1–3 and their synthetic gene sequences are provided in Supplementary Data 1. The remaining data are available in the main text or the Supplementary Information. Correspondence and requests for materials should be addressed to J.M.

Supplementary Information

Cryptic Enzymatic Assembly of Peptides Armed with β-Lactone Warheads

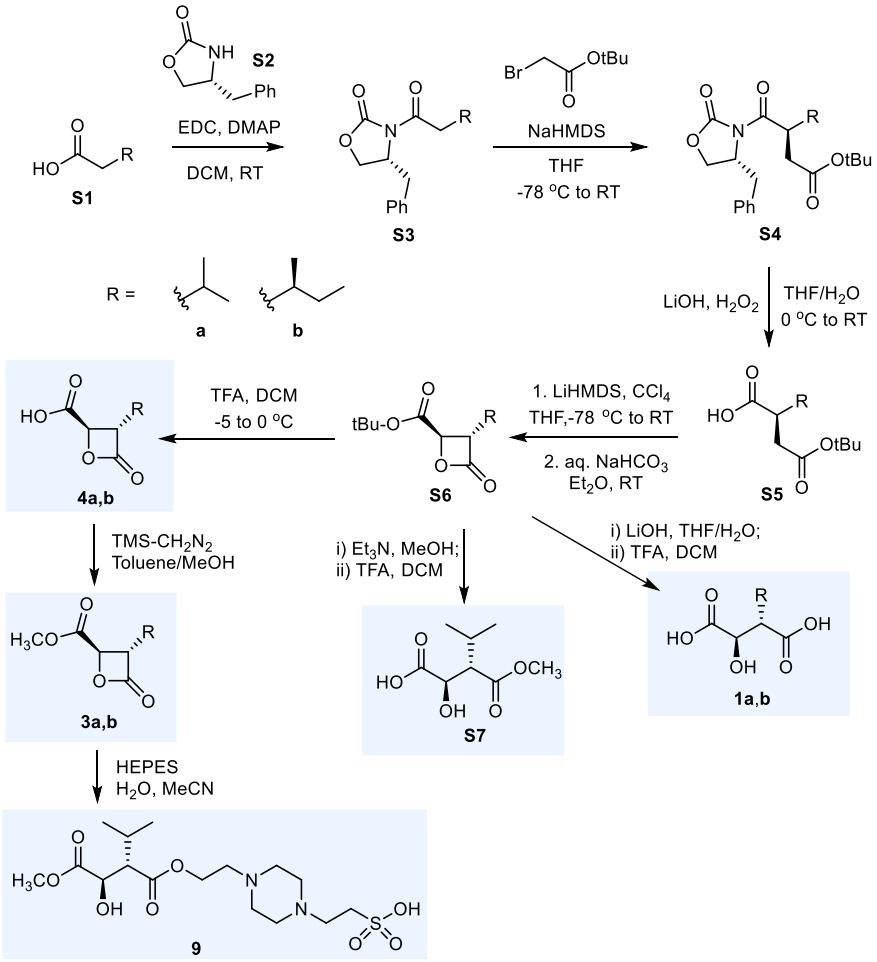
Guangcai Xu, Daniele Torri, Sebastian Cuesta-Hoyos, Deepanjan Panda, Luke R. L. Yates, Rémi Zallot, Kehan Bian, Dongxu Jia, Andreea I. Iorgu, Colin Levy, Sarah A. Shepherd and Jason Micklefield*

*Corresponding author, email: jason.micklefield@manchester.ac.uk

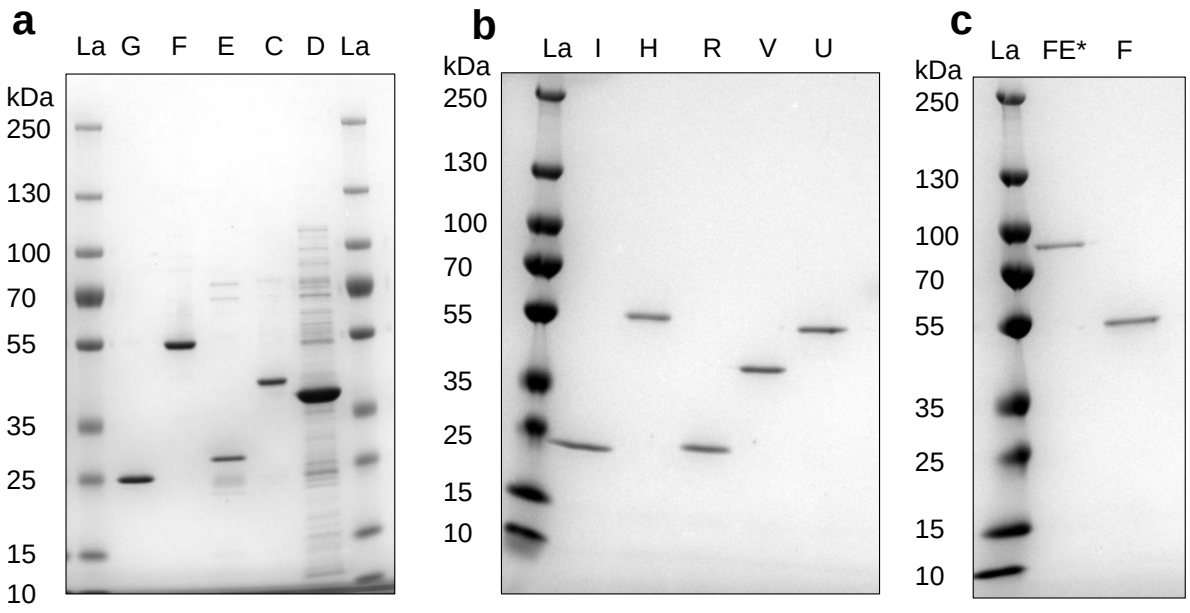
Affiliation: Department of Chemistry and Manchester Institute of Biotechnology, The University of Manchester, 131 Princess Street, Manchester, M1 7DN, UK.

Contents

1. Supplementary figures and tables



Supplementary Fig. 1. Synthesis of standards **1a/b**, **3a/b**, **4a/b**, **9**, and **S7** for the validation of β-lactone warhead assembly pathway. **1a** is also commercially available. For detailed synthesis procedure see method section 5 (synthesis of substrates and standards).



Supplementary Fig. 2. SDS-PAGE gel image of all purified enzymes in this study. (a) All purified enzyme in the Cys pathway. **(b)** All purified enzyme in the Bel pathway. **(c)** Comparison of fused-CysFE and CysF. La = protein ladder, G = CysG (26.5 kDa), F = CysF (54.3 kDa), E = CysE (30.5 kDa), C = CysC (43.0 kDa), D = CysD (37.9 kDa), I = BelI (25.3 kDa), H = BelH (54.3 kDa), R = BelR (29.3 kDa), V = BelV (43.1 kDa), U = BelU (47.4 kDa), FE* = *Bh*CysFE fusion (84.4 kDa).

Supplementary Table 1. Cys pathway[1] proteins

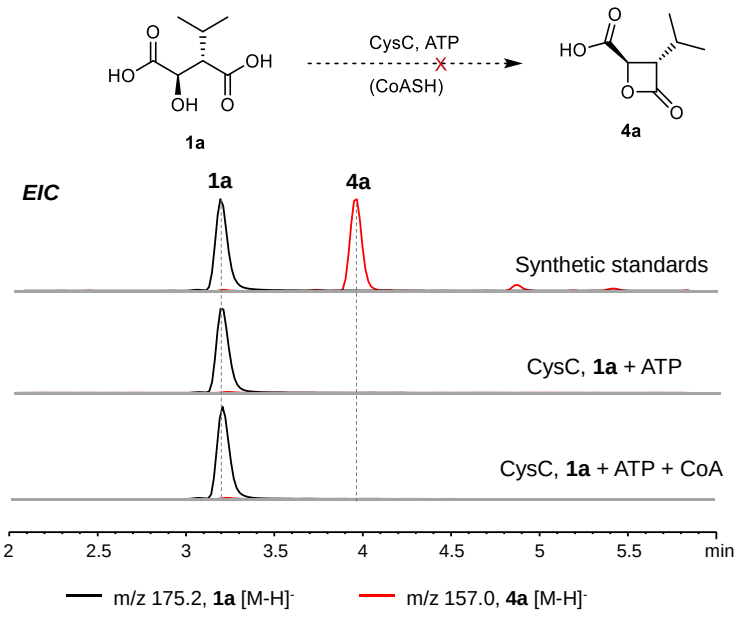
Name	AA[2]	Function[4]	Expression vector	Uniprot ID	His-tag terminal	MW (kDa)[3]
CysG	228	Methyltransferase	pET-28a(+)	A0A1W6R556	C	26.53
CysF	491	Lactone synthetase	pET-21a(+)	A0A1W6R555	C	54.28
CysE	271	Hydrolase	pET-28a(+)	A0A1W6R564	C	30.48
CysC	392	Amide bond synthetase	pET-28a(+)	A0A1W6R559	C	43.04
CysD	333	ATP-grasp ligase	pET-21a(+)	A0A1W6R558	C	37.87

[1] Organism: *Kitasatospora cystarginea* NRRL B16505. [2] Amino acid length of the original protein without His-tag. [3] Molecular weight of the protein with His-tag. [4] Functions determined in this study.

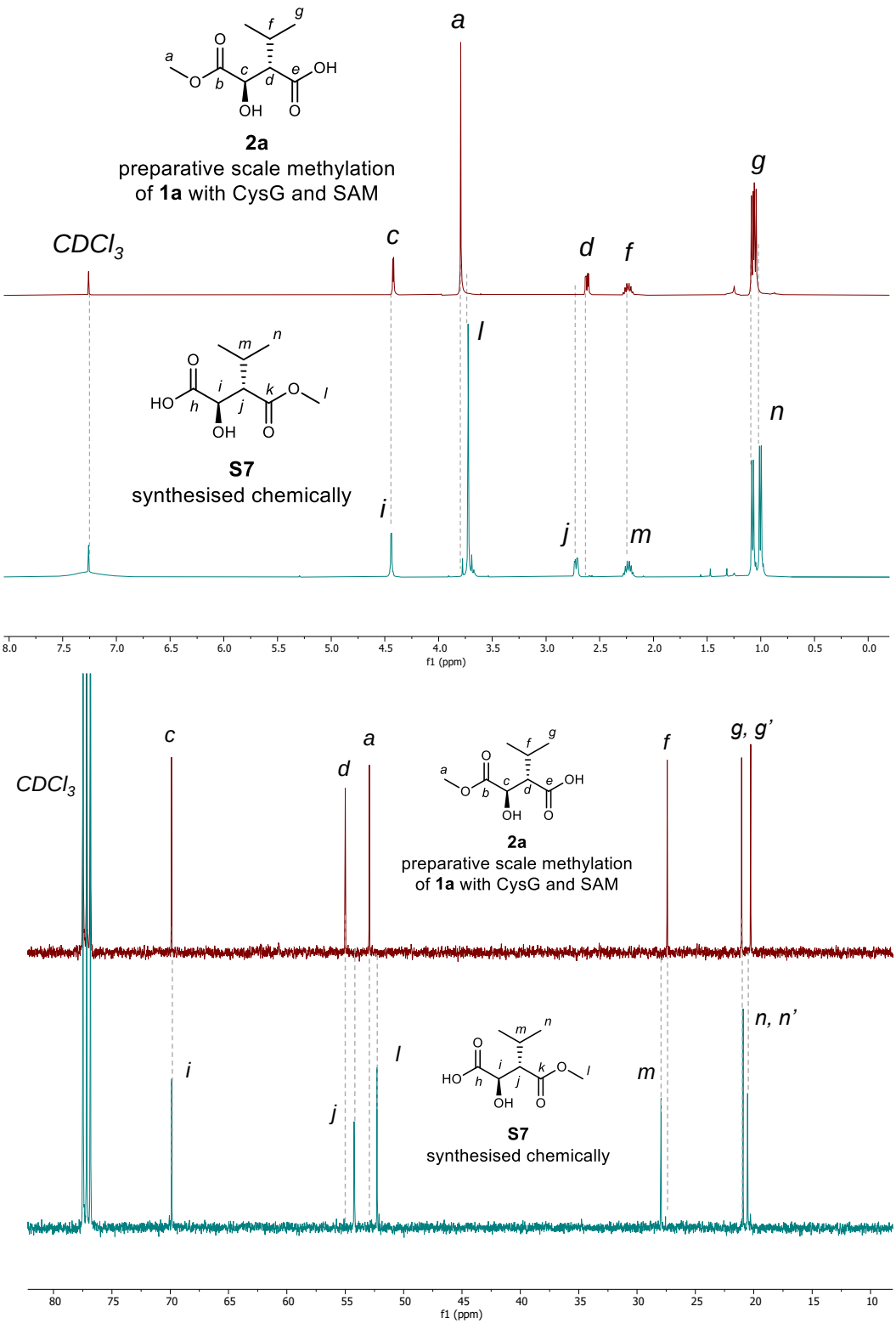
Supplementary Table 2. Bel pathway[1] proteins

Name	AA[2]	Function[4]	Expression vector	UniProt ID	His-tag terminal	MW (kDa)[3]
BelI	228	Methyltransferase	pET-28a(+)	A0A1W6R583	C	25.32
BelH	491	Lactone synthetase	pET-28a(+)	A0A1W6R584	N	54.26
BelR	269	Hydrolase	pET-28a(+)	A0A1W6R592	C	29.27
BelV	389	Amide bond synthetase	pET-28a(+)	A0A1W6R589	C	43.10
BelU	420	ATP-grasp ligase	pET-21a(+)	A0A1W6R594	C	47.43

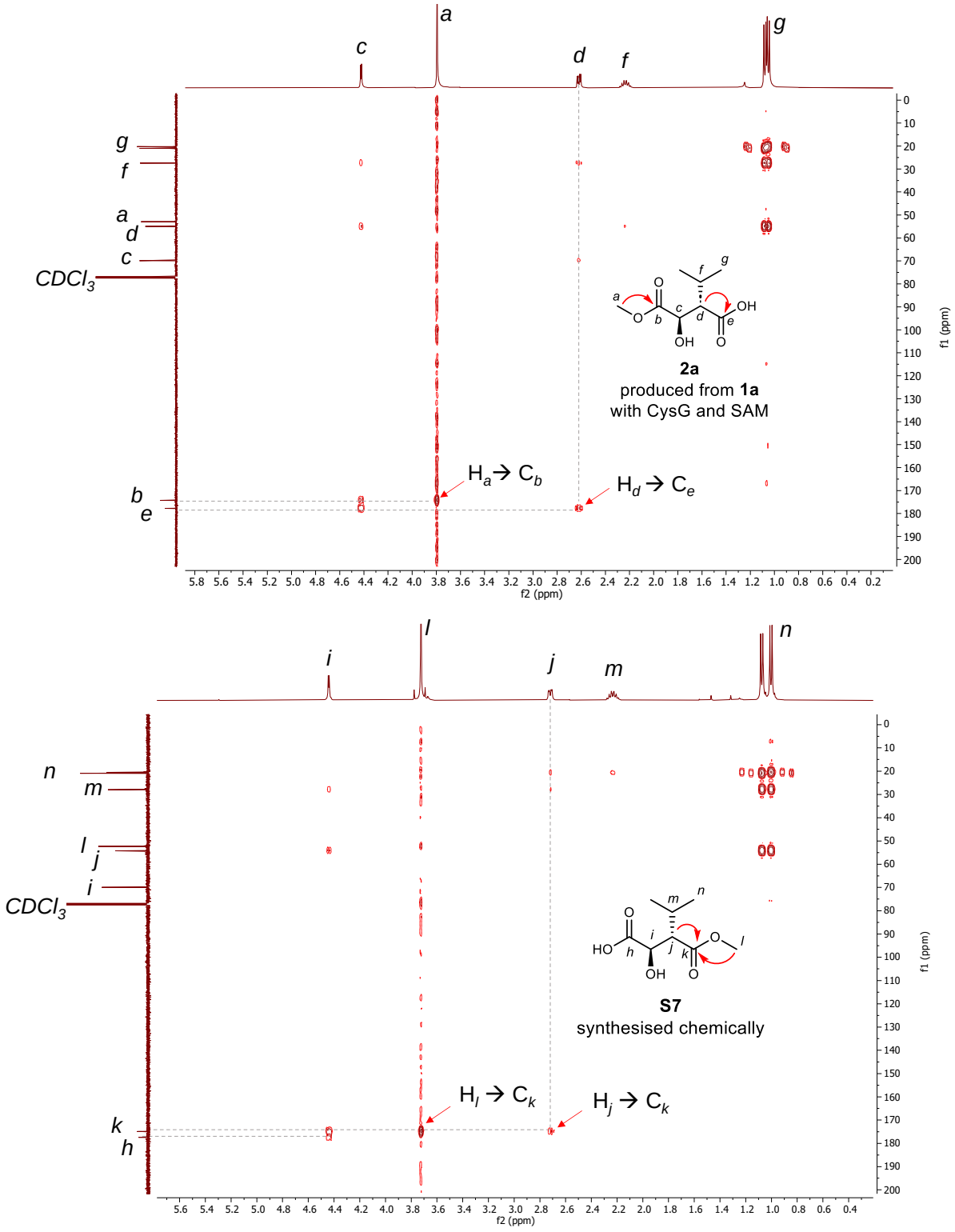
[1] Organism: *Streptomyces* sp. UCK 14. [2] Amino acid length of the original protein without His-tag. [3] Molecular weight of the protein with His-tag. [4] Functions determined in this study.



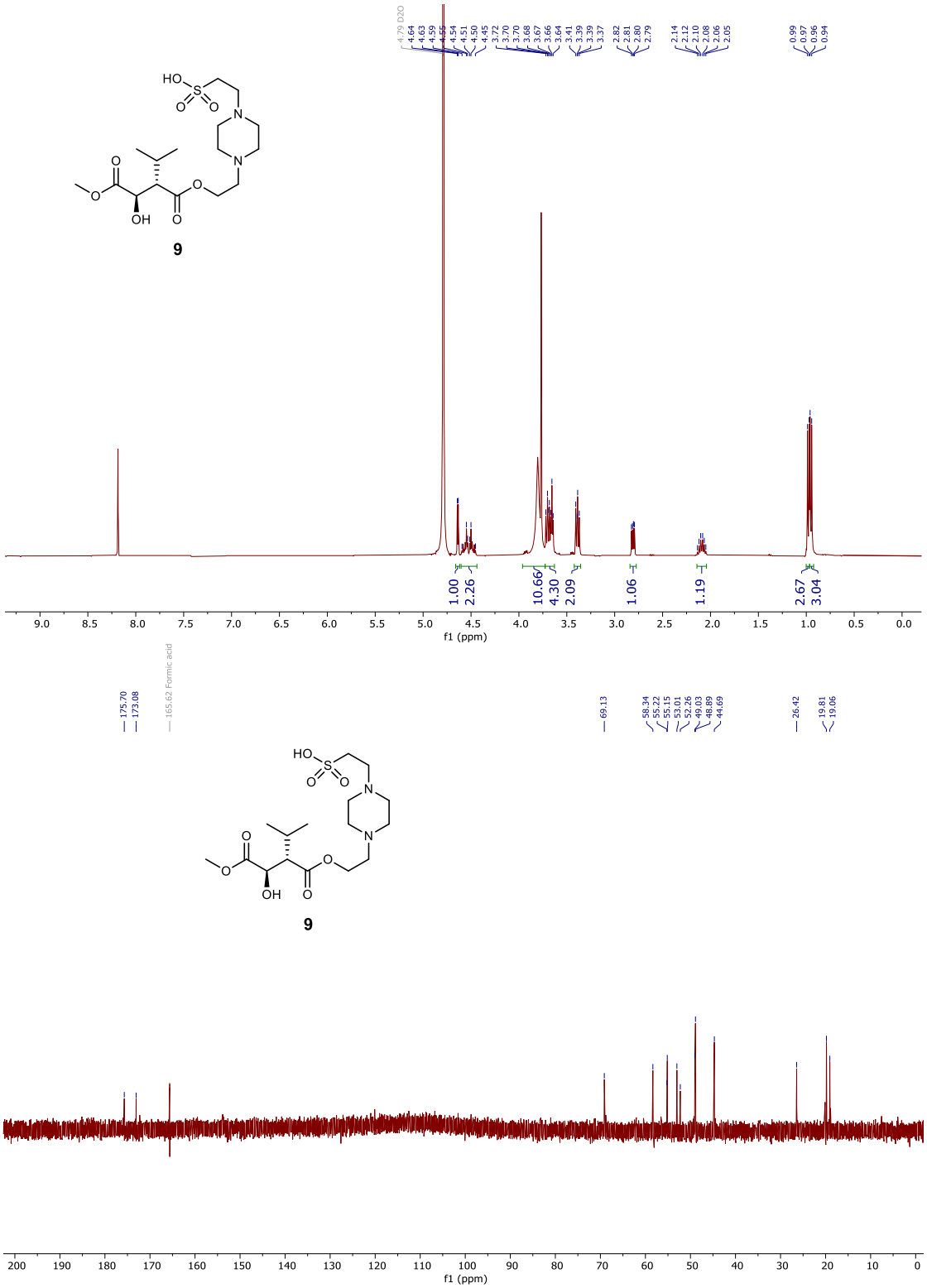
Supplementary Fig. 3. Testing the proposed lactonisation activity of CysC. For reaction conditions see experimental method (enzyme assays).



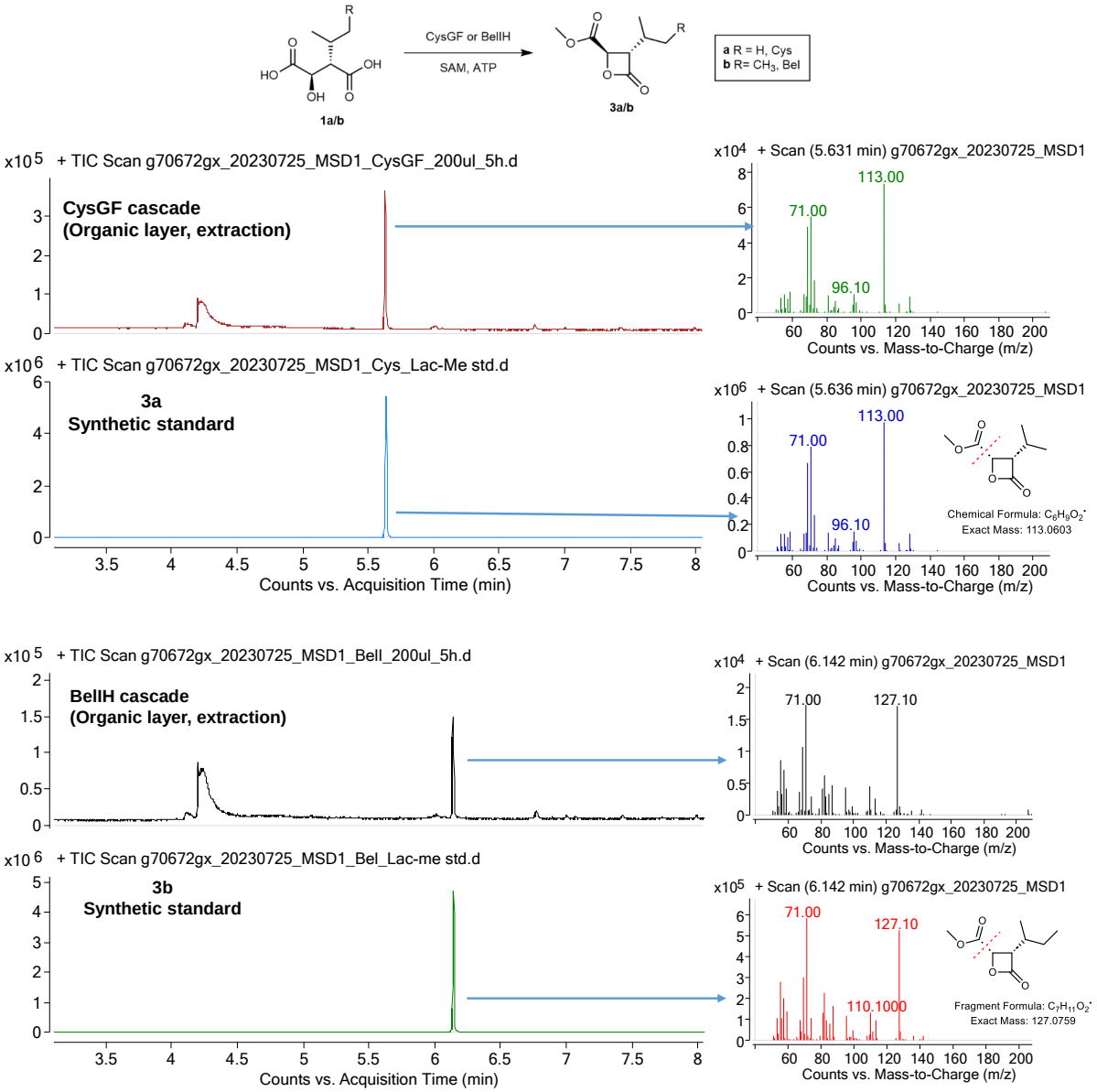
Supplementary Fig. 4. Comparison of the ^{1}H and ^{13}C NMR spectra between enzymatic (CysG) product **2a** and the synthetic standard **S7**.



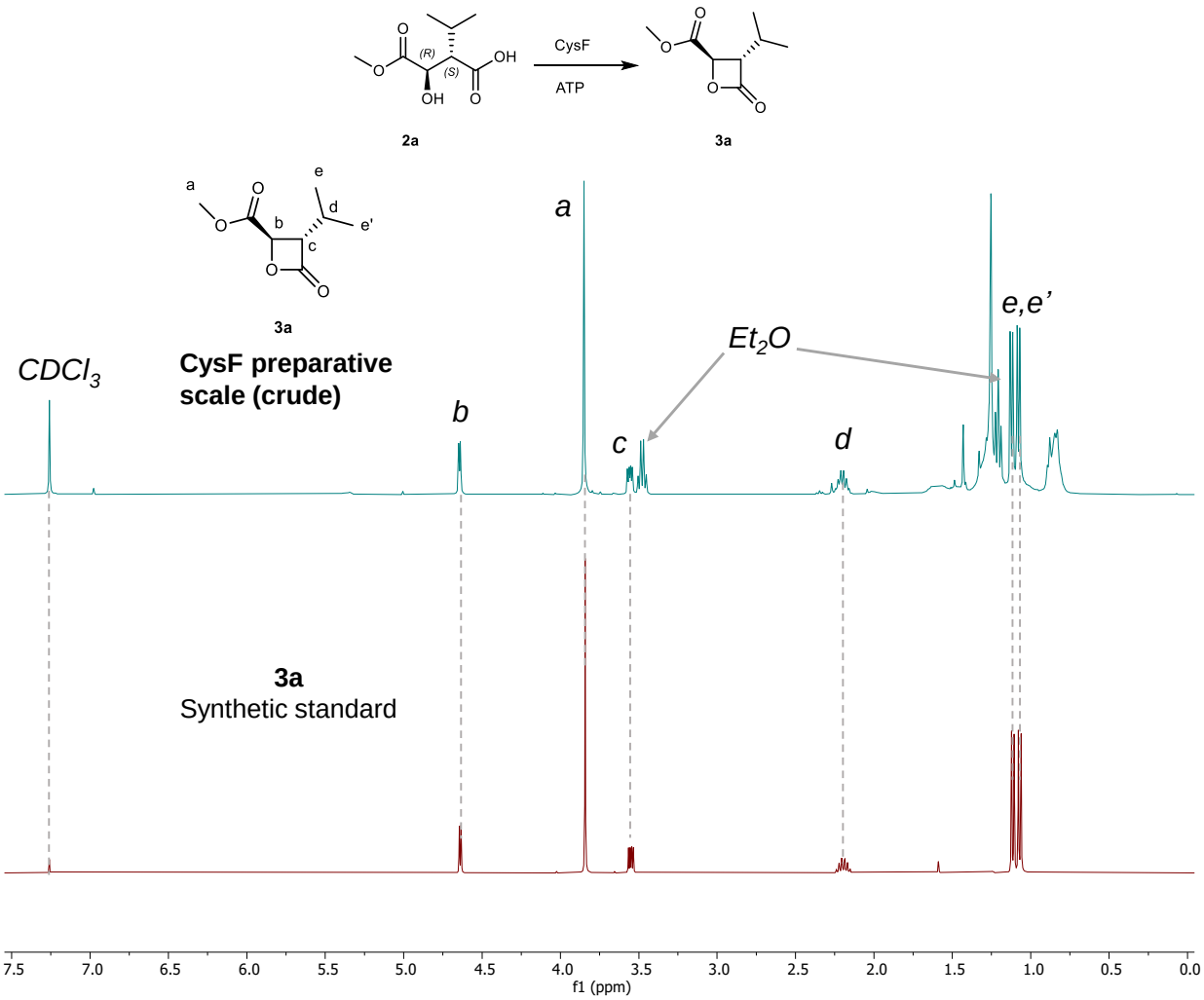
Supplementary Fig. 5. Comparison of the HMBC spectra of **2a** (enzymatic, top) and **S7** (synthetic standard, bottom).



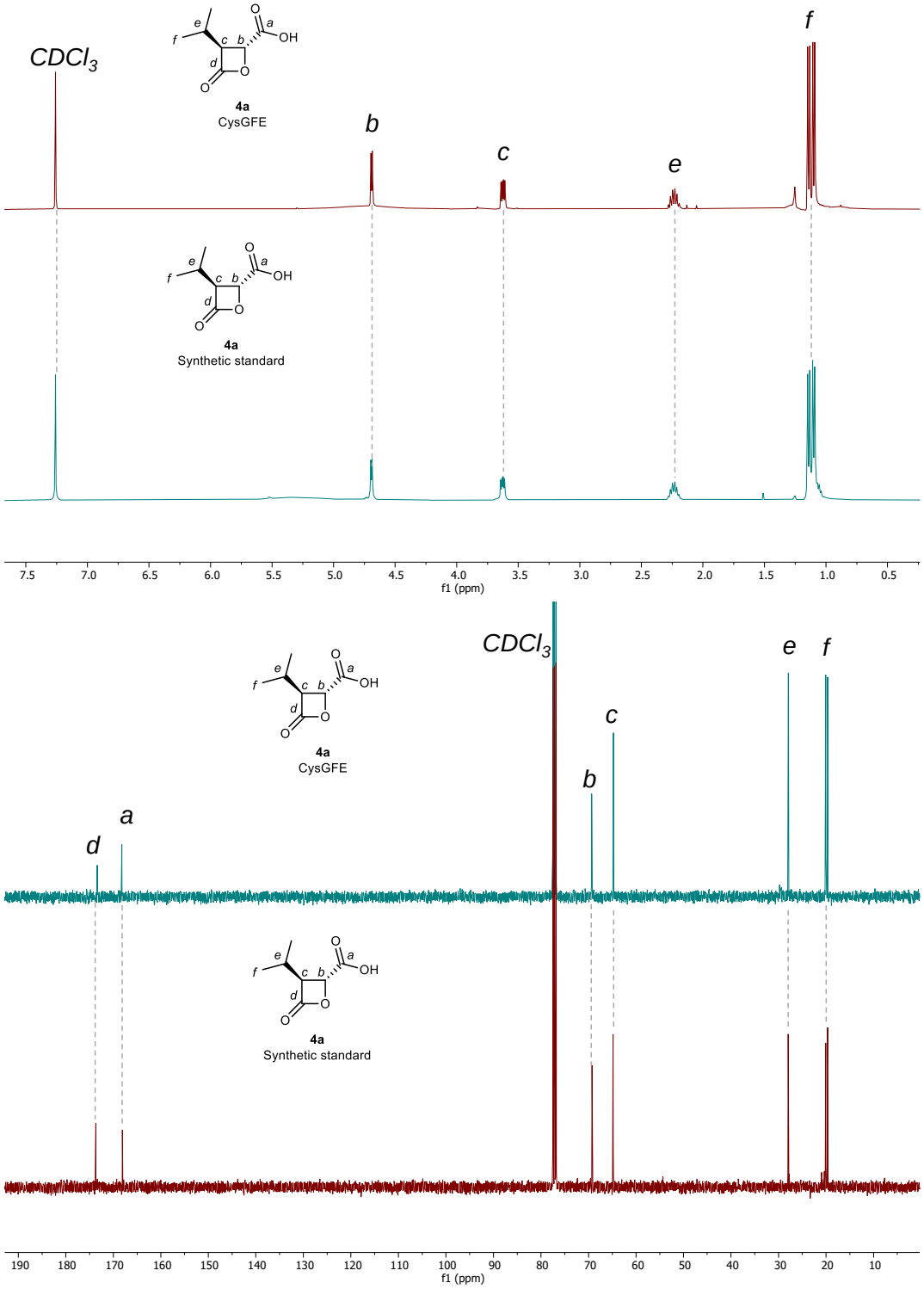
Supplementary Fig. 6. ^{1}H and ^{13}C NMR spectra of the HEPES buffer adduct **9**.



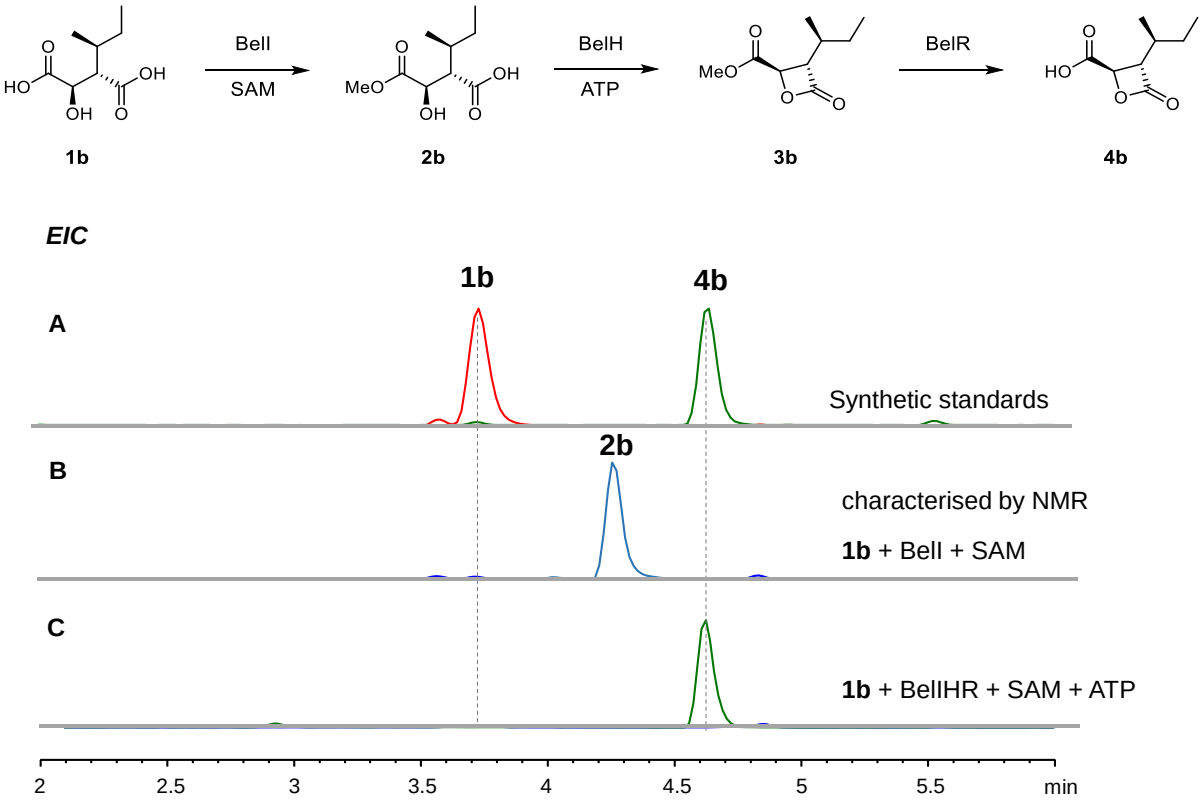
Supplementary Fig. 7. GC-MS analysis of the CysGF and BelIH cascade confirmed the formation of **3a/b**.



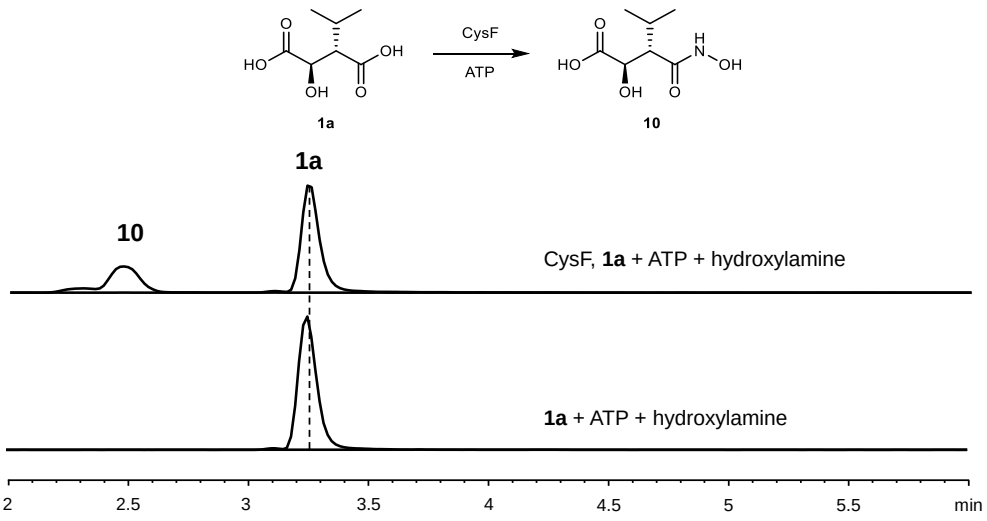
Supplementary Fig. 8. Comparison of the ^{1}H NMR of **3a** prepared enzymatically (crude product) using CysF, and chemically by methylation of **4a** using TMS-diazomethane (column purified).



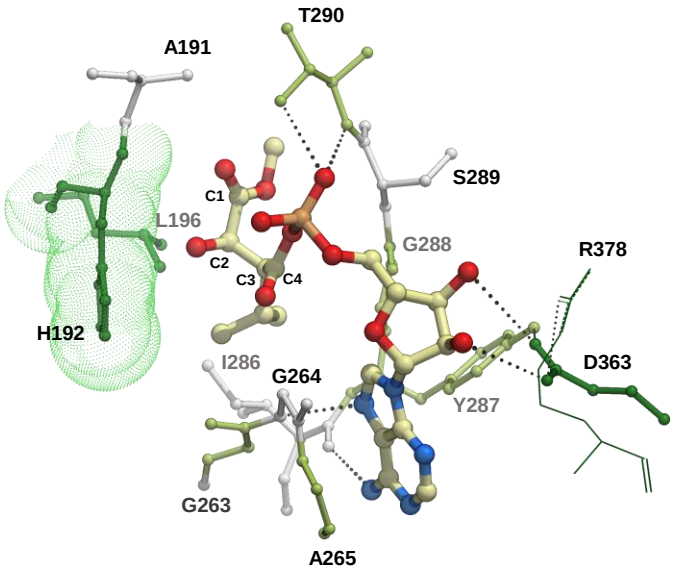
Supplementary Fig. 9. Comparison of the ^{1}H and ^{13}C NMR of **4a** between enzymatically (CysGFE cascade) produced and chemically synthesised.

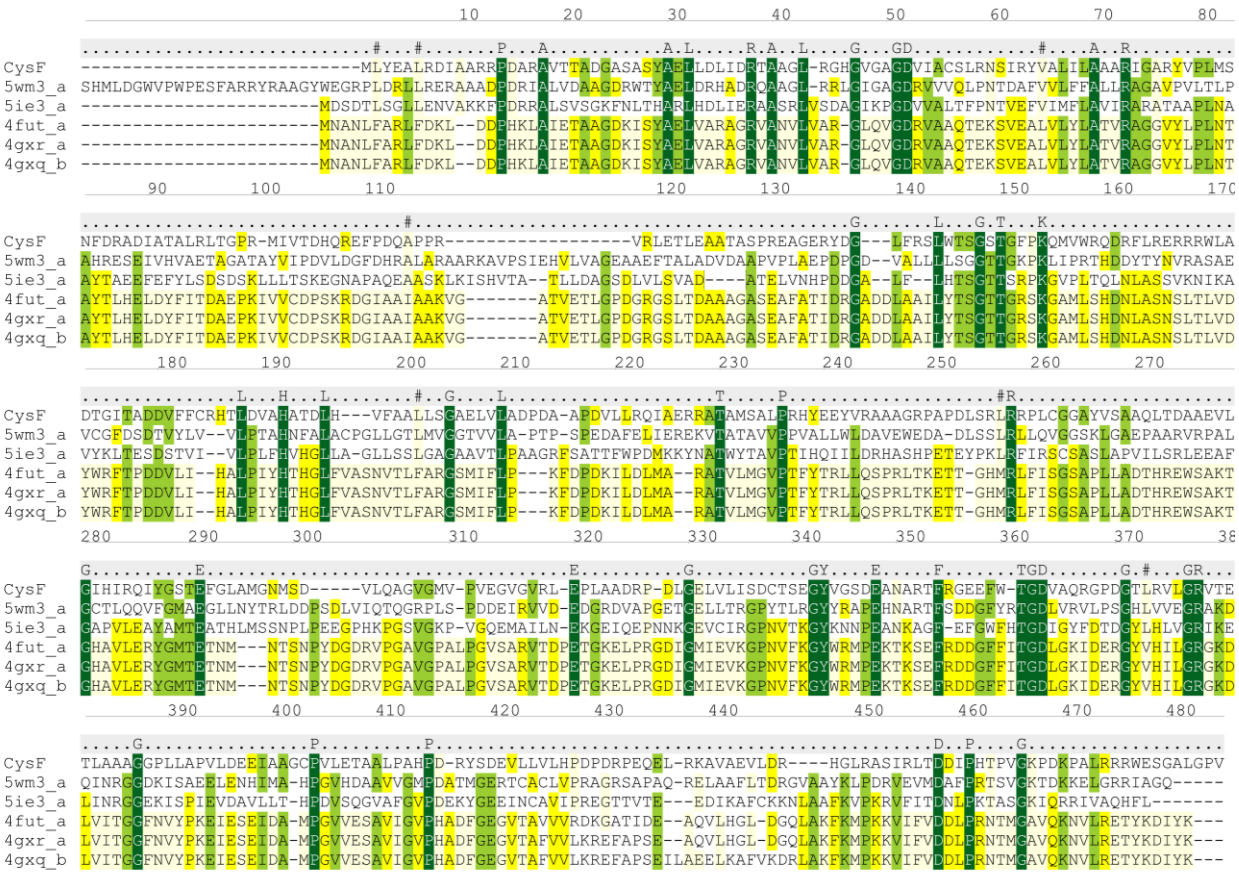


Supplementary Fig. 10. Bel pathway lactone formation LC-MS assay. See experimental information (enzyme assay) for experimental details.



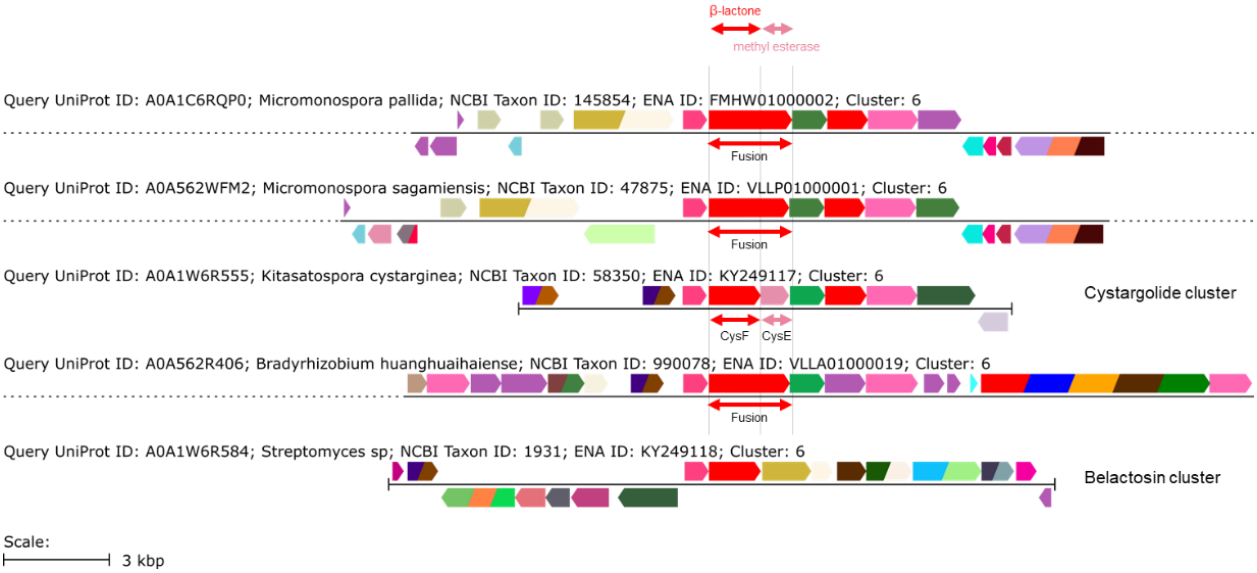
Supplementary Fig. 11. CysF catalysed reaction of hydroxylamine with **1a.**





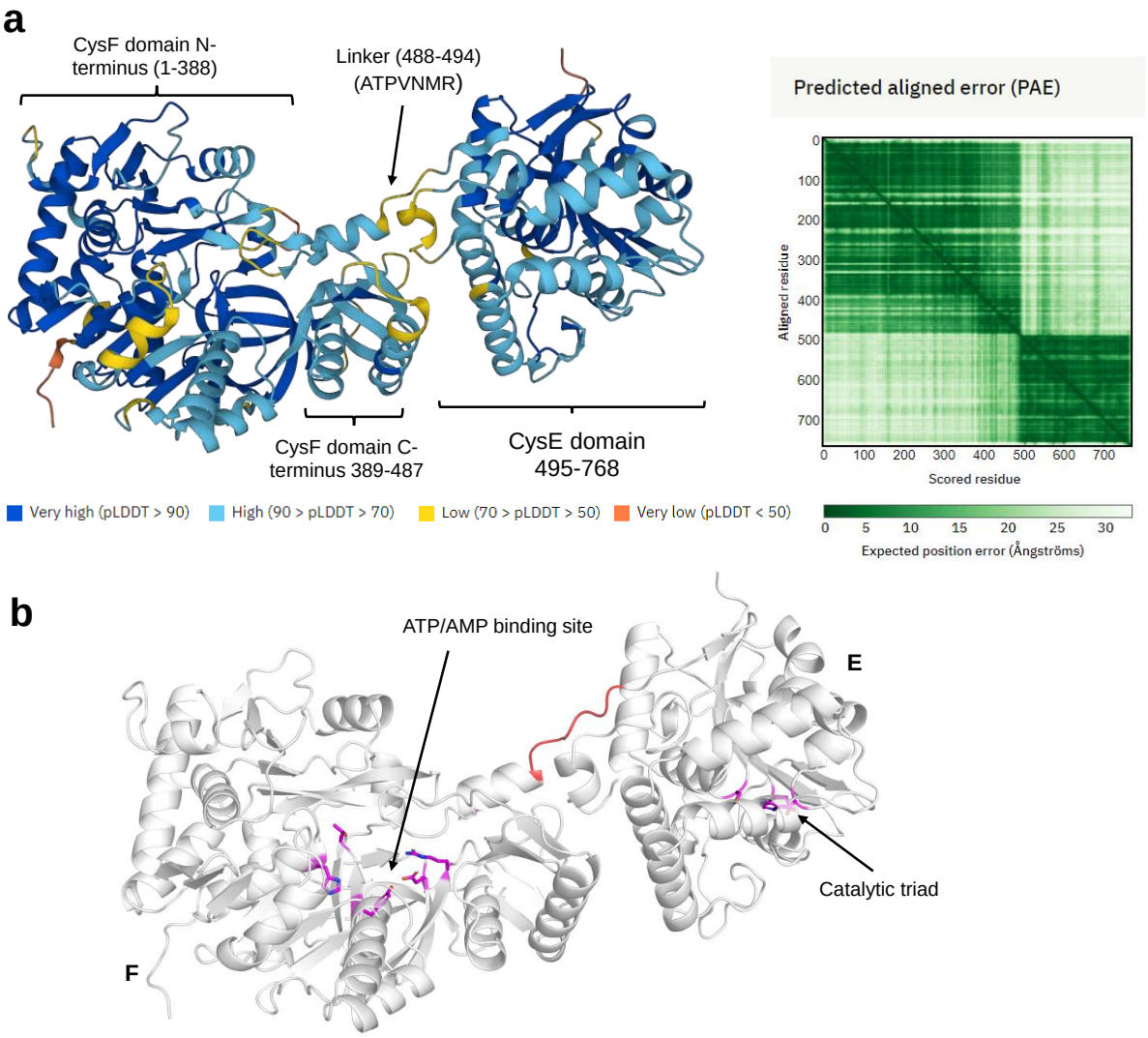
Supplementary Fig. 12. *In silico* docking of **2a**-AMP in the crystal structural of CysF and sequence alignment of CysF with other adenylate-forming enzymes. Pairwise levels of sequence identities between CysF and the related adenylate-forming enzymes (5IE3, 4GXR, 5WM3, 4GXQ & 4FUT)[39–43] are between 18 – 30% and the key residues involved in the binding of AMP are highly conserved (T290, D363, R378). The docked conformation of **2a**-AMP suggests H192

may perform the role of a general base during β-lactone formation. The close proximity of the C1-methyl ester to the phosphoryl group, of the docked **2a**-AMP, further underscores the potential for electrostatic repulsion if the C1-carboxylate group was not methylated. Polyethylene glycol (PEG4) is observed to occupy a region of the putative active site in the crystal structure of CysF, which is likely to be an artifact of the crystallisation process. Additional electron density is present in the crystal structure which is consistent with carboxylation resulting in an N-terminal carbamate stabilised by interactions to R131. Carboxylation is suggested to be a common, spontaneous and reversible post translation modification.[S1] Given the CysF N-terminus is not located near the active site, it is unlikely that this modification has any effect on the catalytic mechanism of the enzyme.

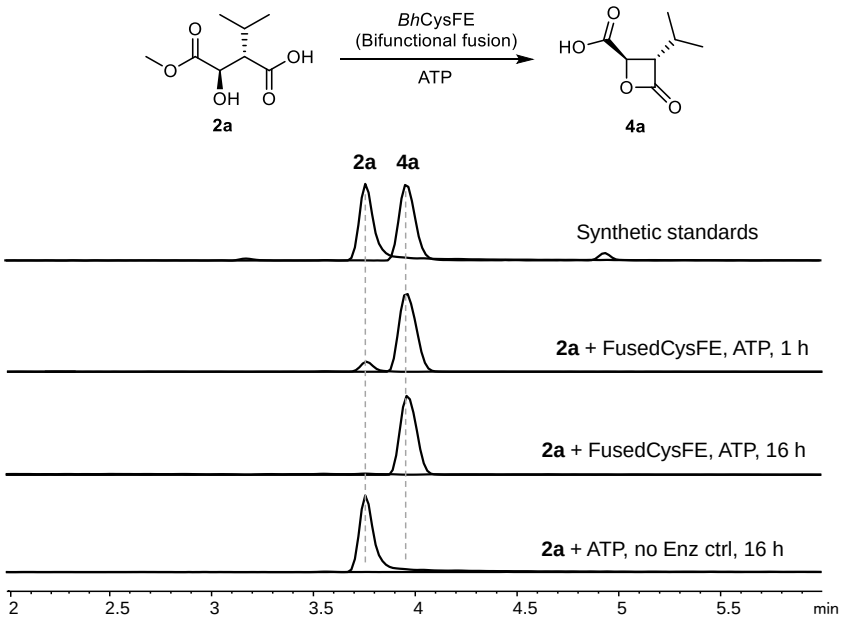


Name[1]	AA[2]	Organism	UniProt ID	MW[3] (*k*Da)
*Bh*CysFE	768	*Bradyrhizobium huanghuaihaiense*	A0A562R406	84.3
*Ms*CysFE	762	*Micromonospora sagamiensis*	A0A562WFM2	83.3
*Mp*CysFE	788	*Micromonospora pallida*	A0A1C6RQP0	85.8

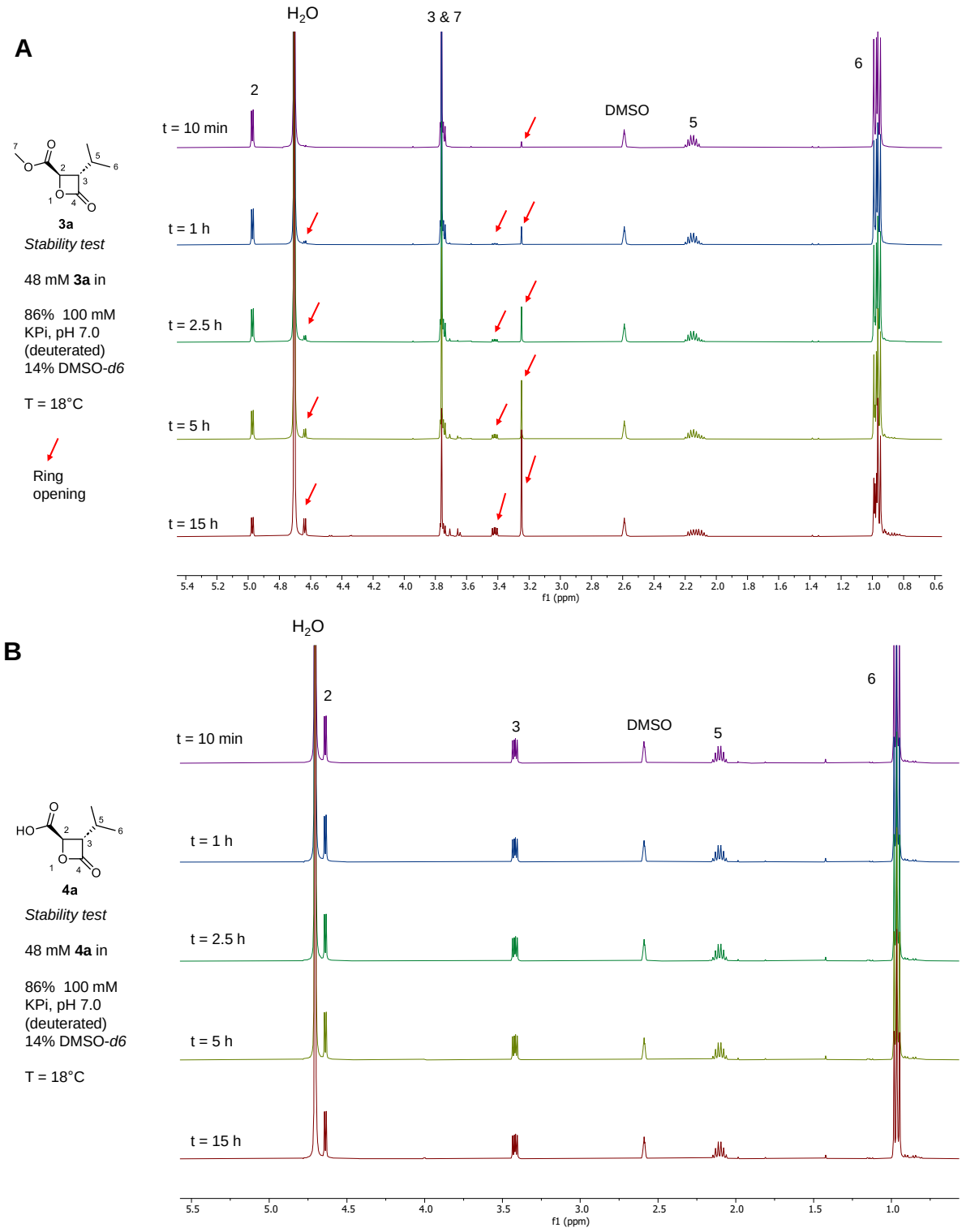
Supplementary Fig. 13. Examples of BGCs that contain fused-CysFE in nature. [1] bifunctional lactone synthetase-hydrolase. [2]: Amino acid length of the original protein without His-tag. [3] Molecular weight of the protein with His-tag.



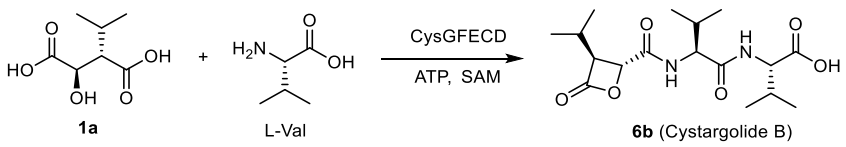
Supplementary Fig. 14. AlphaFold2 model of bifunctional enzyme *Bh*CysFE. (a) Structural prediction revealed two domains similar to CysF (aa. 1-487) and CysE (aa. 495-768), connected by a short linker (sequence ATPVNMR) colour coded showing regions with lower (orange) to higher accuracy (blue). (b) The position of the predicted active site residues of the F domain are highlighted in magenta based on sequence alignment and docking as described in supplementary fig. 12. The predicted catalytic triad present in the E domain Glu607, His739 and Ser582, which is highly conserved in hydrolase enzymes is also highlighted in magenta.

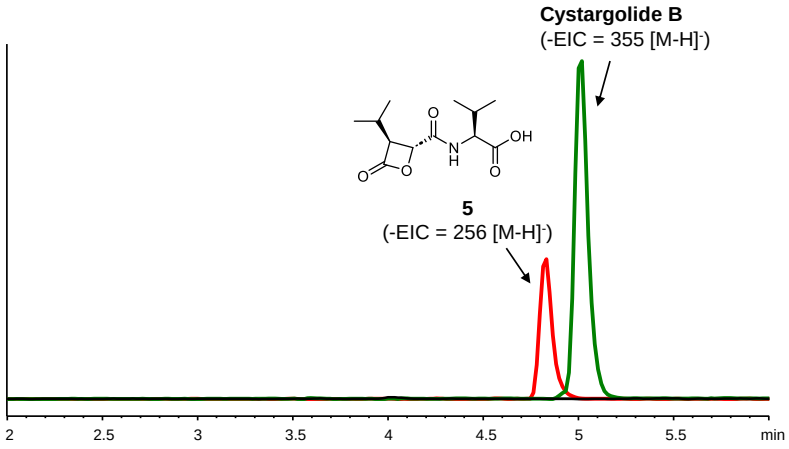


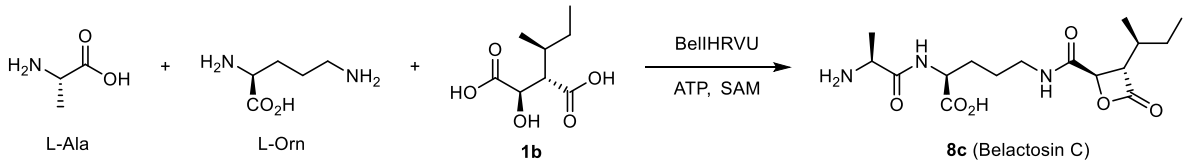
Supplementary Fig. 15. Assay of the bifunctional fusion *Bh*CysFE with **2a** results in direct formation of **4a**.

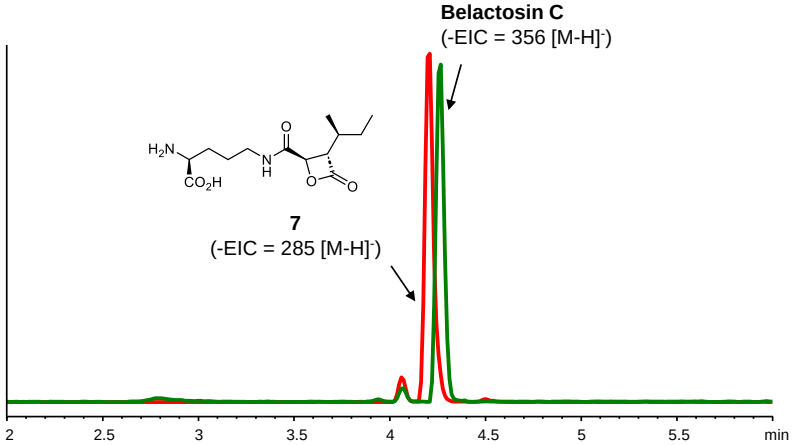


Supplementary Fig. 16. Stability comparison of **3a** and **4a** in KPi buffer. 0.02 mmol of **3a**/**4a** was dissolved in 90 µL of DMSO-*d6* and mixed with 540 µL of 100 mM deuterated KPi buffer. The mixture was transferred to NMR tube immediately and ^{1}H-NMR was recorded at different time points (T = 18 °C). Hydrolysis product formation is indicated by the red arrows.

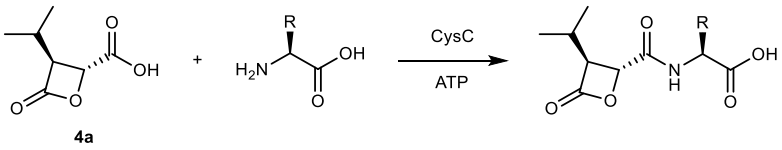






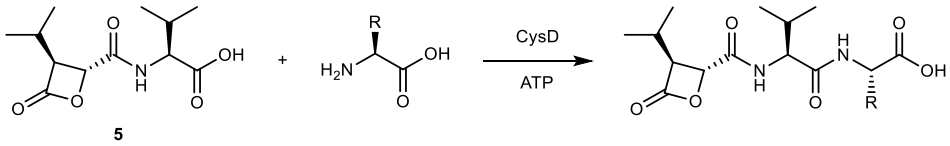


Supplementary Fig. 17. Reconstitution of the Cys and Bel pathway *in vitro*. LC-MS data. For detailed experimental procedure, see method section 3 (enzyme assay).



Accepted
(Product confirmed by LCMS) L-amino acids: Val, Ile, Leu, Met, Ala, Ser, Thr, Asn, Arg, His, Lys

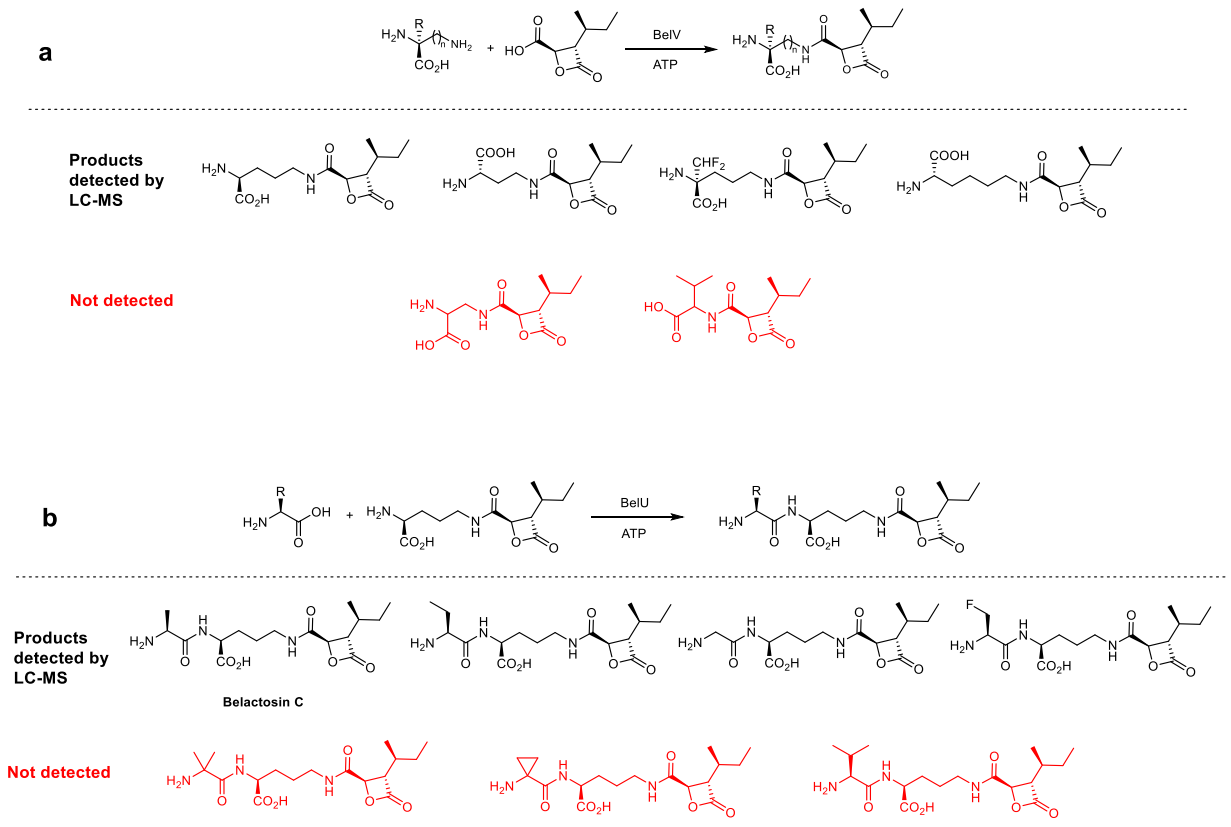
Not accepted
(No product detected) L-amino acids: Phe, Tyr, Trp, Cys, Gly, Pro, Gln, Asp, Glu
D-amino acids: Val



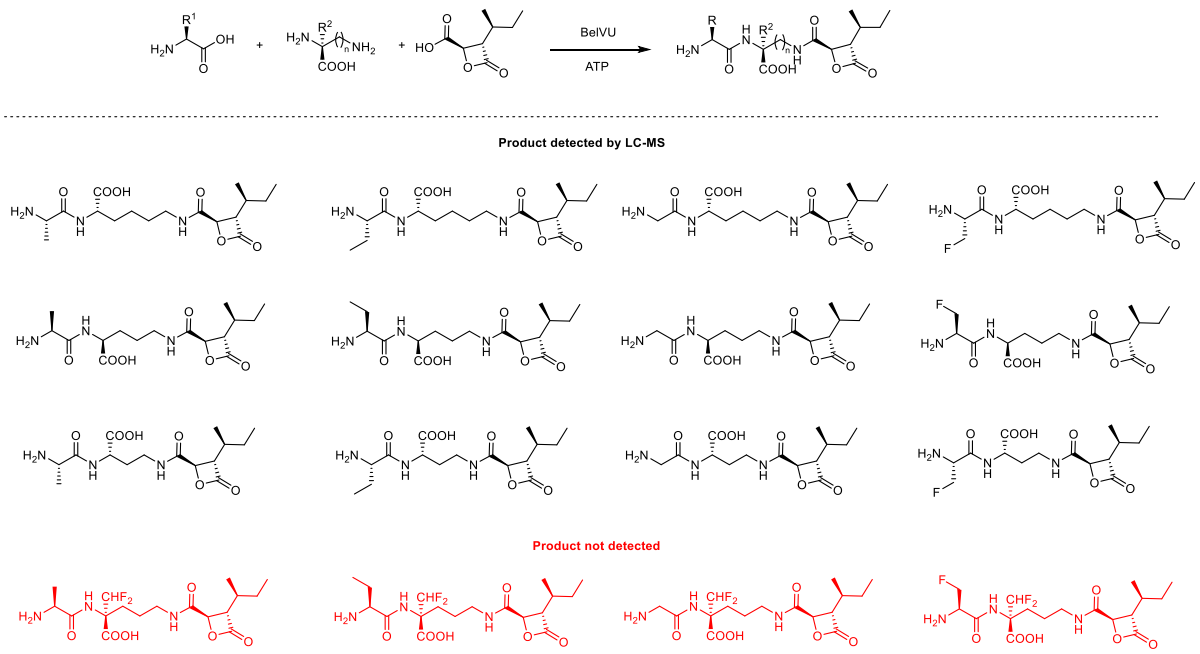
Accepted
(Product confirmed by LCMS) L-amino acids: Val, Ile, Leu, Met, Ala, Gly, Thr

Not accepted
(No product detected) L-amino acids: Phe, Tyr, Trp, Cys, Pro, Ser, Asn, Gln, Arg, His, Lys, Asp, Glu
D-amino acids: Val

Supplementary Fig. 18. Ligation of natural donor carboxylic acid **4a** or **5** with acceptor proteinogenic amino acids catalysed by CysC or CysD. Reactions were carried out on analytical scale (see method section 3 enzyme assays). Tandem mass (MS/MS) data of products are shown in section 9.



Supplementary Fig. 19. Amino acid substrates tested for BelV and BelU. The reaction was analysed by LC-MS. Black colour compounds have anticipated product mass detected, while the red colour indicated no product mass was detected. For assay conditions see methods section 3 (enzyme assays). Tandem mass (MS/MS) data of products are shown in section 9.



Supplementary Fig. 20. BelVU enzyme cascade for the synthesis of belactosin analogues. The reaction was analysed by LC-MS. Black colour compounds have anticipated product mass detected, while the red colour product mass was not detected. Detailed experimental conditions see 'enzyme assay' section. Tandem mass (MS/MS) data of products are shown in section 9.

Supplementary Table 3. Crystallographic data collection and refinement statistics. Statistics for the highest-resolution shell are shown in parentheses.

	CysF
Wavelength	0.9537
Resolution range	50.65 - 1.89 (1.94 - 1.89)
Space group	P 1 21 1
Unit cell	44.1 68.8 75.4 90 96.8 90
Total reflections	248200 (18265)
Unique reflections	35831 (2630)
Multiplicity	6.9 (6.9)
Completeness (%)	99.46 (95.43)
Mean I/sigma(I)	5.03 (0.60)
Wilson B-factor	26.08
R-merge	0.2418 (2.187)
R-meas	0.2616 (2.368)
R-pim	0.099 (0.89)
CC1/2	0.99 (0.38)
CC*	0.998 (0.74)
Reflections used in refinement	35745 (2612)
Reflections used for R-free	1763 (142)
R-work	0.1928 (0.3186)
R-free	0.2374 (0.3521)
Number of non-hydrogen atoms	4033
macromolecules	3752
ligands	26
solvent	255
Protein residues	485
RMS(bonds)	0.008
RMS(angles)	0.85
Ramachandran favored (%)	97.30
Ramachandran allowed (%)	2.7
Ramachandran outliers (%)	0.00
Rotamer outliers (%)	0.26
Clashscore	2.38
Average B-factor	33.15
macromolecules	33.13
ligands	39.93
solvent	32.75

2. Materials and general methods

General information

Chemicals were purchased from Sigma-Aldrich, Fluorochem, Acros Organics, Fisher Scientific UK, Bachem, Alfa Aesar or Apollo Scientific and used without further purification unless otherwise stated. Molecular biology enzymes were purchased from New England Biolabs unless otherwise stated. Codon optimised genes were synthesised by Twist Bioscience. Proteins were analysed by SDS-PAGE on precast gels (Invitrogen Novex™ WedgeWell™ 8-16% Tris-Glycine Gel). Low resolution LC-MS was performed on Agilent 1260 LC system fitted with Agilent 6130 Quadruple MS. High resolution Mass spectrometry (LC-HRMS) and tandem mass spectrometry (LC-MS/MS) was recorded on Agilent 6560 Q-TOF + Agilent 1290 Infinity LC system. GC-MS was performed on Agilent GC 7890B coupled with 5975 Series MSD. NMR spectra were recorded on Bruker machines.

Cloning

Synthetic genes for CysF, BelH, CysD and BelU were initially cloned into the pET28-a(+) plasmid (using *NdeI* and *XhoI* sites) with an N-terminal histidine tag sequence. However, N-terminal histidine tagged CysF produced insoluble protein. N-terminal histidine tagged CysD and BelU produced soluble expression but with no activity. CysF, CysD and BelU were therefore cloned into the pET21-a(+) plasmid (using *NdeI* and *XhoI* site) to insert a C-terminal histidine tag (Supplementary table 1-3). Synthetic genes for CysG, CysE, CysC, BelI, BelR, BelV and *Bh*CysFE were inserted into the pET28-a(+) plasmid (using *NcoI* and *XhoI* site) carrying C-terminal histidine tag sequences (Supplementary table 1-2).

Protein expression general method

LB media (10 mL) with 50 μg/mL of kanamycin (pET28-a(+)) or 100 μg/ml ampicillin (pET21-a(+)) was inoculated with *E. coli* BL21(DE3) containing a plasmid and incubated for 18 hrs at 37 °C, 180 rpm. The resulting seed culture (8 mL) was then used to inoculate 800 mL of autoinduction 2YT broth (with trace elements from Formedium) containing 50 μg/mL of kanamycin (pET28a-(+)) or 100 μg/mL of ampicillin (pET21-a(+)). After incubation for 4 hrs at 37 °C, with shaking (180 rpm), the temperature was reduced to 20 ºC and incubation was continued for a further 18 hrs (180 rpm). The cultures were then harvested by centrifugation (3000 × *g*, 4 °C, 10 min), cells were resuspended in 80 mL PBS, transferred into 2 × 50 mL falcon tubes, pelleted by centrifugation (3000 × *g*, 4 °C, 15 min) and then frozen until further usage. CysE (with C-terminal His-tag) produced mainly insoluble protein. A small amount of soluble CysE was, however, obtained (ca. 5 mg of protein from 1 L of cell culture) after the standard protein purification method described below. For the three fused bifunctional CysFE enzymes tested, only *Bh*CysFE was produced in soluble form. *Ms*CysFE afforded a very low level of expression, whilst *Mp*CysFE showed good expression, but formed inclusion bodies.

Protein purification general method

Frozen cell pellets were thawed and resuspended in 40 mL lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, 10% glycerol, pH 7.8) and lysed by sonication (10 min, 50% pulse, 60% amplitude). The lysate was then cleared *via* centrifugation (23224 × *g*, 40 min, 4 °C), 1 mL of pre-equilibrated Ni-NTA resin was added and the tube incubated at 4 °C for 30 min. The lysate with resin was loaded onto a gravity-flow column (Bio-Rad) and flowthrough was collected. The resin was washed with 10 mL of lysis buffer, followed by 2 × 10 mL of wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 40 mM imidazole, 10% glycerol, pH 7.8), and eluted with 10 mL of elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 100 mM imidazole, 10% glycerol, pH 7.8). The eluent from the column was monitored by Bio-Rad protein assay and elution continued until no further protein could be seen eluting. Sample of wash and elution fractions were analysed by SDS-PAGE. The wash and elution fractions that contained the protein of interested (with good purity) were combined and concentrated down to 2.5 mL using a centrifugal concentrator (Sartorius, Vivaspin 20 MWCO 30 kDa or 10 kDa depending on the size of the protein of interest). The 2.5 mL solution was applied to an equilibrated PD-10 desalting column (GE Healthcare, performed according to manufacturer's instructions). The column was eluted with 3.5 mL of protein storage buffer (20 mM NaH_2PO_4, 300 mM NaCl, 10% glycerol, pH 7.8). If necessary, the resulting eluent was further concentrated to approximately 10 mg/mL or more. Protein was aliquoted into single use tubes, flash frozen using liquid nitrogen and stored at -80 °C.

Further purification for protein crystallisation

For the protein crystallisation, the plasmid encoding CysF gene was transformed with NiCo21(DE3) competent *E. coli* cells. Following transformation, glycerol stocks were prepared from the overnight cultures of a single colony (sequencing confirmed). The enzyme was expressed and first purified using Ni-NTA column according to the general procedure described above. The wash/eluted protein fraction with good purity (judged by SDS-PAGE) were combined and concentrated using a Vivaspin centrifugal concentrator (MWCO 30 kDa) up to 20 mg/mL. For further purification, approximately 0.6 mL of this concentrated protein solution was injected into a 2 mL loop attached to the ÄKTA pure™ protein purification system. The protein was passed down a Superdex® 200 Increase 10/300 GL gel filtration column which was preequilibrated with gel filtration buffer (50 mM Tris-HCl, 300 mM NaCl, 10 mM $MgCl_2$, 10% glycerol, pH 7.5). CysF was eluted with gel filtration buffer and buffer exchanged to crystallisation buffer (50 mM Tris-HCl, 100 mM NaCl, 10 mM $MgCl_2$, 5% glycerol, pH 7.5) using a PD-10 desalting column (preequilibrated with crystallisation buffer). The resulting fraction was concentrated to 8-10 mg/mL using an Amicon Ultra 0.5 mL (30 kDa MWCO) concentrator.

Crystallogenesis

Single crystals of CysF were prepared by mixing 200 nL of 8 mg/mL protein in crystallisation buffer with equal volumes of precipitant. All trials were conducted by sitting drop vapour diffusion and incubated at 4°C. Crystals of the apo protein were formed in 0.1 M sodium citrate pH 5.5, 20% w/v PEG 3000 (JCSG + Eco A2, Molecular Dimensions). Individual crystals were cryoprotected in mother liquor supplemented with 25% PEG 200 prior to flash cooling in liquid nitrogen. Data were collected from single crystals at Diamond Light Source (MX31850-i04) and

subsequently scaled and reduced with Xia2. Preliminary phasing was performed by molecular replacement in Phaser using a search model generated in AlphaFold. Iterative cycles of rebuilding and refinement were performed in COOT and Phenix.refine, respectively. Structure validation with MolProbity and PDBREDO were integrated into the iterative rebuild and refinement process. The resolution cut of the data was determined by paired refinement as implemented in PDBREDO. Complete data collection and refinement statistics can be found in the Supplementary Table 3. Coordinates and structure factors have been deposited in the Protein Data Bank under accession code 8RA0.

Bioinformatic analysis of CysF and the discovery of fused CysFE

Bioinformatic analysis was conducted on the Cys and Bel pathway enzymes using the Enzyme Function Initiative tools.[S2] The analyses utilised databases UniProt 2022-04, InterPro 91, and corresponding ENA information. The identification of the fused CysFE configuration began with a BLAST query using Uniprot ID A0A1W6R555 for CysF to generate a Sequence Similarity Network (SSN) with default settings. An alignment score threshold of 106 was chosen to identify a sub-network containing both CysF and BelH as they are isofunctional in their enzymatic capabilities. The obtained SSN was further analysed with the Genome Neighbourhoods Tool, leading to the obtention of genome neighbourhood diagrams (Supplementary Fig. 13). Manual inspection of these diagrams revealed three genomes exhibiting a CysF and CysE fusion, indicating a bifunctional enzyme.

CysF closed conformation modelling

The crystal structure of CysF was used to search for homologues via Foldseek and Dali. Identified homologues were observed that displayed structures in both open and closed conformations. Crystal structures of 4-Coumarate CoA Ligase in both open and closed conformations (PDB-ID 5BSM open and PDB-ID 5BSR closed) were utilised as templates upon which to model a closed state of CysF. 4-Coumarate CoA Ligase structure revealed that domain closure is achieved through a hinge like rigid body reorganisation of the two domains. A closed model of CysF was therefore modelled by superposition of the N-terminal domain of CysF with the corresponding N-terminal domain of the template (RMSD of 1.86 Å after secondary structure alignment). Subsequently, the C-terminal domain of CysF was superimposed with the C-terminal domain of the template in the closed conformation (RMSD of 3.03 Å after secondary structure alignment) followed by an energy minimisation (Yasara).

CysF docking studies.

A model of **2a**-AMP was constructed and docked into the crystal structure of CysF. The pocket for docking was identified using ICM Pocket Finder as implemented in ICM-Pro.[S3] The top hit from the docking as ranked by RTCNN score (-29.78) is presented in Supplementary fig. 12. The docked pose was compared to structures of similar enzymes (5WM3, 5IE3, 4FUT, 4GXR, & 4GXQ). A multiple sequence alignment of these enzymes is presented in Supplementary fig. 13.

Chromatography

LC-MS analysis was performed using a Kinetex XB-C18 Core Shell column (100 mm x 4.6 mm, 5 µm, Phenomenex), flow = 1.0 mL/min, column oven temperature = 40 °C, mobile phase A = water with 0.1% (v/v) formic acid, mobile phase B = MeCN with 0.1% (v/v) formic acid, mobile phase gradient see table below.

Method 1:

Time	A [%]	B [%]
0:00	95.0	5.0
6:00	5.0	95.0
7:00	5.0	95.0
7:50	95.0	5.0
11:00	95.0	5.0

Method 2:

Time	A [%]	B [%]
0:00	95.0	5.0
15:00	5.0	95.0
16:00	5.0	95.0
16:50	95.0	5.0
20:00	95.0	5.0

GC-MS analysis was performed using a VF-5ht column (30 m x 0.25 mm x 0.1 µm, J&W), inlet temperature = 240 °C, split ratio = 100:1, inlet pressure = 7.6 psi, flow = 1 mL/min, carrier gas = helium, oven temperature programme see below.

	Rate [°C/min]	Value [°C]	Hold Time [min]
Initial		50	2
Ramp	30	350	3

3. Enzyme assays

Enzyme assays for testing proposed pathway.[32]

CysC was initially assayed for the lactonisation of **1a** to **4a** (Fig. 2a) in a reaction mixture (100 µL) containing 2 mM **1a**, 5 mM ATP, with or without 5 mM CoASH, 10 mM $MgCl_2$, and 20 µM CysC in 100 mM KPi buffer (pH 7.8) incubated at 25 °C for 12 h. The reaction was quenched by addition of 100 µL MeCN and the precipitated protein was removed by centrifugation. The supernatant was subjected to LC-MS analysis (method 1).

CysGE and BelIR were tested for methylation-lactonisation activity (**1a** to 4a, Fig, 2a) in reactions (100 µL) containing 2 mM **1a**, 5 mM SAM, 10 mM $MgCl_2$, 20 µM CysG/BelI and 20 µM CysE/BelR in 100 mM KPi buffer (pH 7.8) incubated at 25 °C for 12 h. The reactions were quenched and analysed by LC-MS analysis (as above).

The proposed amide-bond forming activities of CysF and BelH (Fig 4a and 4d) were tested in reactions (100 µL) containing 2 mM **4a** (CysF) or **4b** (BelH), 5 mM H-Val-Val-OH (CysF) or Ala-Orn (**12**, BelH), 5 mM ATP, 10 mM $MgCl_2$ and 20 µM CysF or BelH in 100 mM KPi buffer (pH 7.8, with 300 mM NaCl and 10% v/v glycerol) incubated at 25 °C for 12 h. The reactions were quenched and subjected to LC-MS analysis (as above).

The proposed amide-bond forming activities of CysD and BelU (Fig 4a and 4d) were also tested in reactions (100 µL) with 2 mM L-Val (for CysD) or 2 mM L-Orn and 2 mM L-Ala (for BelU), 5 mM ATP, 10 mM $MgCl_2$ and 20 µM CysD or BelU in 100 mM KPi buffer (pH 7.8) incubated at 25 °C for 12 h. The reactions were quenched and analysed (as above).

Enzyme assays for new pathways (Fig. 2 and 4)

CysG and BelI assay mixtures (100 µL) containing 2 mM **1a** or **1b**, 5 mM SAM, 10 mM $MgCl_2$, and 20 µM CysG or BelI in 100 mM KPi buffer (pH 7.8) were incubated at 25 °C for 12 h. The reactions were quenched by addition of 100 µL MeCN and the precipitated protein was removed by centrifugation. The supernatant was subjected to LC-MS analysis (method 1).

CysGF and BelIH cascade reactions (100 µL) containing 2 mM **1a** or **1b**, 5 mM SAM, 5 mM ATP, 10 mM $MgCl_2$, 20 µM CysG/BelI and 20 µM CysF/BelH in 100 mM HEPES or KPi buffer (pH 7.8) were incubated at 25 °C for 12 h. The reaction was quenched by addition of 100 µL MeCN and the precipitated protein was removed by centrifugation. The supernatant was subjected to LC-MS analysis (method 1). For GC-MS analysis, the reaction was scaled up to 200 µL, and reaction time was 5 h. The reaction was quenched by extracting with EtOAc (200 µL). The organic layer was dried over $MgSO_4$ and analysed by GC-MS.

CysFE and BelHR cascade reactions (100 µL) containing 2 mM **2a** or **2b**, 5 mM ATP, 10 mM $MgCl_2$, 20 µM CysF or BelH and 20 µM CysE or BelR in 100 mM KPi buffer (pH 7.8) were incubated at 25 °C for 12 h. The reaction was quenched by addition of 100 µL MeCN and the precipitated protein was removed by centrifugation. The supernatant was subjected to LC-MS analysis (method 1).

CysGFE and BelIHR cascade reactions (100 µL) containing 2 mM **1a** or **1b**, 5 mM SAM, 5 mM ATP, 10 mM $MgCl_2$, 20 µM CysG or BelI, 20 µM CysF or BelH and 20 µM CysE or BelR in 100 mM KPi buffer (pH 7.8) were incubated at 25 °C for 12 h. The reaction was quenched by addition of 100 µL MeCN and the precipitated protein was removed by centrifugation. The supernatant was subjected to LC-MS analysis (method 1).

CysC and BelV assays (100 µL) containing 2 mM **4a** or **4b**, 4 mM L-Val or L-Orn, 5 mM ATP, 10 mM $MgCl_2$, and 20 µM CysC or BelV in 100 mM KPi buffer (pH 7.8) were incubated at 25 °C for 2 h. The reaction was quenched and analysed (see directly above).

CysD and BelU assays (100 µL) containing 2 mM **5** or **7**, 4 mM L-Val or L-Ala, 5 mM ATP, 10 mM $MgCl_2$, and 20 µM CysD or BelU in 100 mM KPi buffer (pH 7.8) were incubated at 25 °C for 2 h. The reaction was quenched (as above) subjected to LC-MS analysis (method 1 for CysD and method 2 for BelU reactions). We noted that potassium ions serve as an activator for CysD catalysis and are essential for the activity of BelU.

CysCD and BelVU cascade reactions (100 µL) containing 2 mM **4a** or **4b**, 6 mM L-Val (CysC) or 4 mM L-Orn with 4 mM L-Ala (BelG), 10 mM ATP, 10 mM $MgCl_2$, 20 µM CysC or BelV, and 20 µM CysD or BelU in 100 mM KPi buffer (pH 7.8) at 25 °C for 12 h. The reaction was quenched (as above) and subjected to LC-MS analysis (method 1 for CysCD; method 2 for BelVU reactions).

Substrate scope analysis (Fig. 5)

CysGFECD enzyme cascades reactions (100 µL) containing 2 mM **1a**, 3 mM SAM, 9 mM ATP, 6 mM L-Val (or other amino acids), 10 mM $MgCl_2$, 20 µM CysG, 20 µM *Bh*CysFE, 20 µM CysC, and 20 µM CysD in 100 mM KPi buffer (pH 7.8) were incubated at 25 °C for 6 h. The reaction was quenched by addition of 100 µL MeCN and the precipitated protein was removed by centrifugation. The supernatant was subjected to LC-MS analysis (method 1). Reaction yields (average of three replicates) were determined from calibration curves using synthetic product standards.

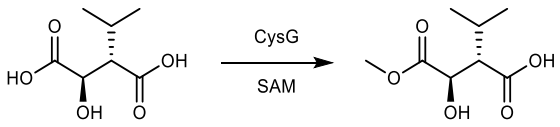
BelIHRVU enzyme cascade reactions (100 µL) containing 2 mM **1b**, 3 mM SAM, 9 mM ATP, 5 mM L-Orn (or other amino acids), 5 mM L-Ala (or other amino acids), 10 mM $MgCl_2$, 15 µM BelI, 15 µM BelH, 15 µM BelR, 15 µM BelV, and 15 µM BelU in 100 mM KPi buffer (pH 7.8) were incubated at 25 °C for 6 h. The reactions were the quenched and analysed (as above).

CysC and BelV reactions mixture (100 µL) containing 2 mM **4a** or **4b**, 3 mM amino acid, 3 mM ATP, 10 mM $MgCl_2$, and 20 µM CysC or BelV in 100 mM KPi buffer (pH 7.8) were incubated at 25 °C for 5 h. The reaction was quenched by addition of 100 µL MeCN and the precipitated protein was removed by centrifugation. The supernatant was diluted 10 times with MeCN/H_2O (50:50) and subjected to LC-MS analysis (method 1). Product yields for CysC reactions (average of three replicates) were determined from calibration curves using synthetic product standards. The yield of **25** was determined by ^{19}F-NMR.

CysD and BelU reaction mixture (100 µL) containing 2 mM **5** or **7**, 3 mM amino acid, 3 mM ATP, 10 mM $MgCl_2$, and 20 µM CysD or BelU in 100 mM KPi buffer (pH 7.8) were incubated at 25 °C for 5 h. The reaction was quenched as described above. The supernatant was diluted 10 times with MeCN/H_2O (50:50) and subjected to LC-MS analysis (method 1). Product yields for CysD reactions (average of three replicates) were determined using calibration curves. The yield of **36** was determined by ^{19}F-NMR.

4. Preparative scale enzymatic reactions

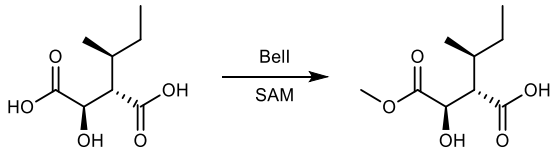
Preparative scale CysG catalysed methylation of **1a**



1a **2a**

A reaction mixture (20 mL) consisting of **1a** (35.2 mg, 10 mM), SAM (Sigma, ≥ 75% purity, 12 mM), $MgCl_2$ (10 mM), purified CysG (5 µM) in 100 mM KPi buffer (pH 7.5) was incubated at 30 °C with 200 rpm shaking for 18 h. A sample (50 µL) was taken for LC-MS analysis and indicated reaction completion before acidifying the reaction with 2 mL 1 M HCl to pH 2-3. The reaction was then extracted three times with EtOAc (20 mL, 10 mL, 10 mL). The organic layers were combined, washed with brine, dried over $MgSO_4$, and concentrated *in vacuo* to obtain 30.7 mg of **2a** as colourless oil in 81% yield.
^{1}H NMR (400 MHz, $CDCl_3$) δ 4.42 (d, *J* = 3.4 Hz, 1H), 3.79 (s, 3H), 2.62 (dd, *J* = 8.6, 3.4 Hz, 1H), 2.30 – 2.16 (m, 1H), 1.07 (dd, *J* = 11.1, 6.7 Hz, 6H).
^{13}C NMR (101 MHz, $CDCl_3$) δ 177.8, 174.3, 69.9, 55.0, 52.9, 27.4, 21.0, 20.3.
HR-MS, *m/z* (ESI+) calcd for $C_8H_{14}O_5$: 213.0739 $[M+Na]^+$, found 213.0732.
The NMR spectra agree with literature.[S4]

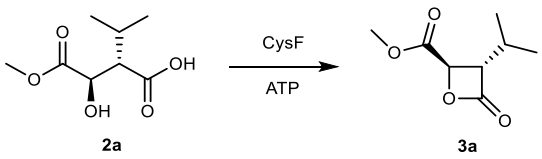
Preparative scale BelI catalysed methylation of **1b**



1b **2b**

A reaction mixture (10 mL) consisting of **1b** (19 mg, 10 mM), SAM (≥ 75% purity, treated as 75%, 12 mM), $MgCl_2$ (10 mM), purified BelI (5 µM) in 100 mM KPi buffer (pH 7.5) was incubated at 30 °C with 200 rpm shaking for 18 h. A sample (10 µL) was taken for LC-MS analysis and indicated reaction completion before acidifying the reaction with 2 mL 1 M HCl to pH 2-3. The reaction was then extracted with EtOAc three times (20 mL, 10 mL, 10 mL). The organic layers were combined, washed with brine, dried over $MgSO_4$ and concentrated *in vacuo* to obtain 15.6 mg **2b** as colourless oil in 76% yield.
^{1}H NMR (400 MHz, $CDCl_3$) δ 4.42 (d, *J* = 3.1 Hz, 1H), 3.80 (s, 3H), 2.77 (dd, *J* = 8.3, 3.0 Hz, 1H), 2.13 – 1.98 (m, 1H), 1.59 – 1.46 (m, 1H), 1.33 – 1.23 (m, 1H), 1.06 (d, *J* = 6.7 Hz, 3H), 0.94 (t, *J* = 7.4 Hz, 3H).
^{13}C NMR (101 MHz, $CDCl_3$) δ 178.3, 174.4, 69.5, 53.2, 53.0, 33.7, 27.6, 16.4, 11.4.
HR-MS, *m/z* (ESI+) calcd for $C_9H_{16}O_5$: 227.0890 $[M+Na]^+$, found 227.0889.

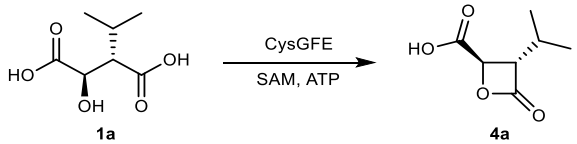
Preparative scale CysF catalysed lactonisation of **2a**



2a **3a**

A reaction mixture (15 mL) consisting of **2a** (28.5 mg, 10 mM), ATP (50 mM), $MgCl_2$ (10 mM), purified CysF (10 µM) in 100 mM KPi buffer (pH 7.5) was incubated at 30 °C with 200 rpm shaking for 18 h before extracting three times with Et_2O (3 x 15 mL). The organic layers were combined, washed with brine, dried over $MgSO_4$, and concentrated *in vacuo* to obtain 5 mg crude product. The crude product was subjected to ^{1}H-NMR analysis (see supplementary Fig. 8).

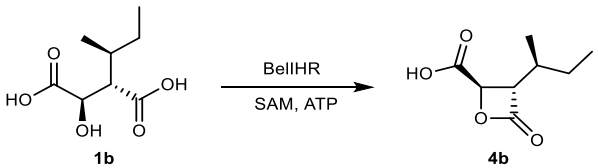
Preparative scale CysGFE enzymatic cascade for the synthesis of **4a**



1a **4a**

A reaction mixture (25 mL) consisting of **1a** (22.0 mg, 5 mM), SAM (≥ 75% purity, 6 mM), ATP (10 mM), $MgCl_2$ (10 mM), purified CysG (20 µM) and purified *Bh*CysFE (20 µM) in 100 mM KPi buffer (containing 10% glycerol v/v, pH 7.5) was incubated at 25 °C with 200 rpm shaking for 24 h. A sample (50 µL) was taken for LC-MS analysis and indicated reaction completion before acidifying the reaction with 5 mL 1 M HCl to pH 2-3. The reaction was then extracted with EtOAc (3 x 15 mL). The organic layers were combined, washed with brine (15 mL), dried over $MgSO_4$, and concentrated *in vacuo* to obtain 18.5 mg **4a** as light-yellow oil in 94% yield.
^{1}H NMR (400 MHz, $CDCl_3$) δ 4.69 (d, *J* = 4.4 Hz, 1H), 3.62 (dd, *J* = 8.4, 4.4 Hz, 1H), 2.29 – 2.17 (m, 1H), 1.14 (d, *J* = 6.7 Hz, 3H), 1.10 (d, *J* = 6.7 Hz, 3H).
^{13}C NMR (101 MHz, $CDCl_3$) δ 173.4, 168.2, 69.3, 64.8, 28.0, 20.1, 19.7.
The NMR spectra agree with chemically synthesised **4a**.
HR-MS, *m/z* (ESI-) calcd for $C_7H_{10}O_4$: 157.0506 $[M-H]^-$, found 157.0508.

Preparative scale BelIHR enzymatic cascade for the synthesis of **4b**

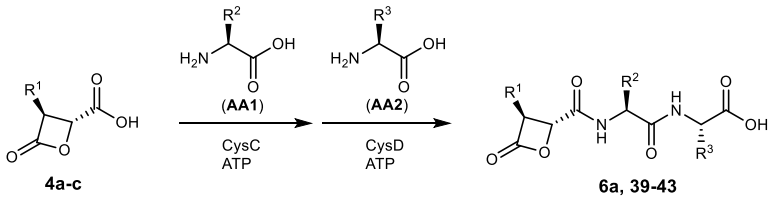


1b **4b**

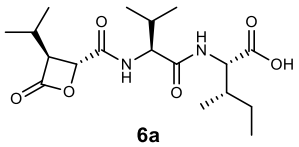
A reaction mixture (25 mL) consisting of **1b** (23.7 mg, 5 mM), SAM (≥ 75% purity, 6 mM), ATP (10 mM), $MgCl_2$ (10 mM), purified BelI (20 µM), BelH (20 µM), and BelR (20 µM) in 100 mM KPi buffer (pH 7.5) was incubated at 25 °C with 200 rpm shaking for 24 h. A sample (50 µL) was taken for LC-MS analysis and indicated reaction completion before acidifying the reaction with 5 mL 1 M HCl to pH 2-3. The reaction was then extracted with EtOAc (3 x 15 mL). The organic layers were combined, washed with brine (15 mL), dried over $MgSO_4$, and concentrated *in vacuo* to obtain 18.7 mg **4b** as light-yellow oil in 87% yield.

^{1}H NMR (400 MHz, $CDCl_3$) δ 4.71 (d, *J* = 4.5 Hz, 1H), 3.76 (dd, *J* = 7.9, 4.5 Hz, 1H), 2.08 – 1.98 (m, 1H), 1.72 – 1.59 (m, 1H), 1.39 – 1.29 (m, 1H), 1.07 (d, *J* = 6.7 Hz, 3H), 0.95 (t, *J* = 7.5 Hz, 3H).
^{13}C NMR (101 MHz, $CDCl_3$) δ 173.75, 168.51, 68.79, 63.33, 33.83, 26.85, 16.37, 11.10.
The NMR spectra agree with chemically synthesised **4b**.
HR-MS, *m/z* (ESI-) calcd for $C_8H_{12}O_4$: 171.0663 [M-H]$^-$, found 171.0667.

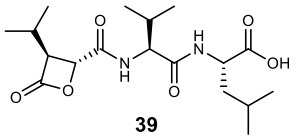
Preparative scale CysCD enzymatic cascade (stepwise)



General method: A reaction mixture (20 mL) consisting of **4a-c** (4 mM, 1.0 equiv, 0.08 mmol), **AA1** (4 mM, 1.0 equiv), ATP (4.4 mM, 1.1 equiv), $MgCl_2$ (10 mM), 100 mM NaCl, and purified CysC (20 μM) in 100 mM KPi buffer (pH 7.8) was incubated at 25 °C with 200 rpm shaking. After 12 h of incubation, additional ATP (1.1 equiv, 0.88 mL of 100 mM stock solution in KPi buffer), **AA2** (1.25 equiv, 1.0 mL of 100 mM stock solution in KPi buffer), and purified CysD (20 μM) were then added. The reaction mixture was incubated at 25 °C with 200 rpm shaking for another 12 h. The reaction mixture was then cooled to 0 °C on ice before acidifying with 1.0 M HCl (1.8 mL) to pH 5-6. The resulting mixture was extracted with ice-cold EtOAc (3 x 40 mL). During extraction, the mixture was kept ice-cold to minimise acid catalysed β-lactone ring opening in water. The organic layers were combined, washed with brine (20 mL), dried over $MgSO_4$, and concentrated *in vacuo* to obtain crude product. The crude product was purified by silica gel column (DCM/MeOH with 0.1% AcOH, 1-4% MeOH) to obtain purified warhead-peptides.

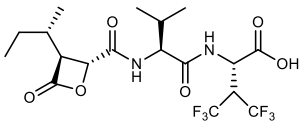


6a was obtained as white solid, 16 mg, 54% yield over two steps.
^{1}H NMR (500 MHz, MeOD) δ 4.81 (d, *J* = 4.3 Hz, 1H), 4.37 (d, *J* = 5.8 Hz, 1H), 4.30 (d, *J* = 7.9 Hz, 1H), 3.56 (dd, *J* = 8.5, 4.4 Hz, 1H), 2.23 – 2.15 (m, 1H), 2.14 – 2.07 (m, 1H), 1.95 – 1.86 (m, 1H), 1.58 – 1.48 (m, 1H), 1.29 – 1.21 (m, 1H), 1.10 (d, *J* = 6.8 Hz, 3H), 1.06 (d, *J* = 6.7 Hz, 3H), 1.00 – 0.91 (m, 12H).
^{13}C NMR (126 MHz, MeOD) δ 174.5, 173.3, 170.6, 170.4, 72.2, 65.2, 60.2, 58.22, 38.3, 32.0, 28.8, 26.2, 20.4, 19.8, 19.7, 18.8, 16.0, 11.8.
HR-MS, *m/z* (ESI-) calcd for $C_{18}H_{30}N_2O_6$: 369.2031 [M-H]$^-$, found 369.2054.



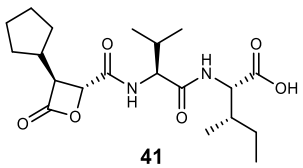
39
Enzymatic

39 was obtained as white solid, 8 mg, 27% yield over two steps.
^{1}H NMR (500 MHz, MeOD) δ 4.81 (d, *J* = 4.4 Hz, 1H), 4.45 (t, *J* = 7.5 Hz, 1H), 4.26 (d, *J* = 7.8 Hz, 1H), 3.55 (dd, *J* = 8.5, 4.4 Hz, 1H), 2.22 – 2.08 (m, 2H), 1.72 (d, *J* = 6.7 Hz, 1H), 1.65 (t, *J* = 7.3 Hz, 2H), 1.10 (d, *J* = 6.7 Hz, 3H), 1.06 (d, *J* = 6.7 Hz, 3H), 1.00 (d, *J* = 6.7 Hz, 3H), 0.97 (d, *J* = 1.4 Hz, 3H), 0.95 (d, *J* = 1.6 Hz, 3H), 0.91 (d, *J* = 6.4 Hz, 3H).
^{13}C NMR (126 MHz, MeOD) δ 175.7, 173.2, 170.6, 170.4, 72.2, 65.3, 60.2, 52.0, 41.6, 32.0, 28.8, 25.9, 23.4, 21.8, 20.4, 19.8, 19.7, 18.8.
HR-MS, *m/z* (ESI-) calcd for $C_{18}H_{30}N_2O_6$: 369.2031 [M-H]$^-$, found 369.2096.



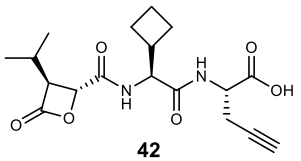
40
Enzymatic

40 was obtained as a light yellow solid, 24 mg, 63% yield over two steps. Unlike other examples, the racemic amino acid (**AA2**, 2.5 equiv) was used in this case. We observed the formation of only one diastereomer, indicating a complete kinetic resolution by the enzyme.
^{1}H NMR (500 MHz, MeOD) δ 5.41 (s, 1H), 4.82 (d, *J* = 4.4 Hz, 1H), 4.38 (d, *J* = 8.3 Hz, 1H), 4.36 – 4.27 (m, 1H), 3.66 (dd, *J* = 8.0, 4.4 Hz, 1H), 2.20 – 2.12 (m, 1H), 2.02 – 1.93 (m, 1H), 1.68 – 1.59 (m, 1H), 1.36 – 1.30 (m, 1H), 1.04 (d, *J* = 6.7 Hz, 3H), 0.99 (d, *J* = 6.7 Hz, 3H), 0.97 – 0.91 (m, 6H).
^{13}C NMR (126 MHz, MeOD) δ 173.5, 170.8, 170.5, 71.6, 63.8, 60.2, 34.8, 31.6, 27.7, 19.6, 18.6, 16.6, 11.3.
^{19}F NMR (471 MHz, MeOD) δ -63.6 (q, *J* = 9.3 Hz), -67.1 (q, *J* = 8.9 Hz).
HR-MS, *m/z* (ESI-) calcd for $C_{18}H_{24}F_6N_2O_6$: 477.1460 [M-H]$^-$, found 477.1453.



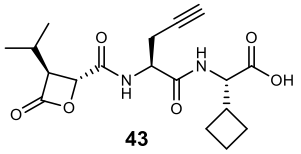
41
Enzymatic

41 was obtained as a light yellow solid, 10 mg, 32% yield over two steps.
^{1}H NMR (500 MHz, MeOD) δ 4.76 (d, *J* = 4.3 Hz, 1H), 4.37 (d, *J* = 5.8 Hz, 1H), 4.29 (d, *J* = 7.9 Hz, 1H), 3.70 (dd, *J* = 8.8, 4.3 Hz, 1H), 2.42 – 2.33 (m, 1H), 2.15 – 2.06 (m, 1H), 1.97 – 1.83 (m, 3H), 1.75 – 1.59 (m, 4H), 1.57 – 1.50 (m, 1H), 1.46 – 1.36 (m, 2H), 1.27 – 1.23 (m, 1H), 1.01 – 0.91 (m, 12H).
^{13}C NMR (126 MHz, MeOD) δ 174.6, 173.2, 170.9, 170.5, 72.9, 62.9, 60.2, 58.3, 39.5, 38.3, 31.9, 31.2, 30.8, 26.2, 26.0, 25.7, 19.7, 18.9, 16.0, 11.8.
HR-MS, *m/z* (ESI-) calcd for $C_{20}H_{32}N_2O_6$: 395.2182 [M-H]$^-$, found 395.2147.



42
Enzymatic

42 was obtained as a white solid, 10 mg, 34% yield over two steps.
^{1}H NMR (500 MHz, MeOD) δ 4.79 (d, *J* = 4.4 Hz, 1H), 4.48 – 4.43 (m, 2H), 3.62 (dd, *J* = 8.4, 4.4 Hz, 1H), 2.79 – 2.67 (m, 3H), 2.32 (t, *J* = 2.7 Hz, 1H), 2.22 – 2.13 (m, 1H), 2.08 – 1.97 (m, 3H), 1.93 – 1.81 (m, 3H), 1.10 (d, *J* = 6.7 Hz, 3H), 1.06 (d, *J* = 6.7 Hz, 3H).
^{13}C NMR (126 MHz, MeOD) δ 172.4, 170.7, 170.6, 80.3, 72.2, 72.1, 65.4, 59.0, 53.0, 38.6, 28.9, 26.2, 26.2, 22.5, 20.4, 19.8, 18.8.
HR-MS, *m/z* (ESI-) calcd for $C_{18}H_{24}N_2O_6$: 363.1556 [M-H]$^-$, found 363.1570.

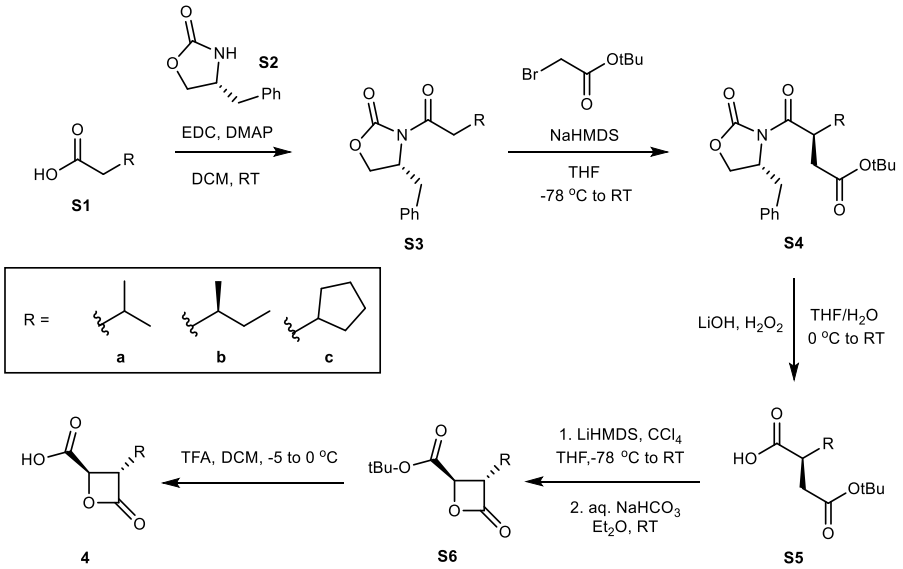


43
Enzymatic

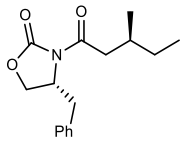
43 was obtained as a white solid, 4 mg, 14% yield over two steps.
^{1}H NMR (500 MHz, MeOD) δ 4.79 (d, *J* = 4.3 Hz, 1H), 4.64 (dd, *J* = 8.7, 5.2 Hz, 1H), 4.34 (d, *J* = 8.4 Hz, 1H), 3.61 (dd, *J* = 8.4, 4.4 Hz, 1H), 2.80 – 2.63 (m, 3H), 2.39 (t, *J* = 2.7 Hz, 1H), 2.22 – 2.14 (m, 1H), 2.07 – 1.95 (m, 3H), 1.93 – 1.79 (m, 3H), 1.11 (d, *J* = 6.7 Hz, 3H), 1.07 (d, *J* = 6.7 Hz, 3H).
^{13}C NMR (126 MHz, MeOD) δ 172.0, 170.7, 170.6, 80.1, 72.4, 72.2, 65.4, 57.9, 53.4, 38.5, 28.9, 26.3, 25.9, 22.6, 20.5, 19.8, 18.7.
HR-MS, *m/z* (ESI-) calcd for $C_{18}H_{24}N_2O_6$: 363.1556 [M-H]$^-$, found 363.1508.

5. Synthesis of substrates and standards

Synthesis of β-lactones **4a-c** (adapted from literature with modifications)[37, S5]

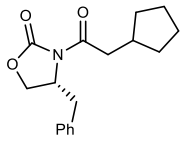


*General procedure synthesis of **S3** (Step 1):* To a solution of Evans' chiral auxiliary **S2** (751 mg, 4.2 mmol, 1.0 equiv) in 10 mL DCM, at room temperature, was added EDC hydrochloride (1.6 g, 8.4 mmol, 2.0 equiv), carboxylic acid **S1** (5.0 mmol, 1.2 equiv), and DMAP (513 mg, 4.2 mmol, 1.0 equiv). The mixture was stirred at RT for 16-20 h. The reaction was diluted with DCM and washed with 1 M HCl, followed by sat. $NaHCO_3$, sat. NH_4Cl, and finally brine. The organic layer was dried over $MgSO_4$ and concentrated *in vacuo*. The crude product was purified by silica gel column (EtOAc/Hexane, 20-30% EtOAc) to obtain **S3b** and **S3c**. A sample of **S3a** was also purchased from Fluorochem (UK) and used for the next step directly.



S3b

S3b was obtained as a white solid, 4.36 g, 89% yield.
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.41 – 7.21 (m, 5H), 4.79 – 4.65 (m, 1H), 4.28 – 4.15 (m, 2H), 3.35 (dd, *J* = 13.3, 3.4 Hz, 1H), 2.99 – 2.71 (m, 3H), 2.12 – 1.96 (m, 1H), 1.55 – 1.41 (m, 1H), 1.38 – 1.22 (m, 1H), 1.03 (d, *J* = 6.7 Hz, 3H), 0.96 (t, *J* = 7.4 Hz, 3H).
^{13}C NMR (101 MHz, $CDCl_3$) δ 173.1, 153.6, 135.5, 129.6, 129.1, 127.5, 66.2, 55.4, 42.3, 38.1, 31.4, 29.5, 19.4, 11.5.
The NMR spectra is in agreement with published data.[S6]
HR-MS, *m/z* (ESI+) calcd for $C_{16}H_{21}NO_3$: 298.1414 $[M+Na]^+$, found 298.1410.

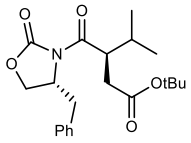


S3c

S3c was obtained as a white solid, 960 mg, 79% yield,
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.42 – 7.20 (m, 5H), 4.76 – 4.63 (m, 1H), 4.28 – 4.12 (m, 2H), 3.33 (dd, *J* = 13.4, 3.3 Hz, 1H), 3.05 (dd, *J* = 16.6, 6.9 Hz, 1H), 2.93 (dd, *J* = 16.6, 7.4 Hz, 1H), 2.78 (dd, *J* = 13.4, 9.6 Hz, 1H), 2.44 – 2.28 (m, 1H), 1.99 – 1.84 (m, 2H), 1.74 – 1.53 (m, 4H), 1.33 – 1.15 (m, 2H).
^{13}C NMR (101 MHz, $CDCl_3$) δ 173.2, 153.6, 135.5, 129.6, 129.1, 127.4, 66.2, 55.3, 41.5, 38.1, 35.9, 32.7, 32.6, 25.1, 25.1.
The NMR spectra is in agreement with published data.[29]
HR-MS, *m/z* (ESI+) calcd for $C_{17}H_{21}NO_3$: 310.1419 $[M+Na]^+$, found 310.1422.

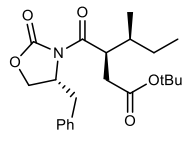
*General procedure for the synthesis of **S4**:* To a solution of **S3** (10.0 mmol, 1.0 equiv) in 50 mL of dry THF (cooled to -78 °C under N_2 atmosphere) was added NaHMDS (7.5 mL of 2 M solution in THF, 30 mmol, 1.5 equiv) dropwise and the solution was kept at -78 °C for 1 h before *tert*-butyl bromoacetate (3.9 g, 20 mmol, 2.0 equiv) was added dropwise. The resulting solution was stirred for 16 h and gradually reached room temperature before quenching with 1.3 mL of acetic acid. The mixture was concentrated *in vacuo* and the residue was added to 100 mL ethyl acetate, followed by washing with 100 mL water and 50 mL brine. The organic layer was dried

over $MgSO_4$ and concentrated *in vacuo* to give crude product, which was either recrystallised in EtOAc/hexane (1/20) or purified by silica gel column to give **S4**.



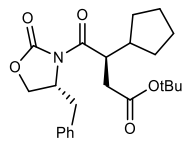
S4a

S4a was obtained as light-yellow needle shape crystal, 3.0 g, 80% yield.
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.37 – 7.26 (m, 5H), 4.72 – 4.60 (m, 1H), 4.20 – 4.08 (m, 3H), 3.35 (dd, *J* = 13.5, 3.3 Hz, 1H), 2.88 – 2.67 (m, 2H), 2.45 (dd, *J* = 16.9, 3.6 Hz, 1H), 2.06 – 1.92 (m, 1H), 1.42 (s, 9H), 1.01 (d, *J* = 6.8 Hz, 3H), 0.91 (d, *J* = 6.9 Hz, 3H).
^{13}C NMR (101 MHz, $CDCl_3$) δ 175.7, 172.0, 136.1, 129.7, 129.1, 127.3, 80.8, 65.9, 55.9, 44.6, 37.6, 33.7, 29.9, 28.2, 20.9, 18.5.
The NMR spectra is in agreement with published data.[S7]
HR-MS, *m/z* (ESI+) calcd for $C_{21}H_{29}NO_5$: 398.1943 $[M+Na]^+$, found 398.1962.



S4b

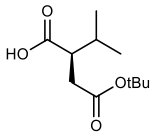
S4b was obtained as light-yellow needle shape crystal, 4.0 g, 64% yield.
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.36 – 7.19 (m, 5H), 4.68 – 4.54 (m, 1H), 4.24 – 4.15 (m, 1H), 4.15 – 4.09 (m, 2H), 3.32 (dd, *J* = 13.5, 3.3 Hz, 1H), 2.88 – 2.65 (m, 2H), 2.34 (dd, *J* = 16.9, 3.5 Hz, 1H), 1.82 – 1.68 (m, 1H), 1.50 – 1.31 (m, 10H), 1.31 – 1.17 (m, 1H), 0.91 (t, *J* = 7.4 Hz, 3H), 0.83 (d, *J* = 6.8 Hz, 3H).
^{13}C NMR (101 MHz, $CDCl_3$) δ 175.7, 172.1, 153.1, 136.0, 129.7, 129.0, 127.3, 80.8, 65.9, 55.9, 43.6, 37.6, 35.8, 32.4, 28.2, 28.0, 14.9, 12.0.
HR-MS, *m/z* (ESI+) calcd for $C_{22}H_{31}NO_5$: 412.2100 $[M+Na]^+$, found 412.2110.



S4c

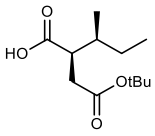
S4c was obtained as white needle shape crystal, 1.0 g, 79% yield.
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.35 – 7.21 (m, 5H), 4.73 – 4.54 (m, 1H), 4.27 – 4.17 (m, 1H), 4.14 – 4.08 (m, 2H), 3.36 (dd, *J* = 13.5, 3.3 Hz, 1H), 2.82 (dd, *J* = 16.8, 11.3 Hz, 1H), 2.69 (dd, *J* = 13.5, 10.3 Hz, 1H), 2.50 (dd, *J* = 16.8, 3.7 Hz, 1H), 2.06 – 1.89 (m, 1H), 1.80 – 1.45 (m, 6H), 1.39 (s, 9H), 1.31 – 1.16 (m, 2H).
^{13}C NMR (101 MHz, $CDCl_3$) δ 176.4, 171.6, 153.3, 136.2, 129.6, 129.0, 127.3, 80.8, 65.9, 56.0, 42.9, 42.7, 37.5, 37.0, 30.2, 30.0, 28.2, 25.0, 24.9.
HR-MS, *m/z* (ESI+) calcd for $C_{23}H_{31}NO_5$: 424.2100 $[M+Na]^+$, found 424.2102.
The NMR spectra is in agreement with published data.[29]

*General Procedure for the synthesis of **S5**:* A solution of **S4a** (3.0 g, 8.0 mmol, 1.0 equiv) in 40 mL of THF and 7 mL of water was cooled to 0 °C and treated with 5.6 mL of a 30% wt. H_2O_2 solution. The mixture was stirred for 5 min, followed by addition of LiOH (monohydrate, 0.85 g, 20 mmol, 2.5 equiv, freshly prepared in 13 mL of water) dropwise. The reaction mixture was stirred and allowed to gradually reach room temperature for 16 h before cooling back to 0 °C and quenching with 9 mL of a 2 M $Na_2S_2O_3$ solution. The reaction mixture was concentrated *in vacuo* to remove THF and the resulting solution was diluted with 20 mL water. The pH was adjusted to 8-9 using 2 M NaOH solution (if necessary). The aqueous solution was washed with EtOAc to remove the Evans' chiral auxiliary. The aqueous layer was then acidified with 2 M HCl solution (to pH 1-2) and extracted with EtOAc three times. The organic layers were combined, washed with brine, dried over $MgSO_4$, and concentrated *in vacuo*. The crude product was purified by silica gel column (EtOAc/Hexane, 20% EtOAc) to give 1.5 g of **S5a** in 87% yield.



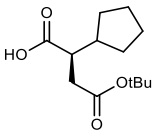
S5a

S5a was obtained as a colourless oil, 1.5 g, 87% yield.
^{1}H NMR (400 MHz, $CDCl_3$) δ 2.75 – 2.57 (m, 2H), 2.36 (dd, *J* = 16.2, 3.7 Hz, 1H), 2.10 – 1.96 (m, 1H), 1.43 (s, 8H), 0.97 (dd, *J* = 11.5, 6.8 Hz, 6H).
^{13}C NMR (101 MHz, $CDCl_3$) δ 171.7, 81.1, 47.6, 34.2, 30.0, 28.1, 20.3, 19.6.
HR-MS, *m/z* (ESI+) calcd for $C_{11}H_{20}O_4$: 239.1259 $[M+Na]^+$, found 239.1266.
The NMR spectra is in agreement with published data.[37]



S5b

S5b was obtained as a colourless oil, 1.2 g, 62% yield.
^{1}H NMR (400 MHz, $CDCl_3$) δ 2.92 – 2.80 (m, 1H), 2.58 (dd, *J* = 16.6, 11.0 Hz, 1H), 2.26 (dd, *J* = 16.6, 3.8 Hz, 1H), 1.92 – 1.75 (m, 1H), 1.50 – 1.32 (m, 10H), 1.30 – 1.16 (m, 1H), 0.96 – 0.81 (m, 6H).
^{13}C NMR (101 MHz, $CDCl_3$) δ 181.3, 171.8, 81.0, 45.8, 36.3, 32.8, 28.1, 27.3, 15.9, 11.9.
HR-MS, *m/z* (ESI+) calcd for $C_{12}H_{22}O_4$: 253.1416 $[M+Na]^+$, found 253.1410.
The NMR spectra is in agreement with published data.[S8]

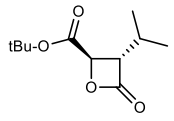


S5c

S5c was obtained as colourless oil, 480 mg, 80% yield.

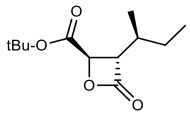
^{1}H NMR (400 MHz, $CDCl_3$) δ 2.69 – 2.57 (m, 2H), 2.49 – 2.37 (m, 1H), 2.02 – 1.90 (m, 1H), 1.84 – 1.70 (m, 2H), 1.68 – 1.59 (m, 2H), 1.58 – 1.49 (m, 2H), 1.42 (s, 9H), 1.38 – 1.29 (m, 1H), 1.27 – 1.14 (m, 1H).
^{13}C NMR (101 MHz, $CDCl_3$) δ 181.1, 181.1, 171.4, 81.2, 46.5, 42.2, 36.8, 30.6, 30.6, 28.1, 25.1, 25.0.
HR-MS, *m/z* (ESI-) calcd for $C_{13}H_{22}O_4$: 241.1445 $[M-H]^-$, found 241.1448.
The NMR spectra is in agreement with published data.[29]

*General Procedure for synthesis of **S6:*** A solution of **S5a** (1.5 g, 6.9 mmol, 1 equiv) in 20 mL of dry THF was cooled to -78 °C and treated with 14 mL of 1 M LiHMDS (14 mmol, 2 equiv) dropwise. The solution was kept at -78 °C for 1 h before dry CCl_4 (1.3 g, 8.3 mmol, 1.2 equiv) was added dropwise. The reaction mixture was stirred and allowed to gradually reach room temperature for 12 h before concentrating *in vacuo*. The resulting gummy solid residue was suspended in 100 mL of Et_2O and 50 mL of 5% (w/v) $NaHCO_3$ solution. The mixture was vigorously stirred at room temperature for 24 h and then diluted with Et_2O, washed with sat. $NaHCO_3$ solution and brine, dried over $MgSO_4$ and concentrated *in vacuo*. The crude product was purified by silica gel column (EtOAc/Hexane, 2%-10% EtOAc) to give **S6a**.



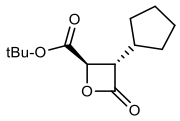
S6a

S6a was obtained as a white solid, 1.0 g, 67% yield.
^{1}H NMR (400 MHz, $CDCl_3$) δ 4.52 (d, *J* = 4.4 Hz, 1H), 3.45 (dd, *J* = 8.4, 4.4 Hz, 1H), 2.27 – 2.08 (m, 1H), 1.51 (s, 9H), 1.12 (d, *J* = 6.7 Hz, 3H), 1.07 (d, *J* = 6.7 Hz, 3H).
^{13}C NMR (101 MHz, $CDCl_3$) δ 168.9, 167.6, 83.7, 70.3, 64.2, 28.1, 27.8, 20.1, 19.8.
HR-MS, *m/z* (ESI+) calcd for $C_{11}H_{18}O_4$: 237.1103 $[M+Na]^+$, found 237.1105.
The NMR spectra is in agreement with published data.[37]



S6b

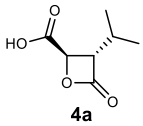
S6b was obtained as a white solid, 760 mg, 64% yield.
^{1}H NMR (400 MHz, $CDCl_3$) δ 4.53 (d, *J* = 4.4 Hz, 1H), 3.60 (dd, *J* = 7.7, 4.5 Hz, 1H), 2.05 – 1.90 (m, 1H), 1.68 – 1.56 (m, 1H), 1.50 (s, 9H), 1.39 – 1.26 (m, 1H), 1.03 (d, *J* = 6.7 Hz, 3H), 0.93 (t, *J* = 7.5 Hz, 3H).
^{13}C NMR (101 MHz, $CDCl_3$) δ 169.2, 167.6, 83.6, 69.6, 62.5, 33.6, 28.0, 26.9, 16.4, 11.2.
The NMR spectra is in agreement with published data.[S5]
HR-MS, *m/z* (ESI+) calcd for $C_{12}H_{20}O_4$: 251.1259 $[M+Na]^+$, found 251.1254.



S6c

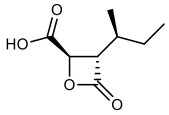
S6c was obtained as a light-yellow solid, 172 mg, 36% yield.
^{1}H NMR (400 MHz, $CDCl_3$) δ 4.49 (d, *J* = 4.3 Hz, 1H), 3.60 (dd, *J* = 8.6, 4.3 Hz, 1H), 2.41 – 2.27 (m, 1H), 1.99 – 1.80 (m, 2H), 1.74 – 1.58 (m, 4H), 1.50 (s, 9H), 1.45 – 1.32 (m, 2H).
^{13}C NMR (101 MHz, $CDCl_3$) δ 169.2, 167.6, 83.6, 71.0, 61.7, 38.3, 30.2, 29.9, 28.1, 25.1, 24.9.
The NMR spectra is in agreement with published data.[29]
HR-MS, *m/z* (ESI+) calcd for $C_{13}H_{20}O_4$: 263.1259 $[M+Na]^+$, found 263.1260.

*General Procedure for the synthesis of **4:*** A solution of **S6a** (262 mg, 1.2 mmol) in 1 mL of DCM was cooled to 0 °C and treated with 1 mL of TFA dropwise. The mixture was stirred at -5 °C to 0 °C for 12 h before removal of the DCM and TFA at 0 °C under high vacuum to give **4a**. (Note that removal of DCM and TFA at higher temperature can lead to partial epimerisation of product). The crude product was purified (in the case of **4c**) by silica gel column (DCM/MeOH with 0.1% acetic acid, 0-5% MeOH).



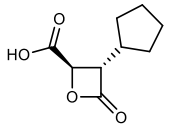
4a

4a was obtained as a viscous oil, 190 mg, in quantitative yield.
^{1}H NMR (400 MHz, $CDCl_3$) δ 4.69 (d, *J* = 4.4 Hz, 1H), 3.63 (dd, *J* = 8.4, 4.4 Hz, 1H), 2.30 – 2.17 (m, 1H), 1.14 (d, *J* = 6.7 Hz, 3H), 1.10 (d, *J* = 6.7 Hz, 3H).
^{13}C NMR (101 MHz, $CDCl_3$) δ 173.7, 168.1, 69.2, 64.8, 28.0, 20.1, 19.7.
HR-MS, *m/z* (ESI-) calcd for $C_7H_{10}O_4$: 157.0506 $[M-H]^-$, found 157.0507.
The NMR spectra is in agreement with published data.[37]



4b

4b was obtained as a viscous oil, 264 mg, in quantitative yield.
^{1}H NMR (400 MHz, $CDCl_3$) δ 4.71 (d, *J* = 4.5 Hz, 1H), 3.76 (dd, *J* = 7.9, 4.5 Hz, 1H), 2.11 – 1.95 (m, 1H), 1.71 – 1.59 (m, 1H), 1.40 – 1.25 (m, 1H), 1.07 (d, *J* = 6.7 Hz, 3H), 0.95 (t, *J* = 7.5 Hz, 3H).
^{13}C NMR (101 MHz, $CDCl_3$) δ 173.9, 168.4, 68.8, 63.4, 33.8, 26.9, 16.4, 11.1.
The NMR data is in agreement with published data.[S9]
HR-MS, *m/z* (ESI-) calcd for $C_8H_{12}O_4$: 171.0663 $[M-H]^-$, found 171.0668.

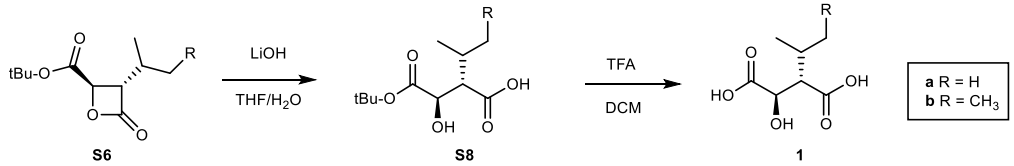


4c

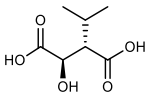
4c was obtained as a viscous light-yellow oil, 45 mg, 60% yield.

^{1}H NMR (400 MHz, $CDCl_3$) δ 4.67 (d, *J* = 4.3 Hz, 1H), 3.78 (dd, *J* = 8.5, 4.2 Hz, 1H), 2.46 – 2.34 (m, 1H), 2.02 – 1.83 (m, 2H), 1.77 – 1.55 (m, 4H), 1.47 – 1.33 (m, 2H).
^{13}C NMR (101 MHz, $CDCl_3$) δ 173.6, 168.5, 70.1, 62.4, 38.3, 30.2, 30.0, 25.2, 25.0.
HR-MS, *m/z* (ESI-) calcd for $C_9H_{12}O_4$: 183.0657 $[M-H]^-$, found 183.0651.

Synthesis of standards for validation of β-lactone assembly pathway

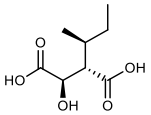


*Synthesis of **1***: To a solution of **S6** (1.4 mmol, 1.0 equiv) in 14 mL THF was added LiOH solution dropwise (2.0 mmol, 1.4 equiv, dissolved in 7 mL of water). The mixture was stirred at room temperature for 4-6 h until reaction complete (judged by TLC). The reaction was diluted with water (40 mL) and concentrated *in vacuo* to remove THF. The resulting aqueous solution was washed with 40 mL EtOAc (this organic layer was discarded), and then adjusted to pH 2 using 1M HCl. The acidified aqueous solution was extracted with EtOAc three times. The organic layers were combined, dried over $MgSO_4$, and concentrated *in vacuo*. The residue was dissolved in 1 mL DCM and cooled to 0 °C. To this solution was added 1 mL of TFA dropwise. The reaction was kept at 0 °C for 16 h before solvent and excess reagent was removed by a stream of nitrogen gas at 0 °C. The residue was dissolved in water and lyophilised to obtain **1**.



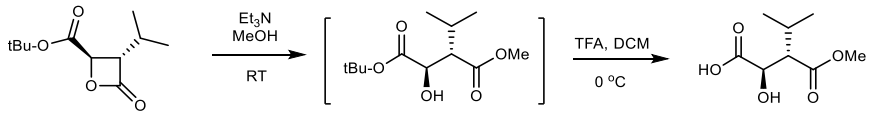
1a

(2*R*,3*S*)-2-hydroxy-3-isopropylsuccinic acid (**1a**) was obtained as a white powder, 201 mg, 82%.
^{1}H NMR (400 MHz, D_2O) δ 4.55 (d, *J* = 4.4 Hz, 1H), 2.64 (dd, *J* = 9.0, 4.4 Hz, 1H), 2.17 – 1.99 (m, 1H), 1.02 (d, *J* = 6.7 Hz, 3H), 0.99 (d, *J* = 6.7 Hz, 3H).
^{13}C NMR (101 MHz, D_2O) δ 176.8, 176.6, 69.3, 55.4, 26.4, 19.9, 19.3.
HR-MS, *m/z* (ESI+) calcd for $C_7H_{12}O_5$: 199.0582 $[M+Na]^+$, found 199.0578.
The NMR spectra is in agreement with published data.[S10]



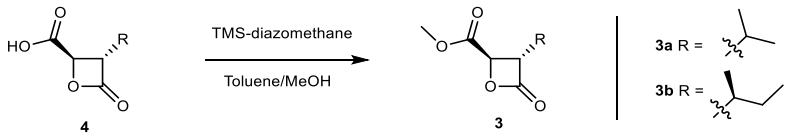
1b

(2*S*,3*R*)-2-((*S*)-*sec*-butyl)-3-hydroxysuccinic acid (**1b**) was a white powder, 72 mg, 89% yield.
^{1}H NMR (400 MHz, D_2O) δ 4.57 (d, *J* = 4.2 Hz, 1H), 2.76 (dd, *J* = 8.9, 4.1 Hz, 1H), 1.98 – 1.86 (m, 1H), 1.55 – 1.43 (m, 1H), 1.29 – 1.16 (m, 1H), 1.01 (d, *J* = 6.8 Hz, 3H), 0.91 (t, *J* = 7.1 Hz, 3H).
^{13}C NMR (101 MHz, D_2O) δ 177.1, 176.8, 69.1, 53.8, 32.7, 26.8, 15.4, 10.4.
The NMR spectra is in agreement with published data.[27]
HR-MS, *m/z* (ESI+) calcd for $C_8H_{14}O_5$: 191.0914 $[M+H]^+$, found 191.0915.



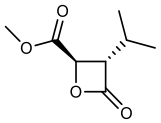
S6a S7

*Synthesis of **S7**:* This procedure adapted from literature.[S11] To a solution of **S6a** (43 mg, 0.2 mmol, 1.0 equiv) in 2 mL MeOH was added triethyl amine (81 mg, 0.8 mmol, 4.0 equiv). The reaction mixture was stirred at room temperature for 16 h before concentrating *in vacuo*. The residue was dissolved in 0.5 mL DCM and cooled to 0 °C. To this solution was added 0.5 mL of TFA dropwise. The reaction mixture was kept at 0 °C for 16 h. The DCM and TFA was removed by a stream of nitrogen gas at 0 °C and the residue was further dried under high vacuum to obtain 39 mg of **S7** in quantitative yield.
^{1}H NMR (400 MHz, $CDCl_3$) δ 4.44 (d, *J* = 3.1 Hz, 1H), 3.72 (s, 3H), 2.72 (dd, *J* = 8.6, 2.0 Hz, 1H), 2.30 – 2.15 (m, 1H), 1.08 (d, *J* = 6.8 Hz, 3H), 1.00 (d, *J* = 6.8 Hz, 3H).
^{13}C NMR (101 MHz, $CDCl_3$) δ 177.4, 174.9, 69.9, 54.2, 52.3, 28.0, 20.9, 20.5.
HR-MS, *m/z* (ESI-) calcd for $C_8H_{14}O_5$: 189.0763 $[M-H]^-$, found 189.0719.



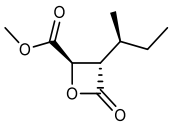
4 3

*Synthesis of **3***: To a solution of **4a** (as an example, 0.21 mmol, 1.0 equiv) in 1.25 mL MeOH and 5 mL Toluene (MeOH/Toluene = 1:4, v/v) was added TMS-diazomethane solution (2 M in Hexane, 300 µL, 3.3 equiv) dropwise at RT. The mixture was stirred at RT for 5-10 min until N_2 bubbling ceased and the yellow colour of the solution persists. The reaction was quenched by adding 50 µL acetic acid and the yellow colour of the solution disappear. The solvent was removed *in vacuo* and the crude product was purified by silica gel column (EtOAc/Hexane, 5-10% EtOAc) to obtain **3** as colourless oil.



3a

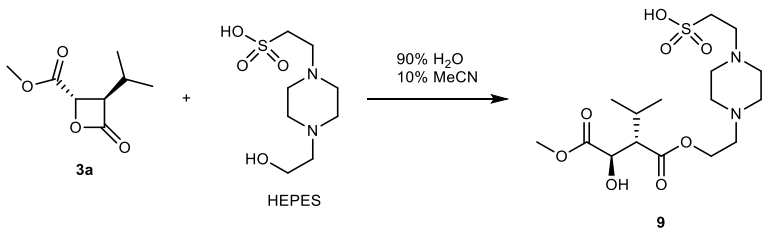
3a was obtained as a colourless oil, 21 mg, 58% yield.
^{1}H NMR (400 MHz, $CDCl_3$) δ 4.64 (d, *J* = 4.4 Hz, 1H), 3.84 (s, 3H), 3.55 (dd, *J* = 8.4, 4.4 Hz, 1H), 2.26 – 2.13 (m, 1H), 1.11 (d, *J* = 6.7 Hz, 3H), 1.07 (d, *J* = 6.8 Hz, 3H).
^{13}C NMR (101 MHz, $CDCl_3$) δ 168.9, 168.4, 69.8, 64.3, 53.1, 27.8, 20.1, 19.6.
HR-MS by ESI was not successful.



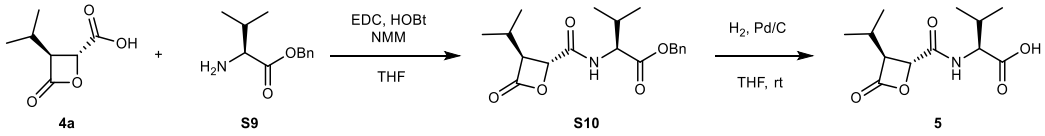
3b

3b was obtained as a colourless oil, 19 mg, 88% yield.

^{1}H NMR (400 MHz, $CDCl_3$) δ 4.66 (d, *J* = 4.4 Hz, 1H), 3.85 (s, 3H), 3.68 (dd, *J* = 7.8, 4.4 Hz, 1H), 2.07 – 1.93 (m, 1H), 1.71 – 1.56 (m, 1H), 1.39 – 1.25 (m, 1H), 1.05 (d, *J* = 6.7 Hz, 3H), 0.94 (t, *J* = 7.5 Hz, 3H).
^{13}C NMR (101 MHz, $CDCl_3$) δ 169.0, 168.8, 69.3, 62.9, 53.1, 33.7, 26.9, 16.4, 11.1.
HR-MS by ESI was not successful.



*Synthesis of **9** (buffer adduct)*: A 100 µL solution of **3a** (200 mM, 3.4 mg, 0.02 mmol) in MeCN was mixed with 900 µL HEPES buffer (200 mM, pH 7.0). The mixture was incubated at room temperature for 3 h and then concentrated in vacuo. The residue was purified by preparative HPLC on a C18 column (mobile phase A = H_2O with 0.1% formic acid, mobile phase B = MeCN with 0.1% formic acid, gradient B: 2-50%) to obtain 1.5 mg of **9** as white solid in 37% isolated yield.
^{1}H NMR (400 MHz, D_2O with 0.3% DCl) δ 4.64 (d, *J* = 4.2 Hz, 1H), 4.60 – 4.44 (m, 2H), 3.80 (brs, 8H), 3.77 (s, 3H), 3.73 – 3.63 (m, 4H), 3.39 (dd, *J* = 8.4, 6.3 Hz, 2H), 2.81 (dd, *J* = 8.3, 4.2 Hz, 1H), 2.14 – 2.04 (m, 1H), 0.98 (d, *J* = 6.8 Hz, 3H), 0.95 (d, *J* = 6.7 Hz, 3H).
^{13}C NMR (201 MHz, D_2O with 0.3% DCl) δ 175.7, 173.1, 69.1, 58.3, 55.2, 55.2, 53.0, 52.3, 49.0, 48.9, 44.7, 26.4, 19.8, 19.1.
HR-MS, *m/z* (ESI-) calcd for $C_{16}H_{30}N_2O_8S$: 409.1645 $[M-H]^-$, found 409.1602.

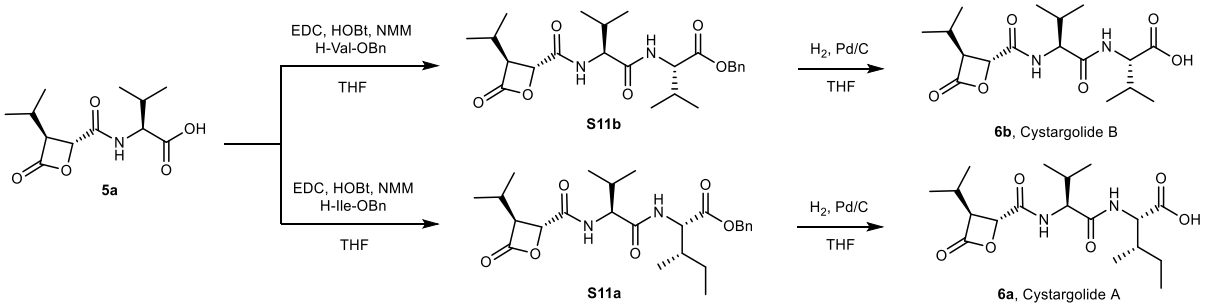


*Synthesis of CysC product **5***: At 0 °C, to a solution of lactone acid **4a** (80 mg, 0.5 mmol, 1.0 equiv) and N-methyl morpholine (NMM, 152 mg, 1.5 mmol, 3.0 equiv) in 10 mL THF was added H-Val-OBn hydrochloride **S9** (122 mg, 0.5 mmol, 1.0 equiv). The solution was stirred at 0 °C for 5 min before EDC·HCl (105 mg, 0.55 mmol, 1.1 equiv), and HOBt·H_2O (74 mg, 0.55 mmol, 1.1 equiv) were quickly added in one portion. The reaction was flushed with nitrogen and the mixture stirred and allowed to gradually reach room temperature for 16 h. The reaction mixture was concentrated *in vacuo*. The residue was taken up with EtOAc (50 mL), washed with 5% (wt) citric acid solution, sat. $NaHCO_3$, brine, dried over $MgSO_4$ and concentrated *in vacuo*. The crude product was purified by silica gel column (EtOAc/Hexane, 5-20% EtOAc) to obtain 95 mg **S10** as colourless oil in 55% yield.
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.42 – 7.29 (m, 5H), 6.73 (d, *J* = 9.0 Hz, 1H), 5.24 (d, *J* = 12.2 Hz, 1H), 5.12 (d, *J* = 12.1 Hz, 1H), 4.62 (d, *J* = 4.5 Hz, 1H), 4.58 (dd, *J* = 8.9, 4.8 Hz, 1H), 3.42 (dd, *J* = 8.2, 4.6 Hz, 1H), 2.35 – 2.10 (m, 2H), 1.09 (dd, *J* = 9.3, 6.7 Hz, 6H), 0.95 (d, *J* = 6.8 Hz, 3H), 0.89 (d, *J* = 7.0 Hz, 3H).

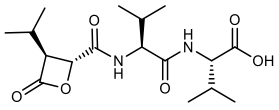
^{13}C NMR (101 MHz, $CDCl_3$) δ 170.9, 168.6, 168.1, 135.2, 128.8, 128.8, 128.6, 71.2, 67.4, 64.6, 57.1, 31.0, 28.0, 20.1, 19.5, 19.2, 17.7.
HR-MS, *m/z* (ESI+) calcd for $C_{19}H_{25}NO_5$: 370.1630 $[M+Na]^+$, found 370.1625.

A solution of **S10** (95 mg, 0.27 mmol) in 5 mL of dry THF was treated with Pd/C (45 mg, 10% wt.). The reaction flask was flushed with hydrogen gas and the mixture stirred under hydrogen atmosphere at room temperature for 16 h. The mixture was diluted with THF and filtered through a celite plug, rinsing with THF. The filtrate was concentrated *in vacuo* to give 60 mg of **5** as white solid (yield 87%).
^{1}H NMR (400 MHz, $CDCl_3$) δ 6.77 (d, *J* = 8.8 Hz, 1H), 4.67 (d, *J* = 4.6 Hz, 1H), 4.57 (dd, *J* = 8.9, 4.8 Hz, 1H), 3.55 (dd, *J* = 8.2, 4.5 Hz, 1H), 2.42 – 2.11 (m, 2H), 1.12 (dd, *J* = 10.3, 6.6 Hz, 6H), 1.01 (d, *J* = 6.8 Hz, 3H), 0.98 (d, *J* = 6.8 Hz, 3H).
^{13}C NMR (101 MHz, $CDCl_3$) δ 175.5, 168.8, 168.7, 71.1, 64.6, 57.1, 30.8, 28.0, 20.1, 19.5, 19.2, 17.7.
HR-MS, *m/z* (ESI) calcd for $C_{12}H_{19}NO_5$: 280.1161 $[M+Na]^+$, found 280.1150.

Synthesis of Cystargolide A (**6a**) and B (**6b**)

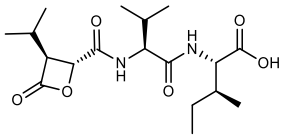


The procedure used for the synthesis of **6a/b** Cystargolide A and B were the same as the synthesis of **5a** mentioned above. Briefly, **5a** was coupled with either H-Val-OBn or H-Ile-OBn to obtain **S11a/b**, which underwent deprotection to obtain **6a/b**.



6b
(Cystargolide B)

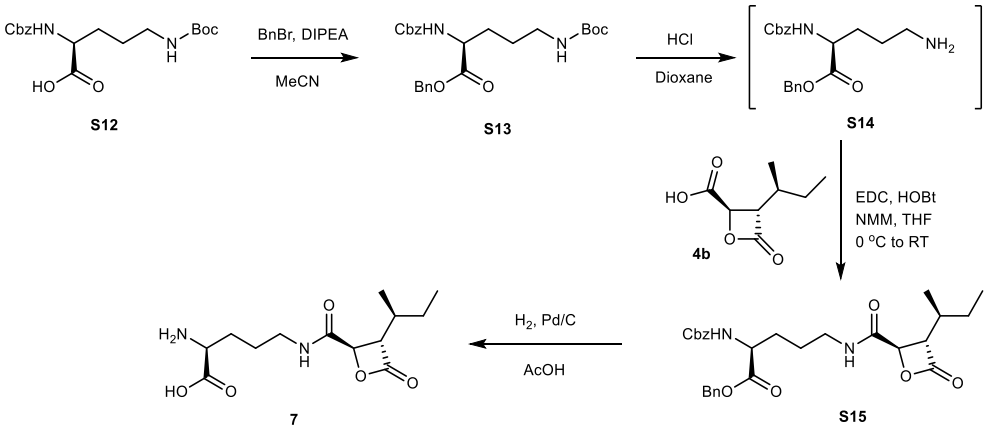
6b was obtained as white solid, 30 mg, 54% yield over two steps.
^{1}H NMR (400 MHz, DMSO-d6) δ 12.56 (brs, 1H), 8.55 (d, *J* = 9.0 Hz, 1H), 8.14 (d, *J* = 8.2 Hz, 1H), 5.02 (d, *J* = 4.4 Hz, 1H), 4.42 – 4.35 (m, 1H), 4.16 – 4.07 (m, 1H), 3.51 (dd, *J* = 8.3, 4.2 Hz, 1H), 2.19 – 1.99 (m, 3H), 1.00 (d, *J* = 6.7 Hz, 3H), 0.97 (d, *J* = 6.7 Hz, 3H), 0.92 – 0.80 (m, 12H).
^{13}C NMR (101 MHz, DMSO-d6) δ 172.8, 170.8, 170.1, 167.3, 70.1, 62.7, 57.33, 57.26, 30.90, 29.60, 26.68, 19.42, 19.29, 19.09, 18.09, 17.88.
HR-MS, *m/z* (ESI-) calcd for $C_{17}H_{28}N_2O_6$: 355.1875 $[M-H]^-$, found 355.1879.
The NMR spectra is in agreement with published data.[37]



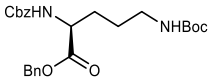
6a
(Cystargolide A)

6a was obtained as white solid, 20 mg, 40% yield over two steps.
^{1}H NMR (400 MHz, MeOD) δ 4.81 (d, *J* = 4.3 Hz, 1H), 4.41 – 4.35 (m, 1H), 4.30 (d, *J* = 7.9 Hz, 1H), 3.56 (dd, *J* = 8.5, 4.4 Hz, 1H), 2.23 – 2.05 (m, 2H), 1.96 – 1.86 (m, 1H), 1.59 – 1.47 (m, 1H), 1.28 (d, *J* = 16.6 Hz, 1H), 1.10 (d, *J* = 6.7 Hz, 3H), 1.06 (d, *J* = 6.7 Hz, 3H), 1.01 – 0.88 (m, 12H).
^{13}C NMR (126 MHz, MeOD) δ 174.6, 173.2, 170.6, 170.4, 72.2, 65.3, 60.2, 58.3, 38.3, 32.0, 28.8, 26.2, 20.4, 19.8, 19.7, 18.8, 16.0, 11.8.
HR-MS, *m/z* (ESI-) calcd for $C_{18}H_{30}N_2O_6$: 369.2031 [M-H]$^-$, found 369.2036.

Synthesis of **7**



*Synthesis of **S13***: Cbz-Orn-(N-Boc) **S12** (672.8 mg, 2.0 mmol, 1.0 equiv) was dissolved in MeCN (17.5 mL). To this solution was added DIPEA (271.4 mg, 1.05 equiv, 2.1 mmol) and benzyl bromide (359.2 mg, 1.05 equiv, 2.1 mmol). The reaction was stirred at room temperature for 14 h, before solvent was removed *in vacuo*, and the residue was re-dissolved in ethyl acetate (25 mL) and washed with water and brine. The organic layer was dried over $MgSO_4$ and concentrated *in vacuo*. The crude product was purified by silica gel column to obtain 680 mg of **S13** in 75% yield.

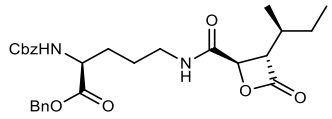


S13

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.41 – 7.28 (m, 10H), 5.37 (d, *J* = 7.1 Hz, 1H), 5.20 – 5.03 (m, 4H), 4.58 – 4.34 (m, 2H), 3.15 – 2.99 (m, 2H), 1.94 – 1.78 (m, 1H), 1.74 – 1.63 (m, 1H), 1.43 (s, 11H).
^{13}C NMR (101 MHz, $CDCl_3$) δ 172.2, 156.0, 136.3, 135.3, 128.8, 128.7, 128.5, 128.3, 128.2, 79.4, 67.4, 67.2, 53.8, 40.1, 30.1, 28.5, 26.1.

The NMR spectra is in agreement with published data.[S12]
HR-MS, *m/z* (ESI+) calcd for $C_{25}H_{32}N_2O_6$: 479.2158 $[M+Na]^+$, found 479.2150.

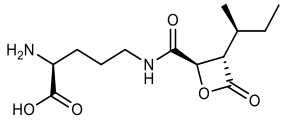
*Synthesis of **S15***: A solution of **S13** (137 mg, 0.3 mmol, 1.0 equiv) in 3 mL 4 M HCl in dioxane was stirred at room temperature for over 3 h. Afterwards, TLC indicated complete removal of Boc protecting group and the reaction was concentrated *in vacuo*. The residue (**S14**) was taken up in 1 mL THF and added to a solution of lactone acid **4b** (52 mg, 0.3 mmol, 1.0 equiv) and N-methyl morpholine (NMM, 91 mg, 0.9 mmol, 3.0 equiv) in 4 mL THF at 0 °C. The solution was stirred at 0 °C for 5 min before EDC·HCl (69 mg, 0.36 mmol, 1.2 equiv), and HOBt·H_2O (49 mg, 0.36 mmol, 1.2 equiv) were quickly added in one portion. The reaction was flushed with nitrogen and the mixture stirred and allowed to gradually reach room temperature for 16 h. The reaction mixture was concentrated *in vacuo* and the residue was taken up with EtOAc (50 mL), washed with 5% (wt) citric acid solution, sat. $NaHCO_3$, brine, dried over $MgSO_4$ and concentrated *in vacuo*. The crude product was purified by silica gel column (EtOAc/Hexane, 20-50% EtOAc) to give 80 mg **S15** as white solid in 52% yield over two steps.



S15

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.45 – 7.28 (m, 10H), 6.38 (s, 1H), 5.38 (d, *J* = 8.9 Hz, 1H), 5.23 – 5.07 (m, 4H), 4.56 (d, *J* = 4.5 Hz, 1H), 4.49 – 4.36 (m, 1H), 3.57 (dd, *J* = 7.6, 4.6 Hz, 1H), 3.37 – 3.21 (m, 2H), 2.03 – 1.83 (m, 2H), 1.76 – 1.62 (m, 2H), 1.57 – 1.42 (m, 2H), 1.33 (d, *J* = 21.2 Hz, 1H), 1.07 (d, *J* = 6.7 Hz, 3H), 0.95 (t, *J* = 7.5 Hz, 3H).
^{13}C NMR (101 MHz, $CDCl_3$) δ 172.0, 169.3, 168.2, 156.0, 136.2, 135.2, 128.8, 128.7, 128.7, 128.5, 128.4, 128.3, 70.8, 67.5, 67.2, 63.0, 53.5, 38.6, 33.9, 30.1, 26.8, 25.3, 16.4, 11.1.
HR-MS, *m/z* (ESI+) calcd for $C_{28}H_{34}N_2O_7$: 533.2264 $[M+Na]^+$, found 533.2261.

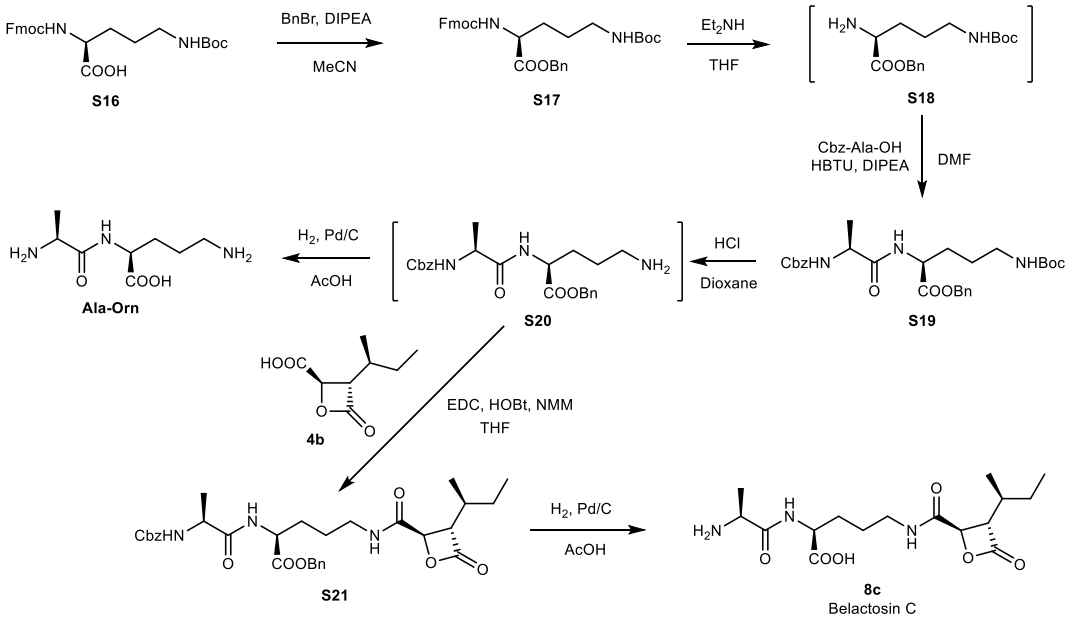
*Synthesis of **7***: A solution of **S15** (79 mg, 0.15 mmol) in acetic acid (10 mL) was treated with Pd/C (45 mg, 10% wt.). The reaction was flushed with hydrogen gas and the mixture stirred under a hydrogen atmosphere (balloon) at room temperature for 16 h. The mixture was filtered through a short celite plug, rinsing with acetic acid. The collected filtrate was concentrated *in vacuo* and the residual was purified by Bond Elut^{TM} C18 column (Agilent, eluent H_2O/MeCN, 0-30% MeCN). The fractions containing pure product were combined, concentrated *in vacuo* to remove MeCN and then lyophilised to obtain 28 mg **7** as white powder in 63% yield.



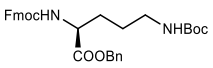
7

^{1}H NMR (400 MHz, D_2O) δ 4.93 (d, *J* = 4.3 Hz, 1H), 3.86 (dd, *J* = 7.5, 4.4 Hz, 1H), 3.76 (t, *J* = 6.1 Hz, 1H), 3.34 (t, *J* = 6.8 Hz, 2H), 2.13 – 1.99 (m, 1H), 1.98 – 1.80 (m, 2H), 1.77 – 1.49 (m, 3H), 1.41 – 1.26 (m, 1H), 1.04 (d, *J* = 6.7 Hz, 3H), 0.92 (t, *J* = 7.5 Hz, 3H).
^{13}C NMR (101 MHz, D_2O) δ 174.3, 172.6, 170.4, 71.0, 62.0, 54.4, 38.6, 32.8, 27.7, 26.2, 24.2, 15.4, 10.3.
HR-MS, *m/z* (ESI+) calcd for $C_{13}H_{22}N_2O_5$: 287.1607 $[M+H]^+$, found 287.1609.

Synthesis of **Ala-Orn** and belactosin C (**8c**)



*Synthesis of **S17***: Fmoc-Orn-(N-Boc) **S16** (909 mg, 2.0 mmol, 1.0 equiv) was added to MeCN (17.5 mL). To the mixture was added DIPEA (271.4 mg, 1.05 equiv, 2.1 mmol) and benzyl bromide (359.2 mg, 1.05 equiv, 2.1 mmol). The reaction was stirred at room temperature for 14 h before solvent was removed *in vacuo*, and the residue was taken up by ethyl acetate (25 mL) and washed with 5% wt. citric acid and brine. The organic layer was dried over $MgSO_4$ and concentrated *in vacuo*. The crude product was purified by silica gel column (EtOAc/Hexane, 20-50% EtOAc) to obtain 710 mg of **S17** in 65% yield.

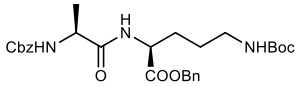


S17

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.77 (d, *J* = 7.5 Hz, 2H), 7.60 (d, *J* = 7.6 Hz, 2H), 7.48 – 7.27 (m, 9H), 5.48 (d, *J* = 7.7 Hz, 1H), 5.26 – 5.10 (m, 2H), 4.60 – 4.32 (m, 4H), 4.21 (t, *J* = 7.1 Hz, 1H), 3.18 – 2.96 (m, 2H), 1.97 – 1.82 (m, 1H), 1.77 – 1.66 (m, 1H), 1.58 – 1.45 (m, 2H), 1.44 (s, 9H).
^{13}C NMR (101 MHz, $CDCl_3$) δ 172.3, 156.1, 144.0, 143.8, 141.4, 135.3, 128.8, 128.7, 128.5, 127.8, 127.2, 125.2, 120.1, 79.4, 67.4, 67.1, 53.8, 47.3, 40.0, 30.0, 28.5, 26.2.
HR-MS, *m/z* (ESI+) calcd for $C_{32}H_{36}N_2O_6$: 567.2471 $[M+Na]^+$, found 567.2465.

*Synthesis of **S19***: A solution of **S17** (700 mg, 1.3 mmol, 1.0 equiv) in 2.5 mL THF was treated with 2.5 mL diethylamine. The reaction was stirred at room temperature for 1 h before another 1 mL diethylamine was added. The mixture was stirred at room temperature for additional 2 h before concentrating *in vacuo*. The residue was azeotroped with toluene (2 × 2 mL) under reduced pressure, and the crude product was taken up in DMF (2 mL). To this solution was added Cbz-Alanine (290 mg, 1.3 mmol, 1.0 equiv), HBTU (606 mg, 1.6 mmol, 1.2 equiv) and DIPEA (504 mg, 3.9 mmol, 3.0 equiv). The reaction mixture was stirred at room temperature for 16 h before diluting with DCM (50 mL). The resulting solution was washed with 5% wt. citric acid, brine,

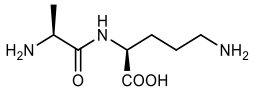
dried over $MgSO_4$ and concentrated *in vacuo*. The crude product was purified by silica gel column (EtOAc/Hexane, 20-50% EtOAc) to give 531 mg **S19** as white solid in 78% yield over two steps.



S19

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.39 – 7.27 (m, 10H), 6.90 (s, 1H), 5.50 (s, 1H), 5.23 – 5.02 (m, 4H), 4.70 – 4.53 (m, 2H), 4.34 – 4.22 (m, 1H), 3.14 – 2.96 (m, 2H), 1.96 – 1.74 (m, 2H), 1.73 – 1.60 (m, 1H), 1.55 – 1.42 (m, 2H), 1.41 (s, 9H), 1.37 (d, *J* = 7.0 Hz, 3H).
^{13}C NMR (101 MHz, $CDCl_3$) δ 172.4, 171.9, 156.3, 156.1, 136.4, 135.3, 128.8, 128.6, 128.5, 128.3, 128.2, 79.5, 67.4, 67.1, 52.3, 50.6, 39.9, 29.0, 28.5, 26.2, 18.6.
HR-MS, *m/z* (ESI+) calcd for $C_{28}H_{37}N_3O_7$: 550.2529 $[M+Na]^+$, found 550.2523.

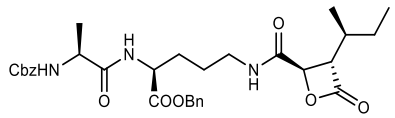
*Synthesis of **Ala-Orn***: A solution of **S19** (200 mg, 0.38 mmol, 1.0 equiv) in 5 mL 4 M HCl (in dioxane) was stirred at room temperature for over 3 h. After TLC indicated complete removal of Boc protecting group the reaction was concentrated *in vacuo*. The residue was taken up in acetic acid (5 mL). To this solution was treated with Pd/C (60 mg, 10% wt.). The reaction was flushed with hydrogen gas and the mixture stirred under a hydrogen atmosphere (balloon) at room temperature for 16 h. The mixture was filtered through a short celite plug, rinsing with acetic acid. The collected filtrate was concentrated *in vacuo* and the residual was purified by Bond Elut^{TM} C18 column (Agilent, eluent H_2O/MeCN with 0.1% HCl, 0-30% MeCN). The fractions containing pure product were combined, concentrated to remove MeCN and then lyophilised to obtain 80 mg **Ala-Orn** di-hydrochloride salt as yellow solid in 76% yield.



Ala-Orn

^{1}H NMR (400 MHz, D_2O) δ 4.45 (dd, *J* = 8.4, 5.3 Hz, 1H), 4.15 (q, *J* = 7.0 Hz, 1H), 3.05 (t, *J* = 7.5 Hz, 2H), 2.07 – 1.95 (m, 1H), 1.91 – 1.71 (m, 3H), 1.57 (d, *J* = 7.2 Hz, 3H).
^{13}C NMR (101 MHz, D_2O) δ 174.7, 170.9, 52.4, 48.9, 38.8, 27.3, 23.2, 16.4.
HR-MS, *m/z* (ESI+) calcd for $C_8H_{17}N_3O_3$: 202.1197 $[M-H]^-$, found 202.1204.

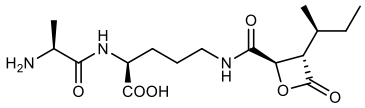
*Synthesis of **S21**:* A solution of **S19** (158 mg, 0.3 mmol, 1.0 equiv) in 3 mL 4M HCl (in dioxane) was stirred at room temperature for over 3 h. After TLC indicated complete removal of Boc protecting group the reaction was concentrated *in vacuo*. The residue (**S20**) was taken up in 1 mL THF and added to a solution of lactone acid **4b** (52 mg, 0.3 mmol, 1.0 equiv) and N-methyl morpholine (NMM, 91 mg, 0.9 mmol, 3.0 equiv) in 4 mL THF at 0 °C. The solution was stirred at 0 °C for 5 min before EDC·HCl (69 mg, 0.36 mmol, 1.2 equiv), and HOBt·H_2O (50 mg, 0.36 mmol, 1.2 equiv) were quickly added in one portion. The reaction was flushed with nitrogen and the mixture stirred and allowed to gradually reach room temperature for 16 h. The reaction mixture was concentrated *in vacuo* and the residue was taken up with EtOAc (50 mL), washed with 5% (wt) citric acid solution, sat. $NaHCO_3$, brine, dried over $MgSO_4$ and concentrated *in vacuo*. The crude product was purified by silica gel column (EtOAc/Hexane, 50-60% EtOAc) to give 85 mg **S21** as white solid in 49% yield over two steps.



S21

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.45 – 7.25 (m, 10H), 6.76 (d, *J* = 8.2 Hz, 1H), 6.58 (s, 1H), 5.41 (d, *J* = 6.8 Hz, 1H), 5.25 – 5.03 (m, 4H), 4.67 – 4.57 (m, 1H), 4.54 (d, *J* = 4.5 Hz, 1H), 4.32 – 4.18 (m, 1H), 3.56 (dd, *J* = 7.7, 4.5 Hz, 1H), 3.34 – 3.13 (m, 2H), 2.03 – 1.81 (m, 2H), 1.78 – 1.58 (m, 4H), 1.57 – 1.41 (m, 2H), 1.37 (d, *J* = 7.1 Hz, 3H), 1.34 – 1.25 (m, 1H), 1.05 (d, *J* = 6.7 Hz, 3H), 0.93 (t, *J* = 7.4 Hz, 3H).
^{13}C NMR (101 MHz, $CDCl_3$) δ 172.7, 171.7, 169.4, 168.3, 156.2, 136.2, 135.2, 128.8, 128.7, 128.6, 128.4, 128.3, 128.1, 70.9, 67.4, 67.2, 62.9, 51.9, 50.5, 38.5, 33.8, 29.3, 26.7, 25.4, 18.7, 16.4, 11.1.
HR-MS, *m/z* (ESI) calcd for $C_{31}H_{39}N_3O_8$: 604.2635 $[M+Na]^+$, found 604.2610.
The NMR spectra is in agreement with published data[S9].

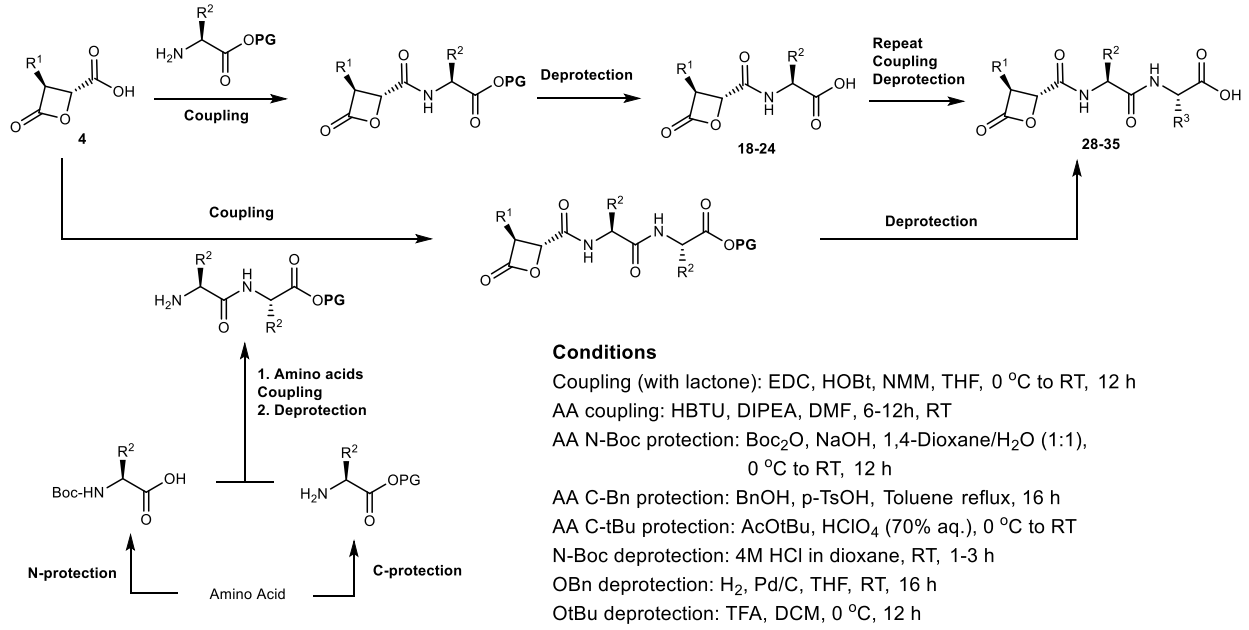
*Synthesis of **8c**:* A solution of **S21** (80 mg, 0.14 mmol) in acetic acid (10 mL) was treated with Pd/C (45 mg, 10% wt.). The reaction was flushed with hydrogen gas and the mixture stirred under a hydrogen atmosphere (balloon) at room temperature for 16 h. The mixture was filtered through a short celite plug, rinsing with acetic acid. The collected filtrate was concentrated *in vacuo* and the residual was purified by Bond Elut^{TM} C18 column (Agilent, eluent H_2O/MeCN, 0-30% MeCN). The fractions containing pure product were combined, concentrated to remove MeCN and then lyophilised to obtain 26 mg Belactosin C **(8c)** as white powder in 53% yield.



8c
(Belactosin C)

^{1}H NMR (400 MHz, D_2O) δ 4.92 (d, *J* = 4.4 Hz, 1H), 4.17 (dd, *J* = 8.1, 5.2 Hz, 1H), 4.11 (q, *J* = 7.1 Hz, 1H), 3.85 (dd, *J* = 7.5, 4.4 Hz, 1H), 3.31 (t, *J* = 6.9 Hz, 2H), 2.13 – 1.98 (m, 1H), 1.89 – 1.78 (m, 1H), 1.78 – 1.67 (m, 1H), 1.66 – 1.49 (m, 6H), 1.40 – 1.28 (m, 1H), 1.03 (d, *J* = 6.7 Hz, 3H), 0.91 (t, *J* = 7.5 Hz, 3H).
^{13}C NMR (101 MHz, D_2O) δ 178.1, 172.7, 170.2, 170.0, 71.0, 62.0, 55.1, 49.0, 38.8, 32.8, 28.6, 26.2, 24.9, 16.4, 15.4, 10.3.
HR-MS, *m/z* (ESI) calcd for $C_{16}H_{27}N_3O_6$: 358.1978 $[M+H]^+$, found 358.1985.
The NMR spectra is in agreement with published data[S13].

Synthesis of Cystargolide analogues



General procedure A: Bn Protection of amino acids. To a solution of amino acid (**AA**) (2.1 mmol, 1.0 equiv) in 5 mL Toluene was added TsOH monohydrate (590 mg, 3.1 mmol, 1.5 equiv) and BnOH (2.2 mL, 21 mmol, 10.0 equiv). The mixture was stirred and heated under reflux for 18 h before diluting with 50 mL water and washed with Et_2O. The aqueous layer was basified with saturated aqueous $NaHCO_3$ and extracted with EtOAc (3 times). The organic layers were combined, dried over $MgSO_4$ and concentrated *in vacuo* to afford the crude **H-AA-OBn**, which was used in the coupling step without further purification.

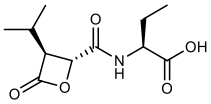
General procedure B: tBu Protection of amino acids. To a solution of amino acid (**AA**) (0.88 mmol) in AcO*t*Bu (2.2 mL), 70% aqueous $HClO_4$ (2.64 mmol, 3.0 equiv) was added dropwise at 0 °C. After completion of the addition of $HClO_4$, the reaction mixture was stirred for 12 h at room temperature, and then quenched with 1 M aqueous HCl. The aqueous layer was diluted with H_2O and washed with Et_2O. Then the aqueous layer was basified with saturated aqueous $NaHCO_3$ (pH 9) and extracted with DCM (3 times). The organic layer was dried over $MgSO_4$ and concentrated to afford the crude **H-AA-O*t*Bu**, which was used in the next reaction without further purification.

General procedure C: N-Boc protection of amino acids. At 0 °C, to a solution of amino acid (**AA**, 4.0 mmol) and NaOH (6.0 mmol, 1.5 equiv) in H_2O/1,4-dioxane (H_2O/dioxane = 1:1, 30 mL) was added Boc_2O (6.0 mmol dissolved in 8 mL 1,4-dioxane). The reaction was then allowed to reach room temperature and kept stirring for 12 h. The mixture was diluted with water and washed with EtOAc. The aqueous layer was adjusted to pH 2-3 with 5% citric acid, extracted with EtOAc, dried and concentrated *in vacuo* to give **Boc-AA-OH**.

General procedure D: coupling of protected amino acids. To a solution of **Boc-AA-OH** (1.0 mmol) and C-protected amino acid (**H-AA-OBn/O*t*Bu**, 1.0 mmol) in 2.5 mL DMF was added

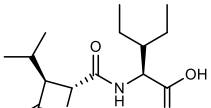
HBTU (1.0 mmol) and DIPEA (2.0 mmol). After stirring at RT for 12 h, the reaction was diluted with EtOAc, washed with sat. $NaHCO_3$, 1 M HCl, brine, dried over $MgSO_4$, and concentrated *in vacuo* to give protected dipeptide.

The coupling of β-lactone containing acid with C-protected amino acids or dipeptides followed the same procedure as the synthesis of CysC product **5** (using EDC, HOBt, NMM, THF). The crude product was deprotected using either H_2, Pd/C (in the case of OBn protection, following the same procedure as described in the synthesis of **5**), or TFA/DCM (in the case of O*t*Bu protection, following the same procedure as described in the synthesis of **4**).



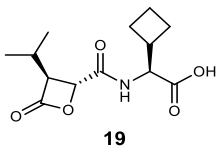
18

18 was prepared as a white solid, 22.3 mg, 29% yield over two steps.
^{1}H NMR (400 MHz, MeOD) δ 4.78 (d, *J* = 4.3 Hz, 1H), 4.42 – 4.32 (m, 1H), 3.58 (dd, *J* = 8.4, 4.3 Hz, 1H), 2.24 – 2.11 (m, 1H), 2.03 – 1.89 (m, 1H), 1.85 – 1.70 (m, 1H), 1.10 (d, *J* = 6.7 Hz, 3H), 1.07 (d, *J* = 6.7 Hz, 3H), 0.98 (t, *J* = 7.4 Hz, 3H).
^{13}C NMR (101 MHz, MeOD) δ 174.7, 170.7, 170.7, 72.2, 65.2, 55.0, 28.8, 25.5, 20.4, 19.8, 10.6.
HR-MS, *m/z* (ESI-) calcd for $C_{11}H_{17}NO_5$: 242.1034 $[M-H]^-$, found 242.1082.



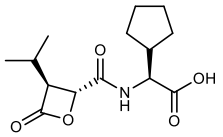
21

21 was prepared as a white solid, 18.2 mg, 32% yield over two steps.
^{1}H NMR (400 MHz, MeOD) δ 4.85 (d, *J* = 4.4 Hz, 1H), 4.66 (d, *J* = 4.6 Hz, 1H), 3.55 (dd, *J* = 8.5, 4.3 Hz, 1H), 2.24 – 2.12 (m, 1H), 1.88 – 1.78 (m, 1H), 1.52 – 1.40 (m, 2H), 1.34 – 1.24 (m, 2H), 1.08 (dd, *J* = 16.5, 6.7 Hz, 6H), 1.01 – 0.87 (m, 6H).
^{13}C NMR (101 MHz, MeOD) δ 169.3, 169.2, 70.6, 63.9, 53.7, 43.6, 27.4, 22.8, 22.4, 22.1, 19.0, 18.5, 10.6, 10.5.
HR-MS, *m/z* (ESI-) calcd for $C_{14}H_{23}NO_5$: 284.1503 $[M-H]^-$, found 284.1569.



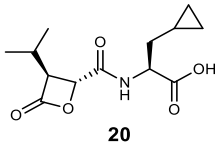
19

19 was prepared as a white solid, 30 mg, 35% yield.
^{1}H NMR (400 MHz, MeOD) δ 4.79 (d, *J* = 4.3 Hz, 1H), 4.40 (d, *J* = 9.1 Hz, 1H), 3.57 (dd, *J* = 8.4, 4.4 Hz, 1H), 2.90 – 2.58 (m, 1H), 2.24 – 2.13 (m, 1H), 2.12 – 1.78 (m, 6H), 1.10 (d, *J* = 6.7 Hz, 3H), 1.06 (d, *J* = 6.7 Hz, 3H).
^{13}C NMR (101 MHz, MeOD) δ 174.0, 170.8, 170.7, 72.1, 65.3, 57.9, 38.1, 28.8, 26.7, 26.2, 20.4, 19.8, 18.7.
HR-MS, *m/z* (ESI-) calcd for $C_{13}H_{19}NO_5$: 268.1190 $[M-H]^-$, found 268.1187.



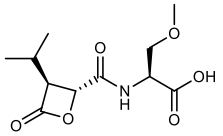
22

22 was prepared as a white solid, 32 mg, 38% yield.
^{1}H NMR (400 MHz, MeOD) δ 4.80 (d, *J* = 4.3 Hz, 1H), 4.39 – 4.32 (m, 1H), 3.56 (dd, *J* = 8.5, 4.3 Hz, 1H), 2.41 – 2.29 (m, 1H), 2.24 – 2.12 (m, 1H), 1.84 – 1.73 (m, 2H), 1.71 – 1.53 (m, 4H), 1.47 – 1.31 (m, 2H), 1.10 (d, *J* = 6.7 Hz, 3H), 1.06 (d, *J* = 6.7 Hz, 3H).
^{13}C NMR (101 MHz, MeOD) δ 174.6, 170.6, 170.5, 72.1, 65.2, 57.3, 42.9, 30.4, 30.0, 28.8, 26.3, 25.9, 20.4, 19.8.
HR-MS, *m/z* (ESI-) calcd for $C_{14}H_{21}NO_5$: 282.1347 [M-H]$^-$, found 282.1393.



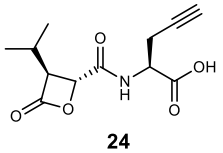
20

20 was prepared as a white solid, 10 mg, 13% yield.
^{1}H NMR (400 MHz, MeOD) δ 4.79 (d, *J* = 4.4 Hz, 1H), 4.53 (dd, *J* = 8.1, 5.5 Hz, 1H), 3.59 (dd, *J* = 8.4, 4.4 Hz, 1H), 2.28 – 2.08 (m, 1H), 1.89 – 1.63 (m, 2H), 1.11 (d, *J* = 6.7 Hz, 3H), 1.07 (d, *J* = 6.7 Hz, 3H), 0.89 – 0.68 (m, 1H), 0.59 – 0.35 (m, 2H), 0.26 – 0.02 (m, 2H).
^{13}C NMR (101 MHz, MeOD) δ 174.7, 170.7, 170.6, 72.3, 65.2, 54.2, 37.2, 28.9, 20.4, 19.8, 8.6, 5.2, 4.5.
HR-MS, *m/z* (ESI-) calcd for $C_{13}H_{19}NO_5$: 268.1185 [M-H]$^-$, found 268.1166.



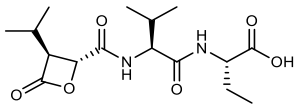
23

23 was prepared as a white solid, 3 mg, 5% yield.
^{1}H NMR (400 MHz, MeOD) δ 4.82 (d, *J* = 4.4 Hz, 1H), 4.65 (dd, *J* = 5.2, 3.6 Hz, 1H), 3.83 (dd, *J* = 9.8, 5.2 Hz, 1H), 3.70 (dd, *J* = 9.8, 3.6 Hz, 1H), 3.60 (dd, *J* = 8.3, 4.4 Hz, 1H), 3.36 (s, 3H), 2.25 – 2.10 (m, 1H), 1.10 (d, *J* = 6.7 Hz, 3H), 1.07 (d, *J* = 6.8 Hz, 3H).
^{13}C NMR (126 MHz, MeOD) δ 172.6, 170.6, 170.6, 72.6, 72.2, 65.3, 59.3, 53.9, 28.9, 20.4, 19.8.
HR-MS, *m/z* (ESI-) calcd for $C_{11}H_{17}NO_6$: 258.0978 [M-H]$^-$, found 258.0959.



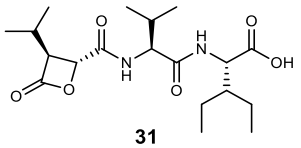
24

24 was prepared as a white solid, 35 mg, 42% yield.
^{1}H NMR (400 MHz, MeOD) δ 4.82 (d, *J* = 4.4 Hz, 1H), 4.62 (dd, *J* = 7.6, 5.0 Hz, 1H), 3.59 (dd, *J* = 8.4, 4.4 Hz, 1H), 2.88 – 2.67 (m, 2H), 2.39 (t, *J* = 2.6 Hz, 1H), 2.26 – 2.12 (m, 1H), 1.11 (d, *J* = 6.7 Hz, 3H), 1.08 (d, *J* = 6.7 Hz, 3H).
^{13}C NMR (101 MHz, MeOD) δ 172.7, 170.6, 79.9, 72.3, 72.1, 65.4, 52.3, 28.9, 22.0, 20.4, 19.8.
HR-MS, *m/z* (ESI-) calcd for $C_{12}H_{15}NO_5$: 252.0877 [M-H]$^-$, found 252.0889.



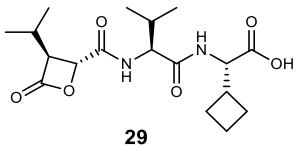
28

28 was prepared as a light-yellow oil, 15.8 mg, 24% yield over two steps.
^{1}H NMR (400 MHz, MeOD) δ 4.81 (d, *J* = 4.3 Hz, 1H), 4.36 – 4.26 (m, 2H), 3.57 (dd, *J* = 8.5, 4.3 Hz, 1H), 2.22 – 2.07 (m, 2H), 1.93 – 1.86 (m, 1H), 1.78 – 1.69 (m, 1H), 1.10 (d, *J* = 6.7 Hz, 3H), 1.06 (d, *J* = 6.7 Hz, 3H), 1.03 – 0.93 (m, 9H).
^{13}C NMR (101 MHz, MeOD) δ 173.6, 171.8, 169.2, 169.0, 70.8, 63.8, 58.7, 53.7, 30.7, 27.4, 24.5, 19.0, 18.4, 18.3, 17.4, 9.2.
HR-MS, *m/z* (ESI-) calcd for $C_{16}H_{26}N_2O_6$: 341.1718 [M-H]$^-$, found 341.1788.



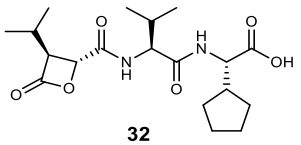
31

31 was prepared as a white solid, 20 mg, 35% yield.
^{1}H NMR (400 MHz, MeOD) δ 4.80 (d, *J* = 4.4 Hz, 1H), 4.61 (d, *J* = 4.8 Hz, 1H), 4.30 (d, *J* = 8.0 Hz, 1H), 3.54 (dd, *J* = 8.5, 4.3 Hz, 1H), 2.23 – 2.03 (m, 2H), 1.81 – 1.68 (m, 1H), 1.50 – 1.26 (m, 4H), 1.10 (d, *J* = 6.7 Hz, 3H), 1.06 (d, *J* = 6.7 Hz, 3H), 1.03 – 0.86 (m, 12H).
^{13}C NMR (101 MHz, MeOD) δ 175.1, 173.6, 170.6, 170.4, 72.2, 65.3, 60.2, 55.2, 45.0, 31.8, 28.8, 23.6, 23.4, 20.4, 19.9, 19.7, 18.8, 12.0, 12.0.
HR-MS, *m/z* (ESI-) calcd for $C_{19}H_{32}N_2O_6$: 383.2182 [M-H]$^-$, found 383.2167.



29

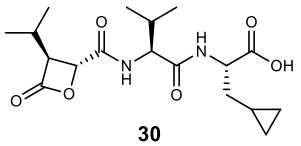
29 was prepared as a white solid, 30 mg, 53% yield.
^{1}H NMR (400 MHz, MeOD) δ 4.81 (d, *J* = 4.4 Hz, 1H), 4.34 (d, *J* = 8.6 Hz, 1H), 4.30 (d, *J* = 7.7 Hz, 1H), 3.56 (dd, *J* = 8.4, 4.4 Hz, 1H), 2.82 – 2.63 (m, 1H), 2.23 – 1.78 (m, 8H), 1.10 (d, *J* = 6.7 Hz, 3H), 1.06 (d, *J* = 6.7 Hz, 3H), 1.00 (d, *J* = 6.8 Hz, 3H), 0.96 (d, *J* = 6.7 Hz, 3H).
^{13}C NMR (101 MHz, MeOD) δ 174.4, 173.3, 170.6, 170.4, 72.2, 65.2, 60.1, 58.0, 38.4, 32.0, 28.8, 26.3, 26.1, 20.4, 19.8, 19.7, 18.7, 18.7.
HR-MS, *m/z* (ESI-) calcd for $C_{18}H_{28}N_2O_6$: 367.1875 [M-H]$^-$, found 367.1869.



32

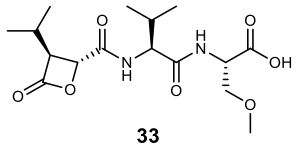
32 was prepared as a white solid, 30 mg, 44% yield.
^{1}H NMR (400 MHz, MeOD) δ 4.80 (d, *J* = 4.4 Hz, 1H), 4.30 (dd, *J* = 9.7, 7.9 Hz, 2H), 3.55 (dd, *J* = 8.5, 4.3 Hz, 1H), 2.35 – 2.24 (m, 1H), 2.21 – 2.04 (m, 2H), 1.82 – 1.71 (m, 2H), 1.70 – 1.62 (m, 2H), 1.61 – 1.52 (m, 2H), 1.46 – 1.31 (m, 2H), 1.10 (d, *J* = 6.7 Hz, 3H), 1.06 (d, *J* = 6.7 Hz, 3H), 0.99 (d, *J* = 6.7 Hz, 3H), 0.95 (d, *J* = 6.7 Hz, 3H).

^{13}C NMR (101 MHz, MeOD) δ 174.9, 173.2, 170.6, 170.4, 72.2, 65.3, 60.1, 57.2, 43.0, 32.0, 30.2, 29.9, 28.8, 26.3, 26.0, 20.4, 19.8, 19.7, 18.8.
HR-MS, *m/z* (ESI-) calcd for $C_{19}H_{30}N_2O_6$: 381.2026 [M-H]$^-$, found 381.2081.



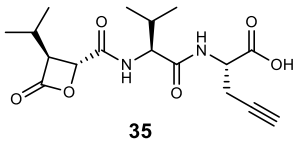
30

30 was prepared as a white solid, 30 mg, 42% yield over two steps.
^{1}H NMR (400 MHz, MeOD) δ 4.81 (d, *J* = 4.3 Hz, 1H), 4.51 – 4.44 (m, 1H), 4.30 (d, *J* = 7.7 Hz, 1H), 3.56 (dd, *J* = 8.4, 4.4 Hz, 1H), 2.24 – 2.06 (m, 2H), 1.78 – 1.61 (m, 2H), 1.10 (d, *J* = 6.7 Hz, 3H), 1.06 (d, *J* = 6.7 Hz, 3H), 1.01 (d, *J* = 6.7 Hz, 3H), 0.96 (d, *J* = 6.8 Hz, 3H), 0.88 – 0.77 (m, 1H), 0.53 – 0.40 (m, 2H), 0.21 – 0.05 (m, 2H).
^{13}C NMR (101 MHz, MeOD) δ 175.0, 173.0, 170.6, 170.4, 72.2, 65.2, 60.1, 54.2, 37.7, 32.1, 28.8, 20.4, 19.8, 19.7, 18.8, 8.6, 5.2, 4.7.
HR-MS, *m/z* (ESI-) calcd for $C_{18}H_{28}N_2O_6$: 367.1869 [M-H]$^-$, found 367.1845.



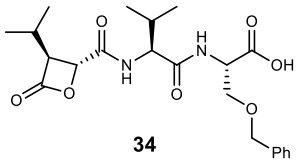
33

33 was prepared as a white solid, 5 mg, 9% yield.
^{1}H NMR (400 MHz, MeOD) δ 4.82 (d, *J* = 4.4 Hz, 1H), 4.52 (s, 1H), 4.35 (d, *J* = 7.2 Hz, 1H), 3.79 (dd, *J* = 9.7, 4.5 Hz, 1H), 3.65 (dd, *J* = 9.7, 3.0 Hz, 1H), 3.60 (dd, *J* = 8.5, 4.3 Hz, 1H), 3.34 (s, 3H), 2.23 – 2.09 (m, 2H), 1.11 (d, *J* = 6.7 Hz, 3H), 1.07 (d, *J* = 6.7 Hz, 3H), 1.00 (d, *J* = 6.8 Hz, 3H), 0.96 (d, *J* = 6.8 Hz, 3H).
^{13}C NMR (126 MHz, MeOD) δ 172.8, 170.6, 170.5, 73.3, 72.2, 65.2, 60.0, 59.3, 32.1, 28.9, 20.4, 19.8, 19.7, 18.6.
HR-MS, *m/z* (ESI-) calcd for $C_{16}H_{26}N_2O_7$: 357.1662 [M-H]$^-$, found 357.1506.



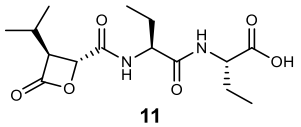
35

35 was prepared as a white solid, 25 mg, 32% yield.
^{1}H NMR (400 MHz, MeOD) δ 8.47 (d, *J* = 7.8 Hz, 1H), 8.19 (d, *J* = 8.7 Hz, 1H), 4.82 (d, *J* = 4.4 Hz, 1H), 4.59 – 4.49 (m, 1H), 4.39 – 4.30 (m, 1H), 3.59 (dd, *J* = 8.5, 4.3 Hz, 1H), 2.80 – 2.65 (m, 2H), 2.34 (s, 1H), 2.24 – 2.09 (m, 2H), 1.10 (d, *J* = 6.7 Hz, 3H), 1.07 (d, *J* = 6.8 Hz, 3H), 1.00 (d, *J* = 6.7 Hz, 3H), 0.96 (d, *J* = 6.7 Hz, 3H).
^{13}C NMR (101 MHz, MeOD) δ 173.0, 170.6, 170.5, 80.1, 72.2, 72.2, 65.3, 60.0, 52.6, 32.2, 28.9, 22.3, 20.4, 19.8, 19.7, 18.6.
HR-MS, *m/z* (ESI-) calcd for $C_{17}H_{24}N_2O_6$: 351.1562 [M-H]$^-$, found 351.1553.



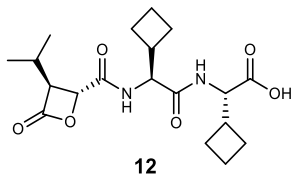
34

34 was prepared as a white solid, 20 mg, 23% yield.
^{1}H NMR (400 MHz, MeOD) δ 7.39 – 7.21 (m, 5H), 4.81 (d, *J* = 4.4 Hz, 1H), 4.67 (t, *J* = 4.4 Hz, 1H), 4.59 – 4.49 (m, 2H), 4.42 – 4.31 (m, 1H), 3.89 (dd, *J* = 9.9, 5.0 Hz, 1H), 3.75 (dd, *J* = 9.9, 3.6 Hz, 1H), 3.57 (dd, *J* = 8.6, 4.4 Hz, 1H), 2.25 – 2.05 (m, 2H), 1.09 (d, *J* = 6.6 Hz, 3H), 1.06 (d, *J* = 6.8 Hz, 3H), 1.00 (d, *J* = 6.8 Hz, 3H), 0.96 (d, *J* = 6.8 Hz, 3H).
^{13}C NMR (101 MHz, MeOD) δ 171.6, 171.4, 169.2, 169.0, 137.8, 128.0, 127.4, 127.3, 72.8, 70.8, 69.2, 63.8, 58.5, 52.7, 48.2, 48.0, 47.8, 47.6, 47.4, 47.2, 47.0, 30.8, 27.4, 19.0, 18.4, 18.3, 17.2.
HR-MS, *m/z* (ESI-) calcd for $C_{22}H_{30}N_2O_7$: 433.1975 $[M-H]^-$, found 433.1964.



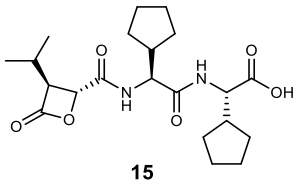
11

11 was prepared as an oil, 39.3 mg, 43% yield over three steps.
^{1}H NMR (400 MHz, MeOD) δ 4.77 (d, *J* = 4.4 Hz, 1H), 4.41 – 4.25 (m, 2H), 3.61 (dd, *J* = 8.4, 4.3 Hz, 1H), 2.24 – 2.09 (m, 1H), 1.93 – 1.85 (m, 2H), 1.81 – 1.66 (m, 2H), 1.10 (d, *J* = 6.7 Hz, 3H), 1.06 (d, *J* = 6.7 Hz, 3H), 0.98 (t, *J* = 7.5 Hz, 6H).
^{13}C NMR (101 MHz, MeOD) δ 175.0, 173.8, 170.6, 170.5, 72.2, 65.2, 56.0, 55.0, 28.8, 26.4, 25.9, 24.0, 20.4, 19.8, 10.6.
HR-MS, *m/z* (ESI-) calcd for $C_{15}H_{24}N_2O_6$: 327.1562 $[M-H]^-$, found 327.1583.



12

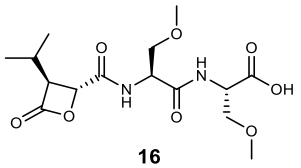
12 was prepared as a white solid, 18.2 mg, 32% yield over three steps.
^{1}H NMR (400 MHz, MeOD) δ 4.78 (d, *J* = 4.4 Hz, 1H), 4.44 (d, *J* = 9.3 Hz, 1H), 4.34 (d, *J* = 8.4 Hz, 1H), 3.58 (dd, *J* = 8.4, 4.4 Hz, 1H), 2.78 – 2.64 (m, 2H), 2.24 – 2.10 (m, 1H), 2.09 – 1.76 (m, 12H), 1.07 (dd, *J* = 16.3, 6.7 Hz, 6H).
^{13}C NMR (101 MHz, MeOD) δ 172.8, 171.4, 169.2, 169.2, 70.8, 63.9, 57.6, 56.3, 37.2, 37.1, 27.4, 25.0, 24.8, 24.5, 24.5, 19.0, 18.4, 17.3, 17.3.
HR-MS, *m/z* (ESI-) calcd for $C_{19}H_{28}N_2O_6$: 379.1875 $[M-H]^-$, found 379.1908.



15

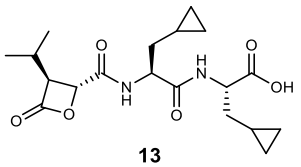
15 was prepared as a white solid, 5 mg, 8% yield over three steps.

^{1}H NMR (400 MHz, MeOD) δ 4.77 (d, *J* = 4.4 Hz, 1H), 4.31 (d, *J* = 7.6 Hz, 1H), 4.28 (d, *J* = 9.6 Hz, 1H), 3.56 (dd, *J* = 8.5, 4.4 Hz, 1H), 2.35 – 2.23 (m, 2H), 2.22 – 2.11 (m, 1H), 1.84 – 1.50 (m, 13H), 1.47 – 1.33 (m, 4H), 1.10 (d, *J* = 6.7 Hz, 3H), 1.06 (d, *J* = 6.7 Hz, 3H).
^{13}C NMR (126 MHz, MeOD) δ 173.4, 170.6, 170.3, 72.2, 65.3, 59.0, 57.4, 43.4, 43.2, 30.4, 30.2, 29.7, 28.8, 26.3, 26.2, 26.0, 25.9, 20.4, 19.8.
HR-MS, *m/z* (ESI-) calcd for $C_{21}H_{32}N_2O_6$: 407.2182 [M-H]$^-$, found 407.2104.



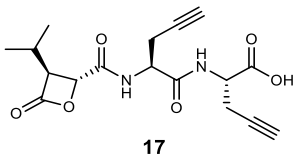
16

16 was prepared as a white solid, 30 mg, 53% yield.
^{1}H NMR (400 MHz, MeOD) δ 4.80 (d, *J* = 4.4 Hz, 1H), 4.69 (t, *J* = 5.6 Hz, 1H), 4.60 (t, *J* = 4.0 Hz, 1H), 3.82 (dd, *J* = 9.8, 4.4 Hz, 1H), 3.71 (d, *J* = 5.6 Hz, 2H), 3.68 – 3.60 (m, 2H), 3.37 (s, 3H), 3.36 (s, 3H), 2.24 – 2.11 (m, 1H), 1.11 (d, *J* = 6.7 Hz, 3H), 1.07 (d, *J* = 6.7 Hz, 3H).
^{13}C NMR (101 MHz, MeOD) δ 172.7, 171.4, 170.7, 170.6, 72.9, 72.7, 72.3, 65.3, 59.4, 59.3, 54.3, 54.1, 28.9, 20.4, 19.8.
HR-MS, *m/z* (ESI-) calcd for $C_{15}H_{24}N_2O_8$: 359.1454 [M-H]$^-$, found 359.1436.



13

13 was prepared as a white solid, 29 mg, 24% yield.
^{1}H NMR (400 MHz, MeOD) δ 4.77 (d, *J* = 4.4 Hz, 1H), 4.60 – 4.45 (m, 2H), 3.61 (dd, *J* = 8.4, 4.4 Hz, 1H), 2.24 – 2.11 (m, 1H), 1.74 – 1.65 (m, 4H), 1.10 (d, *J* = 6.6 Hz, 3H), 1.06 (d, *J* = 6.7 Hz, 3H), 0.88 – 0.73 (m, 2H), 0.55 – 0.41 (m, 4H), 0.21 – 0.07 (m, 4H).
^{13}C NMR (101 MHz, MeOD) δ 173.6, 173.5, 170.6, 170.4, 72.3, 65.2, 55.2, 54.2, 37.9, 37.8, 28.9, 20.4, 19.8, 8.5, 8.5, 5.2, 5.1, 4.7, 4.6.
HR-MS, *m/z* (ESI-) calcd for $C_{19}H_{28}N_2O_6$: 379.1875 [M-H]$^-$, found 379.1912.



17

17 was prepared as a white solid, 20 mg, 30% yield.
^{1}H NMR (400 MHz, MeOD) δ 8.59 (d, *J* = 8.0 Hz, 1H), 8.37 (d, *J* = 7.8 Hz, 1H), 4.81 (d, *J* = 4.4 Hz, 1H), 4.74 – 4.62 (m, 1H), 4.60 – 4.49 (m, 1H), 3.63 (dd, *J* = 8.4, 4.4 Hz, 1H), 2.85 – 2.54 (m, 4H), 2.39 (t, *J* = 2.7 Hz, 1H), 2.35 (t, *J* = 2.6 Hz, 1H), 2.27 – 2.06 (m, 1H), 1.11 (d, *J* = 6.7 Hz, 3H), 1.08 (d, *J* = 6.7 Hz, 3H).
^{13}C NMR (126 MHz, MeOD) δ 172.9, 171.7, 170.7, 170.6, 80.0, 79.9, 72.5, 72.4, 72.2, 65.5, 53.2, 52.6, 28.9, 22.5, 22.4, 20.5, 19.8.
HR-MS, *m/z* (ESI-) calcd for $C_{17}H_{20}N_2O_6$: 347.1243 [M-H]$^-$, found 347.1263.

6. DNA and protein sequences

(DNA sequence codon optimised for E. coli, His-tag shown in bold)

cys pathway

>CysG DNA sequence
ATGGGCCAGACCGCCTCACCGAGAATTGAACTCTGGAATCCCGAGACGTATGACGC
GCTTCGTCGCCAACTGATTCCCAGCTTTGACTTACTGTACGGCTCAGCAGTGAGCGT
TGTGGCCATGTCAGTACCAGCTACAGCCAGAATTCTTGACTTAGGTGCCGGAACTGG
TCTTCTGGGTGCTGCGCTGCGTGAGCGCCTTCCCGATGCGGAACTGCTGCTTCAGGA
TCGATCCCAAGCAATGCTGGAGCAGGCGCGCCAACGTTTTGCCGACGACGACCAAG
TGGCAATTCGCGTAGCTGACCACCTTGATGAGTTACCTGCAGGTCCGTTCGACGCTG
TAGTGTCCGCCTTGTCCATTCATCATTTGGAGCATCAGGACAAGCAGGATTTATTCA
CACGCATTCGTAAGATCCTGCGCCCAGGCGGTATTTTCGTCAACGTTGAGCAAGTGC
TGGCGCCTACTTCAGAACTTGAGAAAATGTATGACCGCCAACACGAGGCTCATGTAC
TGGCTTCTGACACTCCGGCAGAAGAGTGGGCGGCAGGCCGCGAAAGAATGAAGCAC
GACATCCCAATCGATGTGGAGACTCAAATTCAATGGCTGAGAGACGCAGGGTTCAC
TACCGCTGACTGTCTGGCTAAGGACTGGCGGTTTGCAACCTACGCGGGCTGGAATGG
CAGT**CTCGAGCACCACCACCACCACCAC**TGA

>CysG protein sequence
MGQTASPRIELWNPETYDALRRQLIPSFDLLYGSAVSVVAMSVPATARILDLGAGTGLL
GAALRERLPDAELLLQDRSQAMLEQARQRFADDDQVAIRVADHLDELPAGPFDAVVSA
LSIHHLEHQDKQDLFTRIRKILRPGGIFVNVEQVLAPTSELEKMYDRQHEAHVLASDTPA
EEWAAGRERMKHDIPIDVETQIQWLRDAGFTTADCLAKDWRFATYAGWNGS**LEHHH
HHH**

>CysF DNA sequence
ATGTTATATGAAGCCCTGCGCGACATTGCTGCCCGTCGGCCAGATGCGCGTGCCGTG
ACAACTGCAGATGGTGCGAGCGCGTCTTACGCGGAGCTTCTGGATCTGATTGATCGT
ACTGCCGCCGGCTTACGTGGCCATGGAGTGGGTGCAGGTGATGTTATTGCGTGCAGT
CTGCGCAATTCCATCCGTTACGTTGCTCTCATTTTAGCCGCGGCCCGTATTGGTGCCC
GTTACGTGCCCCTGATGTCCAACTTTGATCGTGCCGATATTGCTACGGCGCTGCGGC
TTACCGGCCCGCGCATGATTGTTACGGACCATCAACGGGAATTTCCTGATCAGGCGC
CGCCACGTGTGCGTCTGGAGACGTTAGAGGCGGCCACCGCGTCCCCGCGTGAGGCC
GGCGAACGCTATGATGGTCTGTTTCGTAGTCTGTGGACTTCAGGCTCAACCGGATTC
CCGAAACAGATGGTGTGGCGTCAGGATCGCTTCCTCCGCGAACGTCGTCGCTGGCTT
GCTGACACCGGCATTACTGCAGATGATGTGTTTTTCTGCCGCCATACACTGGACGTG
GCACACGCGACCGACCTGCATGTGTTTGCAGCACTGTTGTCAGGCGCCGAACTTGTC
CTTGCCGATCCGGACGCCGCCCCCGACGTGCTGCTGCGCCAAATCGCCGAGCGGCGT
GCCACCGCCATGTCCGCCCTGCCGCGTCACTACGAGGAATATGTGCGTGCCGCAGCG
GGCCGTCCGGCCCCAGACTTGTCTCGCTTGCGCCGGCCACTGTGTGGCGGGGCGTAT
GTCTCAGCGGCCCAGCTGACTGATGCAGCCGAAGTTCTGGGCATCCACATTCGCCAG
ATTTACGGCTCGACAGAGTTCGGATTAGCCATGGGGAACATGTCGGATGTTCTGCAG
GCAGGTGTGGGCATGGTGCCAGTTGAAGGCGTCGGCGTACGTTTAGAACCGTTGGC
GGCTGATCGTCCCGACCTGGGAGAGTTAGTGTTAATTTCGGATTGCACGTCAGAGGG

ATACGTCGGCTCGGATGAAGCGAACGCACGGACGTTTCGCGGCGAAGAATTTTGGA
CCGGCGATGTGGCCCAGCGCGGTCCGGATGGGACTCTTCGGGTATTGGGCCGTGTGA
CAGAAACTCTCGCTGCCGCCGGCGGCCCACTGCTGGCGCCCGTGCTGGACGAAGAG
ATTGCTGCGGGCTGCCCAGTGCTGGAAACGGCGGCGCTGCCGGCGCACCCTGACCG
CTACAGCGATGAAGTCTTATTAGTCTTGCACCCCGATCCCGATCGCCCGGAACAGGA
GCTGCGGAAAGCAGTTGCGGAGGTCTTGGATCGTCATGGCTTGCGTGCATCGATTCG
GCTGACCGACGATATCCCGCATACGCCGGTCGGCAAGCCAGATAAGCCGGCCTTGC
GTCGTCGGTGGGAGTCTGGTGCTCTGGGTCCCGTAGGTGAATGGCACCATGGC**CTC**
GAGCACCACCACCACCACCACTGA

>CysF protein sequence
MLYEALRDIAARRPDARAVTTADGASASYAELLDLIDRTAAGLRGHGVGAGDVIACSL
RNSIRYVALILAAARIGARYVPLMSNFDRADIATALRLTGPRMIVTDHQREFPDQAPPRV
RLETLEAATASPREAGERYDGLFRSLWTSGSTGFPKQMVWRQDRFLRERRRWLADTGI
TADDVFFCRHTLDVAHATDLHVFAALLSGAELVLADPDAAPDVLLRQIAERRATAMSA
LPRHYEEYVRAAAGRPAPDLSRLRRPLCGGAYVSAAQLTDAAEVLGIHIRQIYGSTEFGL
AMGNMSDVLQAGVGMVPVEGVGVRLEPLAADRPDLGELVLISDCTSEGYVGSDEANA
RTFRGEEFWTGDVAQRGPDGTLRVLGRVTETLAAAGGPLLAPVLDEEIAAGCPVLETAA
LPAHPDRYSDEVLLVLHPDPDRPEQELRKAVAEVLDRHGLRASIRLTDDIPHTPVGKPD
KPALRRRWESGALGPVGEWHHG**LEHHHHHH**

>CysE DNA sequence
ATGGCCGAACTGTTACTTGAGGATGGGCAATCAATTCACTACCAAGAGACTGGAAA
GGGGAGTCCCGTGTTGCTTGTGCATGGGCTTGGAGCCCCCTCAGCGTTCTTGGCAGC
AACAGCCGAAGGACTGGCACGTGACCATCGAGTAGTAACCTTCGACTTACGCGGCC
ACGGTCGCACTCCTTTAGGTACTGGACCAGTAGGAATCGACAGATGTGCGGCGGATT
TGCACGCGGTGGCAGGAAAGTTGGACCTTCGGGCAGTGACTTTAGTGGGCTGGTCCC
TGGGCGCTACGGTTGCCTATCGATACCTGGAACGTTACGGCGCACAACGCGTAGCTC
GCCTGGTGAGCGTAGAGCAAAGCCCATATTTACTTTATGAGGACGGGTGGGAGCAC
GCCGCCTTCGGTCGCTTGACAGCTGCTGATGCTGAGACCGTTCGCCAGAACTTAGCG
GGCACAGATAGAGCGGTCGCGGCAGATCAAGTCGCAGGCTATTTTGCCGAAGGGAC
TGTTCCGGACCCAGACCTGTTGGCCCGTTTGGCTGACGCTGTGGCCACGTGCTCCCC
GGCTGCTCGGCAACAACTCTGGCAAGATGTTGTTCGTCAAGACTGGAGAGAACGGC
TGGCCGCACTGCCGGTACCAGTGCTGTTCGTTCATGGCGCTCGCTCCCGGATTTATC
CATCTGCCGTGGGTAGTCGCCTGGCTGACACCGTACCCGGAGCCCGGCTTGAGGTAT
TTGAGAACAGTGGTCATCTCCCGTTCCTTGAAGAGCCTGAGCGATTCCAACGTACAA
TACGCTCATGGGTTGCCCGC**CTCGAGCACCACCACCACCACCAC**TGA

>CysE protein sequence
MAELLLEDGQSIHYQETGKGSPVLLVHGLGAPSAFLAATAEGLARDHRVVTFDLRGHG
RTPLGTGPVGIDRCAADLHAVAGKLDLRAVTLVGWSLGATVAYRYLERYGAQRVARL
VSVEQSPYLLYEDGWEHAAFGRLTAADAETVRQNLAGTDRAVAADQVAGYFAEGTVP
DPDLLARLADAVATCSPAARQQLWQDVVRQDWRERLAALPVPVLFVHGARSRIYPSA
VGSRLADTVPGARLEVFENSGHLPFLEEPERFQRTIRSWVAR**LEHHHHHH**

>CysC DNA sequence
ATGATCGGACAAATCCCCACCGAGAGATTACTCACTCATGCTCGCACCTTAGCCGAA
GCGGGTGGGGCTCCCGTACCCCCGGCAGCACGACAACTCGCCGACCTTGCTGTTACT
GGTCCTGCAGAGATCTACGCATTAACCGGTGCAGCTGCTCGTAGCGCTGGAAGCGTT
TTAATGTCGTCCGGCGGAACTACGGGCCGTCCGAAGTTAACATACGTCCCGCATGGA
ATGGGATTGACTCGGCTGTTGGAACACTGGCGTCCGCTCCGCCCTGGTAACGTACTT
CTGAACTTATTCAATGCGGGTAGAATGTGGGGCTCTCACTACTATATGCAAACATTG
GCCGAGCGTAGTGGTTGCACAGTTATACCGTCAGGTCCTTATTCACCTGCGGAAGTA
GCCGGTTGGCTGCCTATGTTAGCCGAGGTTGGAGTGGACGCCTTGGCGGGCACGCCT
ACGGGCTTGGCCGACTTCGCACAGGGTGTTATAGACGCAGGTGGCACTCTGCCAGTG
CGAACAGTTATTTGGATGGCCGAGCCTTGGACAGGTAACAAGCGTGAGCTCGTGCG
TCAGGCGTTCCCCGAGGCTGGTTTATGGGGTAACTACGGATCAGTAGAGACGTGGGT
GATGTGCACTAACCAACCCGGATGCGACGAGACTACCCTGCACTTGTTACCGGACC
AGGTTATGGAGCCCGACGAGGATGGGGCTTTGTTAAGTCGCGTTGGCGAGGGCTGG
ACTGTTCCTGTGGTTCGGTACCGTCTGGGCGATCGAGTCGCGCCTGTAGAGTGTCGA
TGTGGTCGCCCTGATGCGCTGCGTGTCTTAGGGCGCGCCGACGACTCAGTCACACTC
AGATCCGCCTTGTTTAAGGTCTCTGAACTGGTGGATCTGGTAAGAGGGGAGCCAGG
GGTCTTAGAGGCTCAATTGCGGCTTACACGTAGCGCCGACTCCCCGAAGGCTGCAA
GTGCCTTGACCTTAGAGTTCACGGGGACAGCAGACGCAGAGGCGGTTCGGGGCCGC
CTTATAGGCGGGTTCTATCACCTCGCCGCAGTTGCCCGTCAATATCCAGACGCACTT
CGAGCCCGCAGAGTTGAGCGTCTGACACGCATTGAACGCACAAATAAGGTGCCGGC
TATGGTATGGCAACAAGCAGAGACAGACGGCGCGGGCGTTGCC**CTCGAGCACCAC**
CACCACCACCACTGA

>CysC protein sequence
MIGQIPTERLLTHARTLAEAGGAPVPPAARQLADLAVTGPAEIYALTGAAARSAGSVLM
SSGGTTGRPKLTYVPHGMGLTRLLEHWRPLRPGNVLLNLFNAGRMWGSHYYMQTLAE
RSGCTVIPSGPYSPAEVAGWLPMLAEVGVDALAGTPTGLADFAQGVIDAGGTLPVRTVI
WMAEPWTGNKRELVRQAFPEAGLWGNYGSVETWVMCTNQPGCDETTLHLLPDQVME
PDEDGALLSRVGEGWTVPVVRYRLGDRVAPVECRCGRPDALRVLGRADDSVTLRSALF
KVSELVDLVRGEPGVLEAQLRLTRSADSPKAASALTLEFTGTADAEAVRGRLIGGFYHL
AAVARQYPDALRARRVERLTRIERTNKVPAMVWQQAETDGAGVA**LEHHHHHH**

>CysD DNA sequence
ATGACAACCGCCCCAGCCGCACCGACCCGCCGTTTATGTTGGATCTACCCTGATCGA
GGAACCGACCGCATGAGAGACAAGGAGTGGAAGCATGTTTGGGGCATTTATCGTGA
AGTAGCCAGAGAAGGCGGTTGGGCTCTGTCTCTGCATAAACCAGAAGAAGTTGCAG
TAGATGTATGCGGCCCTGGCGGGCCTAAAGTTTACCTCGACGGAGAACGCGTTACAC
CTGAGGACACCGTTTTCGTGACGAGTTTGTGGTCCCTTCCTCATCACACTGTGGATGT
ATGTAATCAACTCTATCTCTATACAATATTAGAGCAGGCCGGCTTTTACCTTCCAATT
CCCCCGCATCTCAGCTTCATTGCAAATGATAAAGCAGCGACAATGCTTCACTTGGCG
GATAGCCCACTCCCGCAAGTCCCTACGGTTCGGATTGGAACTGGCCGCGACACTGCC
CAACGGTCATACGAGGCAGCACTGGCTTCGTTATCCTATCCTGCAATCGTGAAGCCA
GCGTATTGGGGTATGGGAATGGGTGTATGCTTGGTGAGAAACGCTGAGGAACTTAA
GGGCGTTGCCGGCATTGCCTCAGGCGCCGATACCGCTCTTGTGTGTCAACCATACCT
CGGCGAAGGAATTAATGACTTTCGGGTTTGGGTAGTTGGTGGAAAGCCACATACGG

TATTACGCCGAATACCTAAAGGTGCTTCCCTCACGGCGAACCTCAGTTCTGGCGGTG
GAATGGAGCACGTGCCATTACCGCCAGAATTAGCCGAGACTGTCGACTACGTAGCT
GCTCGAATGCCTATGCCTTATATTGCTGTTGATTTTCTTTGGGACGGGGAACGGTTTT
GGCTTTCCGAAGTCGAGCCAGATGGTGCCGTAGGATTCGCCGACTCAGAGCAGACG
GAGCGTGAACAGCGGAAGGTAATAGCCGATAGATTTGCGGCCTATGCCGATGCCCA
CCGCCAATTCTTGAATCGTAAAGATGCGATTCGA**CTCGAGCACCACCACCACCAC**
CACTGA

>CysD protein sequence
MTTAPAAPTRRLCWIYPDRGTDRMRDKEWKHVWGIYREVAREGGWALSLHKPEEVAV
DVCGPGGPKVYLDGERVTPEDTVFVTSLWSLPHHTVDVCNQLYLYTILEQAGFYLPIPP
HLSFIANDKAATMLHLADSPLPQVPTVRIGTGRDTAQRSYEAALASLSYPAIVKPAYWG
MGMGVCLVRNAEELKGVAGIASGADTALVCQPYLGEGINDFRVWVVGGKPHTVLRRIP
KGASLTANLSSGGGMEHVPLPPELAETVDYVAARMPMPYIAVDFLWDGERFWLSEVEP
DGAVGFADSEQTEREQRKVIADRFAAYADAHRQFLNRKDAIR**LEHHHHHH**

bel pathway

>BelI DNA sequence
ATGGCTCAAACGTTTGAGATTAAGGGCAATGATCTTTGGGATCCTACAACCTTTGAC
GCGCTCCGTCGTCAACTCATTCCATCATTTGACCTCATTTACGAAGCCGCTGTGAGA
ACGGTAGCCGCTACAGTCCCGACCGCGCCGCGTGTACTGGACTTAGGTGCCGGTACC
GGTCTCTTATCTGCAGCAATTCTTCGAGAACTGCCCGACTCTGAGGTAGTCCTGGTT
GACCGCTCTGAACTTATGCTCACGCAAGCCCGCGGGAGATTCGCTAGCCAAGACGG
TGTTACGGTCCAAACCGGCGATCTCACGGACCCTCTGCCTGAGGGTGGCTTCGATGC
CGTGGTGAGCGGCTTGGCGATTCATCACTTGAGCCACACTGGCAAGAGAGATCTGTT
CCGCCGCATCCGCGAAGCGCTGCGGCCGGGTGGGGTTTTCGTCAACGTTGAGCAAG
TGCAGGGACCGTTGCCCCACTTAGAGAGCCTCTACGACTCACAACACGAGCTTCACG
TAATTCGCGAGCAGGCACCTGCCCACGAGTGGGCAGCGGGGCGTGAACGGATGAAG
TTCGATGTCTGTATCGATCTTGAAACCCAACTCCAATGGCTGCGAGACGCAGGCTTC
CGAAGCGTGGACTGCTTAGCTAAGGATTTCCGTTTCGCTACCTATGCAGGGTGGGTG
TCC**CTCGAGCACCACCACCACCACCAC**TGA

>BelI protein sequence
MAQTFEIKGNDLWDPTTFDALRRQLIPSFDLIYEAAVRTVAATVPTAPRVLDLGAGTGL
LSAAILRELPDSEVVLVDRSELMLTQARGRFASQDGVTVQTGDLTDPLPEGGFDAVVSG
LAIHHLSHTGKRDLFRRIREALRPGGVFVNVEQVQGPLPHLESLYDSQHELHVIREQAPA
HEWAAGRERMKFDVCIDLETQLQWLRDAGFRSVDCLAKDFRFATYAGWVS**LEHHHH**
HH

>BelH DNA sequence
ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGG
CAGCCATATGACGGCTCTTCATGCTGCTGTTCACGAAATCGCACGCCGTCGTCCAGA
TGCGATAGCCGTTGAGACTACTGCGGGAGAGAGAACAACGTATGCTGAACTTTTAG
CTCGCGCTGATAGAATTGCCGCGGGTTTGAGAGCACGCGGGGTTACAGAAGGACGA

GTTGTCGTTTGCTCAGGGCTGGCCAACGATGCTAGTTACTTGGCTTTTCTCTTAGGGC
TGTGTGCCAACGGGGCCGCTTATGTCCCACTTCTGGCAGACTTCGACGCTACCGCTG
TCGACAGAGCATTACGGATGACTCGCCCTGTCTTGTGGGTCGGCCCAGATAATCATC
ACCGAGCAGGGGTTACATTGCCTCGCGTTGAATTAGCCGATCTGGAAACTCCCGCGC
CAGCAACGGCTCCTGCCGCTGGTGGTCGAGCACTCGCCCCCGGAACATTTCGTATGC
TCTGGACGTCGGGGAGTACAAAAGCTCCCAAATTGGTGACGTGGCGCCAGGAGCCC
TTCGTTCGGGAGCGTCGTAGATGGATCGCACATATAGAGGCGACTGAACGGGACGC
ATTCTTCTGTAGACATACATTAGACGTCGCCCATGCTACTGATTTACACGCCTTTGCA
GCCCTTCTTGCAGGAGCCCGGCTTATTTTGGCCGACCCAGCGGCAGACCCTGCCACG
CTGTTGGCCCAATTGGCCGCGACAGGTGCTACTTATACAAGTATGCTTCCTAACCAT
TACGAGGACTTAATTGCCGCCGCGAGACAGCGCCCTGGGACGGATTTGTCTCGCTTA
CGGCGGCCAATGTGTGGTGGAGCGTATGCTAGTCCAGCTCTCATAGCAGACGCTGCT
GATGTTTTAGGAATCCACATTCGCCATATATATGGGTCTACTGAATTTGGTCTGGCG
CTCGGTAATATGGCGGATGAAGTGCAAACTGTCGGTGGTATGCACGAGGTTGCTGG
GGTACGGGCACGTCTGGAGCCCCTTGCTGGCTACGACGGGGATGATTTAGGCCACCT
GGTTTTAACGTCGGATTGCACGTCTGACGGTTACTTGGACGACGACGAAGCGAATGC
CGCCACGTTCCGGGGCCCCGATTTCTGGACGGGAGACGTCGCAAGACGTCTTGATG
ATGGTTCCCTTCGCCTGTTGGGACGGGTCACAGATCTGGTTCTTACAACGGATGGAC
CCCTGGCTGCACCCCATGTAGACGAGCTCGTGGCGCGTCATTGTCCTGTCGCGGAAA
GTGTAACACTGGCAGCCGATCCCGACACACTTGGCAATCGGGTGTTGGTTGTATTAC
GCGCCGCTCCGGGGACTAGTGATGCCGATGCGGTCGGAGCTGTCGACAAGCTTTTAG
ATGCACACGGGTTGACAGGAGTTGTACTGGCATTTGACCGAATCCCTCGCACCGTGG
TTGGAAAGGCTGATCGTGCTTTGCTGCGTCGTCGGCATCTTCCTGCTCCTAGCAGCTC
ATGA

>BelH protein sequence
MGSSHHHHHHSSGLVPRGSHMTALHAAVHEIARRRPDAIAVETTAGERTTYAELLAR
ADRIAAGLRARGVTEGRVVVCSGLANDASYLAFLLGLCANGAAYVPLLADFDATAVD
RALRMTRPVLWVGPDNHHRAGVTLPRVELADLETPAPATAPAAGGRALAPGTFRMLW
TSGSTKAPKLVTWRQEPFVRERRRWIAHIEATERDAFFCRHTLDVAHATDLHAFAALLA
GARLILADPAADPATLLAQLAATGATYTSMLPNHYEDLIAAARQRPGTDLSRLRRPMCG
GAYASPALIADAADVLGIHIRHIYGSTEFGLALGNMADEVQTVGGMHEVAGVRARLEP
LAGYDGDDLGHLVLTSDCTSDGYLDDDEANAATFRGPDFWTGDVARRLDDGSLRLLG
RVTDLVLTTDGPLAAPHVDELVARHCPVAESVTLAADPDTLGNRVLVVLRAAPGTSDA
DAVGAVDKLLDAHGLTGVVLAFDRIPRTVVGKADRALLRRRHLPAPSSS

>BelR DNA sequence
ATGGTAGAGGCGATTCACCACGTTCGGCATCCCGGCACAGGCCCTACCGTCTTATTC
GTACCTGGACTGGGGATGACTTCTCATAGCTTCGCTCCTGTTGTTGCTGCTCTTGGCC
CTGACCATGACGTTGTGACCGTGGATTTACGTGGGCATGGGGACGGACCGCATCCAC
CGCTTGGCTGGACCTTGCAAGATGCGGCTGCCGACCTTGCACAAGTAATAGAATTGC
TTGACCTTCGAGACGTTACGTTGGTCGGATGGTCTTTAGGCGCAACCGTTACTTACA
ATTATCTTGACCGTTACGGGACGGACAGAGTGAGCCGCTTGGTTTCAGTTGAGATGA
GCCCTCACTTACTGCGAGAGGAAGGATGGGAACACGCGGCCTTCGGCGGTCTGGAC
GCAGCAGGAGCTCTCCAAGCGACGCAGCAACAGTGGACGGACCCTGACACATACCT
TACTGCACTGATAGAGAACTGCTTTGCATCTGGCTCTGATCCGGAGCCAGCGTTACT

GCGGCCTTTGTTGACCGAGTCCCTTAAAACTAGCAGACTTGCCCTTCTCGCGCTGTG
GTCAGGCGTGCTGACGCAGGACTGGCGAGAACGCATTGGGGCTTTAACATTGCCCA
CTCTGTTGGTACACGGTGCGCGGAGTCGCATCTTTCCTACTGAGGTGGGCAAATGGC
TCTTAGGCGCGCTGCCCGATGCCCGCTTGGAGATGTTTGCCCATTCAGGACACGCCC
CTTTCTTAGAGGAGACGGACCGTTTCGTGCGGGTTCTGCGTGACTTCGTTGGTGGCA
CCGCCCGCCCA**CTCGAGCACCACCACCACCACCAC**TGA

>BelR protein sequence
MVEAIHHVRHPGTGPTVLFVPGLGMTSHSFAPVVAALGPDHDVVTVDLRGHGDGPHPP
LGWTLQDAAADLAQVIELLDLRDVTLVGWSLGATVTYNYLDRYGTDRVSRLVSVEMS
PHLLREEGWEHAAFGGLDAAGALQATQQQWTDPDTYLTALIENCFASGSDPEPALLRPL
LTESLKTSRLALLALWSGVLTQDWRERIGALTLPTLLVHGARSRIFPTEVGKWLLGALP
DARLEMFAHSGHAPFLEETDRFVRVLRDFVGGTARP**LEHHHHHH**

>BelV DNA sequence
ATGGTAGGAAAGATCGATAGAGACATCTTATTAGCTCATGCCCGTACCATCGCGTCA
AGCGGCGGTAAGGCCGTCCCGGAAGGTGCGACGGAGCTCAGTGACTTGGAAGTCTG
TGGTGCAGCAGAGATCACGGCGATGACACGAGCTGCAGCGGAGCGTGATGGCGGGG
TCCTGGTTTCATCAGGCGGAACTACGGGCACACCTAAACTTACATACTTACCTCACC
ACATGGCTTTTGATCGCCTGCTTCCTGTTTGGCGACCCCTGGGTGTGGGCGACGTTGT
CTTAAATTTGTACGACGCCGGTCGCATGTGGGGTACCCACTATTTTGTGCAGAAGCT
TGCAGAGCGTACAGAGTCTACAGTTATTCCCAGCGGGCCAATGACCGCAGAGGAAT
TAGAGTTACGTTTACCAATGCTGCGGGAAGTCGGAGTTACTGCTCTGGCCTCGGTAC
CGTCAGTTTTAGCAGACTTCGCGGAGATGGTTCTTAAGACGGGTGTTAAGCCGCCTG
TTAAGACGATTATCTGGATGGGCGAGGCATGGACCGAATCCAAGCGGGAACTCGTG
AGACAAGCATTTCCAGACGTCCAATTATGGGGAATTTATGGCTCTGTTGAAACGTGG
GCTATGGGCTCCAATCGCCCGGGTTGCGACGAAACTACATTACATCTCCTGGCAGAC
GAAATTTTCGAGTTGGAAGATGCAGGAGCTCTGCTGACACGTGCGGGAGAGGGTTG
GACCGTGCCGGTAGTTCGTTACCGATTGGGAGATCGAATCGAAGCGGCTAGATGCC
GTTGCGGAAGAGGTGACGCGTTCAGAGTGCTGGGAAGAGCCGATGATGGGTTCGGG
TTATTGGGGAACTACGTTAACATCGCTGATATAGTTGACTCTGCCCGCCAATGCCCA
GATGTGGTGGATGCACAACTGGTCCTGGTGAGAGATACGGATCTCGCGGACTCGGC
TAGAAGTATGACCGTGGAGTACACGGGTACCGCTGCCCCCGAGGCTGTTAGAACGT
GGTTAACGGGAAGCAATGTGCGACTTGGGACGGTCGATTCACAGCATCCCGAGGCA
ATCTCTGCACGGCGTGTAGACCGCCTGCGTAGAATCGCGCGTACAAACAAAACTCC
CGGAGCAGTTTGGCAAGATGCATTAGGCTCTGGCAGT**CTCGAGCACCACCACCAC**
CACCACTGA

>BelV protein sequence
MVGKIDRDILLAHARTIASSGGKAVPEGATELSDLEVCGAAEITAMTRAAAERDGGVLV
SSGGTTGTPKLTYLPHHMAFDRLLPVWRPLGVGDVVLNLYDAGRMWGTHYFVQKLAE
RTESTVIPSGPMTAEELELRLPMLREVGVTALASVPSVLADFAEMVLKTGVKPPVKTIIW
MGEAWTESKRELVRQAFPDVQLWGIYGSVETWAMGSNRPGCDETTLHLLADEIFELED
AGALLTRAGEGWTVPVVRYRLGDRIEAARCRCGRGDAFRVLGRADDGFGLLGNYVNI
ADIVDSARQCPDVVDAQLVLVRDTDLADSARSMTVEYTGTAAPEAVRTWLTGSNVRL
GTVDSQHPEAISARRVDRLRRIARTNKTPGAVWQDALGSGS**LEHHHHHH**

>BelU DNA sequence
ATGAACACGAAGACGGTGGTTCTGGTTGGTGTCCCCTGGAGTGTTCACGAGTTAGAT
GATGCAATACGAGATGCCGCAACTCTTGGTGCTTCCCTTCTGGTTGTAGACACGCCA
GAATCGCTCGCCCAGATCGGAGAGCAAACAGCCGTACGGACCCGCACTGTAAAGGC
CCTTGACCCCTTATTAATCGCGGACTGTGTGCGTGACGATGAGCCAGCCACGGTATT
AGCTATTACGGAGTTCTCAATGGAGCTTGCAGCAGCCGTTCGAGAACTGCTCGGTAT
TCCCGGAACCCCTAGTGCCGTTGAGGCTCGGGTCTTGGATAAAGCCCAGACACGCG
AAGTCCTTCGAGAACATGGGCTTACACGGGTTGGCTTTCATCGCTCAAGTCTTCTGG
CACCTGAAGACTTGTTAGGTGGGCTGGAGCCGCCCGTAGTAGTGAAACCTCGCTCTT
TCTCAGGGTCTCACGGGGTGACTTTCGTCGCAGACCGATCCGAGTTAGAGAGAGTCT
TTGAACCTTATGACCTCGCAGAAACAGACCTGGATGACCGTGATGGACGCGTCGCTC
ATCTCGATGGTGACCACCGAACACACGAAGTCATCGTAGAGGAATACGTACCCGGA
CCTGAAATTAGCGCTGAGGGTTTAGTTGTAGATGGACGCCTTACATTGTTTAGTTTA
ACAGATAAGGTGAACACAGGGATGCCCCATTTCGAAGAGGTTGGTCATCTTGTCCCC
AGTAAGTATACCCGAGAACGCTCCGCGCAGGTCGAAGAATATTTACAGGCCGTTGT
GTCTGCGTTAGGCTTTGTTACATCACCTATGCATGCTGAAATAAAATTGCTCGATGA
TCGTATCGAGTTGGTTGAAATTCATACGCGGTATCCTGGCGACAGAGTTGTGGAGCT
GCTGCAAAGTGCATACGATATTCGTCCATACGAGGCATATTTTGACGCCATGTTGAA
CGGACGGGTGCCCCAAAGACCCAGACCAACAGGGGAGCACTACGGCGTTGGTTTCT
TTAATGGGCCCACAGATGCCCCCTTCGCATGGCCGTCTTATGCATTCCCACACCCTG
AAGCCGTCGTGTCCATTGACGTAGATCGACGCCGTGCACCGAAGGTCTTCGCATACG
AGGGTCTCCGGATTCGCTACTGGAGAGCTGGCCATGCATTATTTGCCCACGAAGACC
ATGCCCGCGTCCAAGAAAATATCGCCTTCTTATTAGATAATACACCTGGGCAAGGCG
GAAGTGGAAGC**CTCGAGCACCACCACCACCACCAC**TGA

>BelU protein sequence
MNTKTVVLVGVPWSVHELDDAIRDAATLGASLLVVDTPESLAQIGEQTAVRTRTVKAL
DPLLIADCVRDDEPATVLAITEFSMELAAAVRELLGIPGTPSAVEARVLDKAQTREVLRE
HGLTRVGFHRSSLLAPEDLLGGLEPPVVVKPRSFSGSHGVTFVADRSELERVFEPYDLAE
TDLDDRDGRVAHLDGDHRTHEVIVEEYVPGPEISAEGLVVDGRLTLFSLTDKVNTGMPH
FEEVGHLVPSKYTRERSAQVEEYLQAVVSALGFVTSPMHAEIKLLDDRIELVEIHTRYPG
DRVVELLQSAYDIRPYEAYFDAMLNGRVPQRPRPTGEHYGVGFFNGPTDAPFAWPSYA
FPHPEAVVSIDVDRRRAPKVFAYEGLRIRYWRAGHALFAHEDHARVQENIAFLLDNTPG
QGGSGS**LEHHHHHH**

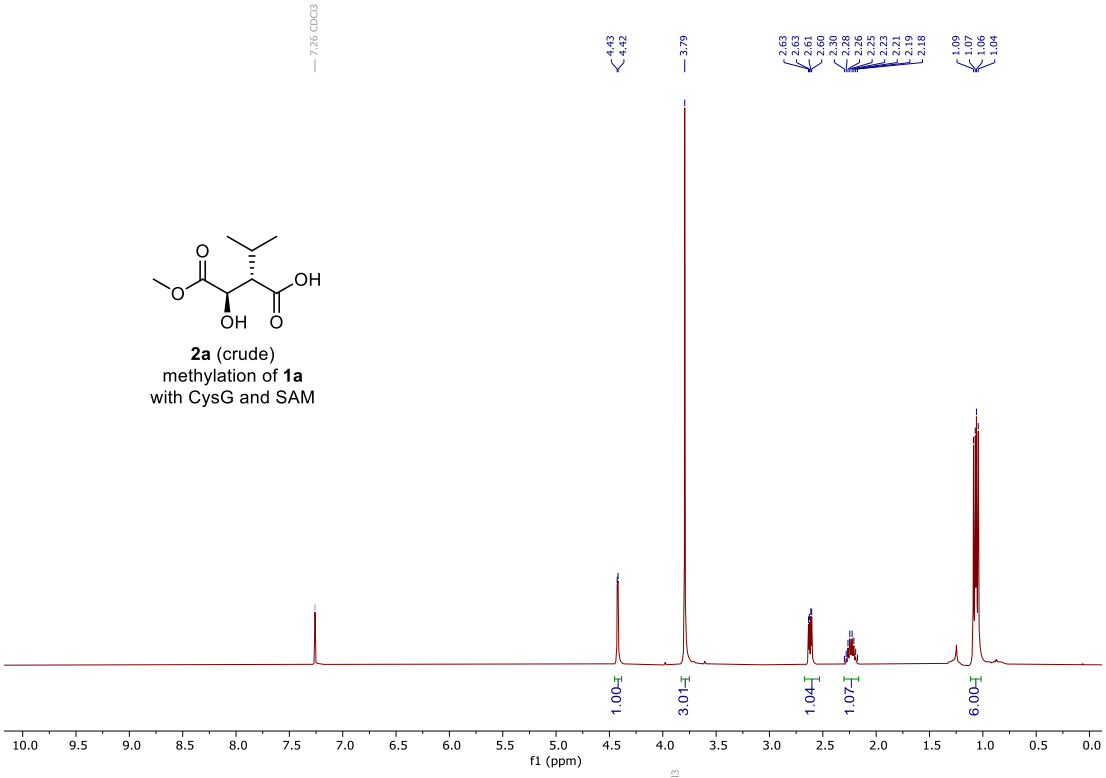
Naturally occurring Fused-CysFE

>*Bh*CysFE DNA sequence

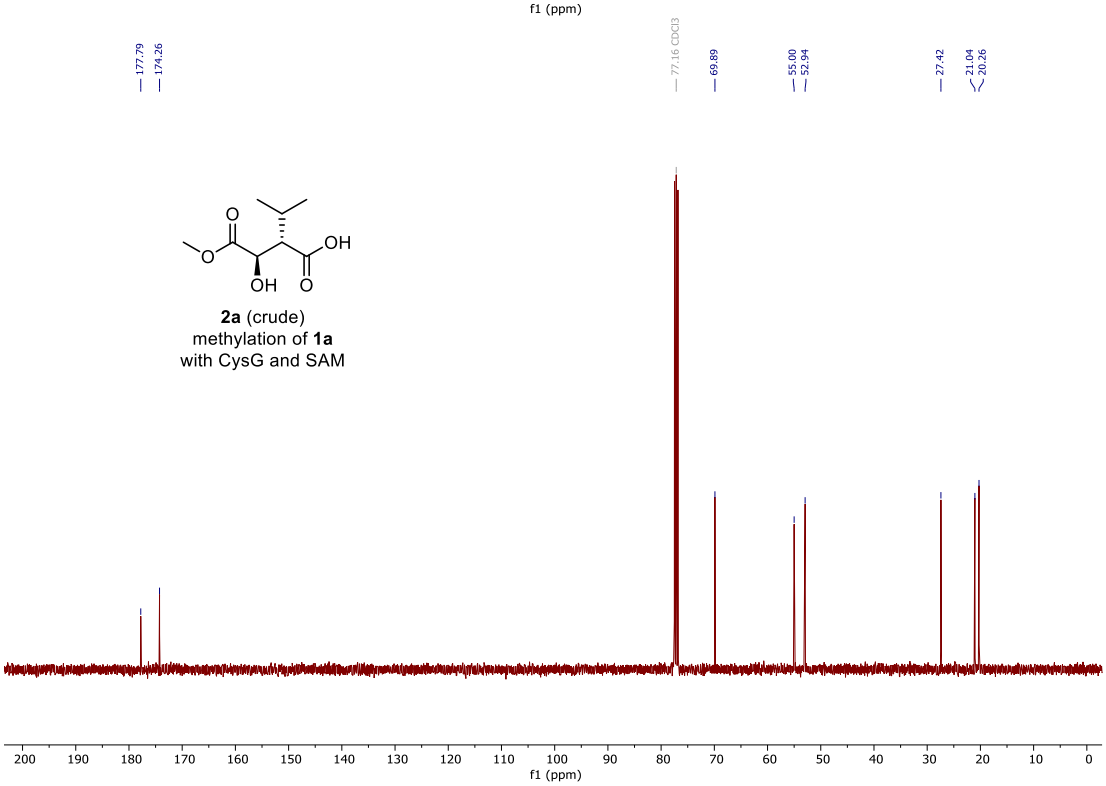
ATGGACAAAACGCGCCAGTCTATAGCTTCATCTATTCCTATCTTCGACGCTATCTCCC
AAATAGCACGTGAGAATGCAAACGGAGTTGCCATCGAGGCTTTAAGTGGAGAAACA
TGTACTTTTGGATCTTTGATCGAGCGTGCAGAAATGCAAGCACGTGGTCTCCACGCT
CTGGGAGTTGGGGCTGGACACTGCGTAGCAGTTTATTGTAAGACTAGCATCTCCTAT
GCTAGTTTGATATTAGCGGTTTGCCGTTTGGGCGCTTCCTACGTACCTATTCTGAATA
ACTTCGATTTAGAAGCACGGCGCCGGGCCTTCCAAATGGCGCAACCTGTATTGGTTG
TCCACGACGGTGTCCGTTCCTATTCTTCTTTCGGACGCCCGGCCGTAGAGATTCGTGC
CCTGATAGAGCCCACCAGCGACCGGGGCTCCCCACCACGTCCTGACGCGGAGCACG
TATTTCGCAAGTTATGGAGCTCGGGTTCCACAGGCGGCAGCAAATTGATAGGTTGGA
CTCAGGGTAAGTTGCTTAAAGAACGGCTGCGCTGGCAAAACCACGTCGGCCTTCGTG
GGTCTGACCGTTACTTCTGCAAGCACACACTGGATGTGGCGCACGCAACCGACTTGC
ACCTGTTTAGCGCGCTGCTCTCAGGCGCCACATGTATTCTGGACGACGTCCACGCTG
GAGACGCAGCATTATGGGAAACTATCGCAGACTGGCGTCCGACGGTGATGTCTGCG
CTCCCAGAGCATTACCGTGACTGGTTACGGCATTACAGAGCGAATGGTACCCGACTG
CCGGGCACTCTCCGTCTTGCAATGTGTGGTGGTACATACGTCTCACCAGAAACGGCG
GCAGATGTGGCCGATGGTCTGGGCTTTCGGCTGAGACAAATCTACGGGTCGACCGA
GTTTGGACTCGCAATGATTAGTGAGGAGAGCGCGGGAGACTTAGTGCTTGTTAACG
GAGTCGGCGCTCGTCTTGAGCCGCTCCCCCAGGGTACGGCTGGTAATCTGGGCCATC
TCATCCTCATCTCAGACTGTACATCCGAAGGGTATCTGGGAGACGCGGACGAGCAC
GCAGCCGCTTTTCGCGGTGAAGAGTATGGCACTGGAGACGTAGCAGAGATGGTTTCT
CCCAAGACTTATCGAATAGTTGGACGTACTAAGGAATTGCTTAATGTAGGTGGGCGC
ATCACCACAACAAGTATGGTTGACCGCAGAATACAAGCGGAACTTCGTCTTCAAAA
CTTCGCCACTGTTGTGGATCCCCGTTCGAGCGAATCAGTAACCGTCTTTGTTGATCAA
CCGCCGGGCACTCCTGTTGAGGAAGAGCGCCTTCAACGAAGAATTACGGCGGCCAC
CGAGCCATTAGGCTTAAAACCAAGTATCGTGTTCTTGCATCCCTTCCCCTACACTGA
GGTTGGAAAGCCGGATAAAGCAGTATTGCGACAGAAACTGGCGACTCCTGTGAATA
TGCGACTTAATTGCCGTGAGCTTGGACGAGGGTTCCCGCTGGTGATTCTGCCAGGTC
TGTGCCTGACTGACGCGATATTTGAGCCGCTTATTGACTTAATACAGGATGAGTACC
GCTTGTTACTTATCGACCTTCCAGGGCATGGTCAGAACCAGAGTCTGCCGCCGGAAG
TTACTGCACGTCCAGGCATCATTGACCGTTTCTGTGAGGGCATCTACTCCTTGTTAGC
GGAACGCGGGATCGAGAAATTTGCCGTAATGGGTGTTTCATTAGGAGCGACAGTAG
CGTATGCGTTAGCCACCGGAAGACATGCAAATCAAATTAGCGCGCTGGTCTCTATCG
AACAGACACCGTTCCTGTTAGCGGACGATGGTTGGGGTCACGCCGCGTTCGGCACTT
TGACGCGTGAAGGGGCGCAGCAGATTTTGGCCGGATTAGCAACTGATTCCGGAGAA
TTCTCCCGGCAAATCATAGCGGCTAGTCTGTTAGAGGCTACCCGTATAGACAAAGCA
CTCAAGGGGCGCATAGTTAGAAGCTCTGCCGCATGCGATCCTAGAGCTATGGCCGC
GCTGCTGGCAGATGCGCTCTCGCAGGACTGGCGTCCGGGCCTTCAAGGGGCAACTC
GGAATGTTATGCTTGTGCACGGTTCCCAATCTGCGGTGTACCCGACCAACGTCGGGT
CCTGGCTTAACGAGAACTGGGATGTGAAAGCACTTCTGCAGCTTAAGACGGGTGGT
CACCTTCCCTTCATCGATGAACCTGTTATGTTCAGTAACACCGTTAAGGCCTTTCTGA
AGTCTACATGTGGAGACCAGAATGACGAGCAT**CTCGAGCACCACCACCACCACCA**
CTGA

>*Bh*CysFE protein sequence
MDKTRQSIASSIPIFDAISQIARENANGVAIEALSGETCTFGSLIERAEMQARGLHALGVG
AGHCVAVYCKTSISYASLILAVCRLGASYVPILNNFDLEARRRAFQMAQPVLVVHDGVR
SYSSFGRPAVEIRALIEPTSDRGSPPRPDAEHVFRKLWSSGSTGGSKLIGWTQGKLLKERL
RWQNHVGLRGSDRYFCKHTLDVAHATDLHLFSALLSGATCILDDVHAGDAALWETIAD
WRPTVMSALPEHYRDWLRHYRANGTRLPGTLRLAMCGGTYVSPETAADVADGLGFRL
RQIYGSTEFGLAMISEESAGDLVLVNGVGARLEPLPQGTAGNLGHLILISDCTSEGYLGD
ADEHAAAFRGEEYGTGDVAEMVSPKTYRIVGRTKELLNVGGRITTTSMVDRRIQAELRL
QNFATVVDPRSSESVTVFVDQPPGTPVEEERLQRRITAATEPLGLKPSIVFLHPFPYTEVG
KPDKAVLRQKLATPVNMRLNCRELGRGFPLVILPGLCLTDAIFEPLIDLIQDEYRLLLIDL
PGHGQNQSLPPEVTARPGIIDRFCEGIYSLLAERGIEKFAVMGVSLGATVAYALATGRHA
NQISALVSIEQTPFLLADDGWGHAAFGTLTREGAQQILAGLATDSGEFSRQIIAASLLEAT
RIDKALKGRIVRSSAACDPRAMAALLADALSQDWRPGLQGATRNVMLVHGSQSAVYP
TNVGSWLNENWDVKALLQLKTGGHLPFIDEPVMFSNTVKAFLKSTCGDQNDEH**LEHH**
HHHH

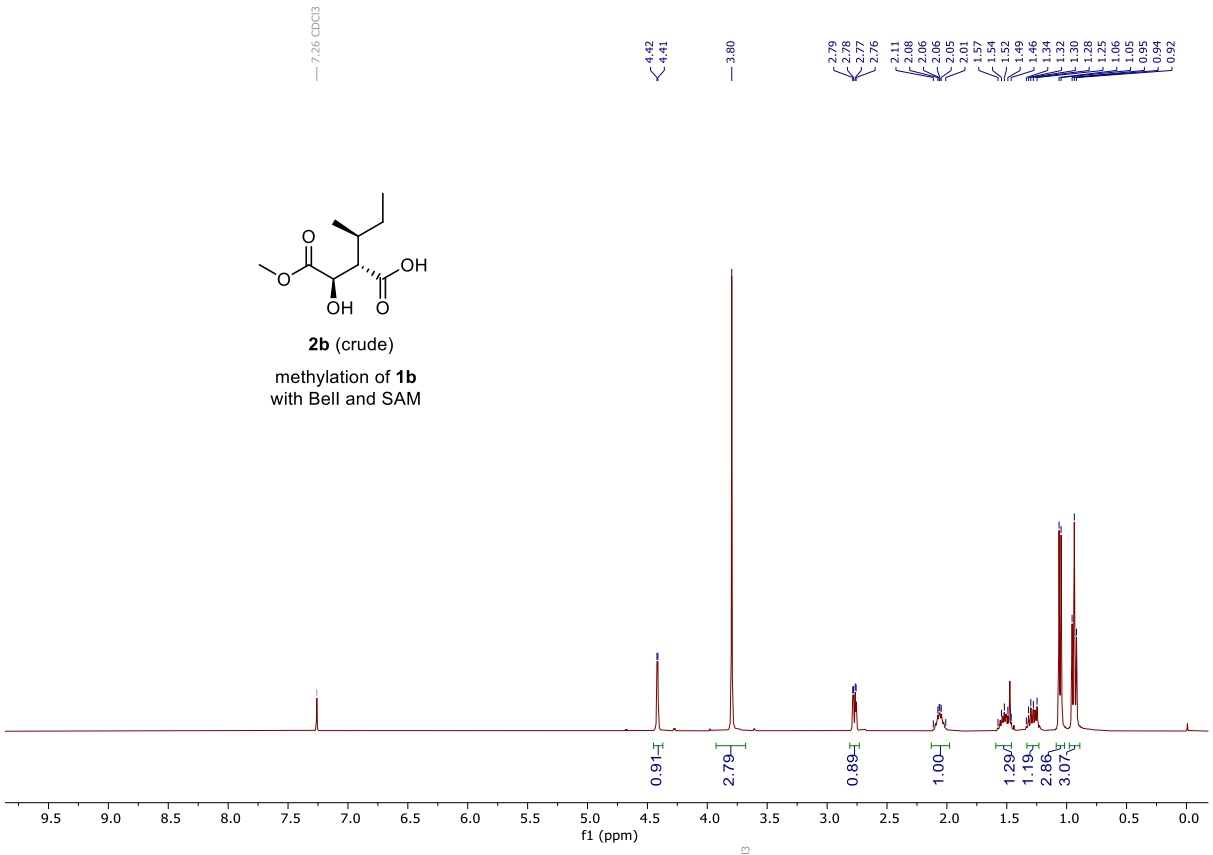
7. NMR spectra of compounds synthesised enzymatically



2a (crude)
methylation of **1a**
with CysG and SAM

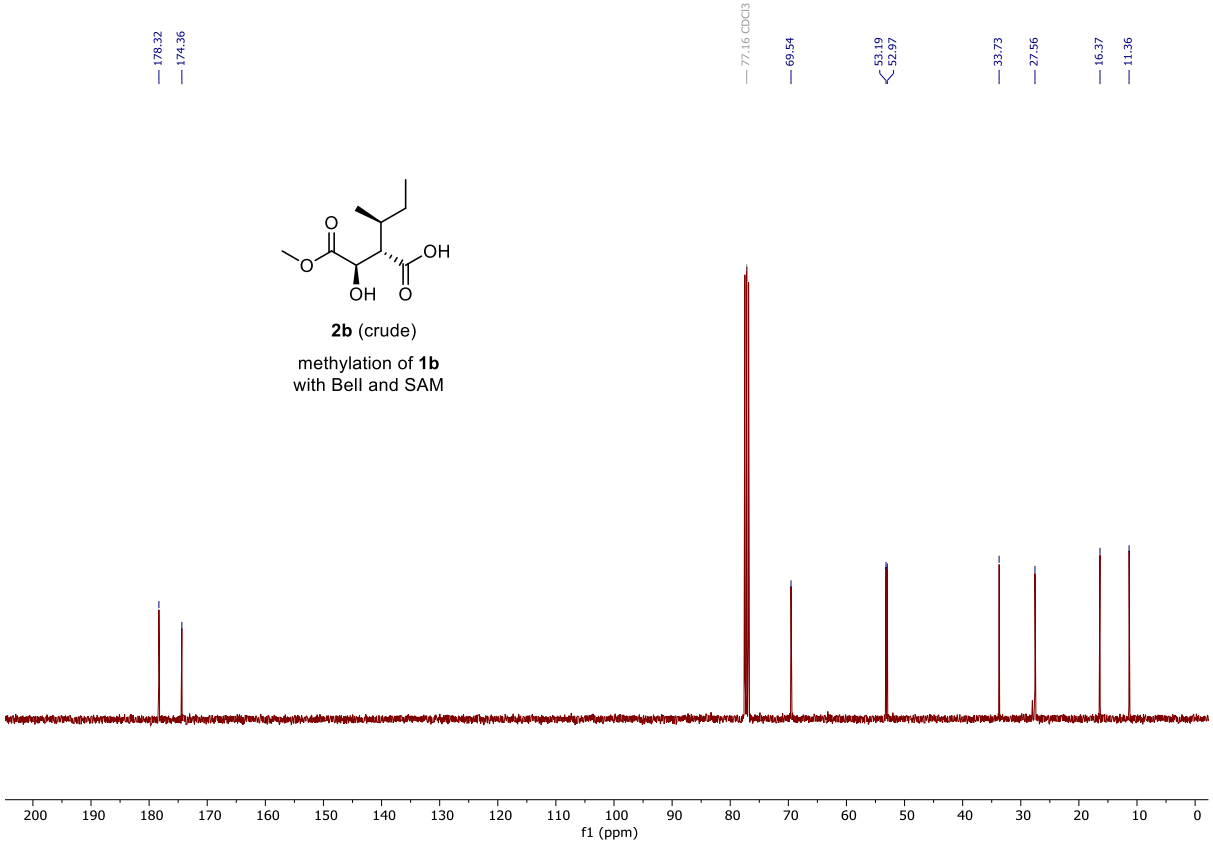


2a (crude)
methylation of **1a**
with CysG and SAM



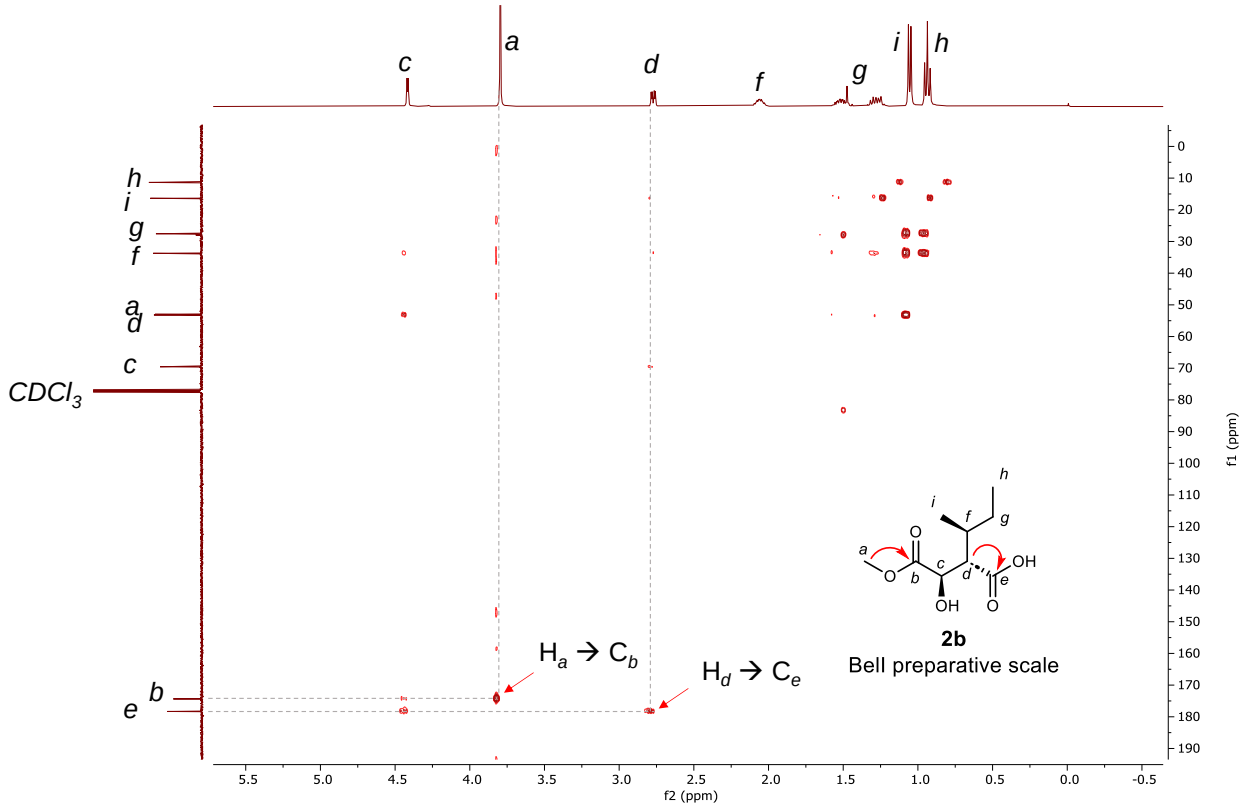
2b (crude)

methylation of **1b**
with Bell and SAM

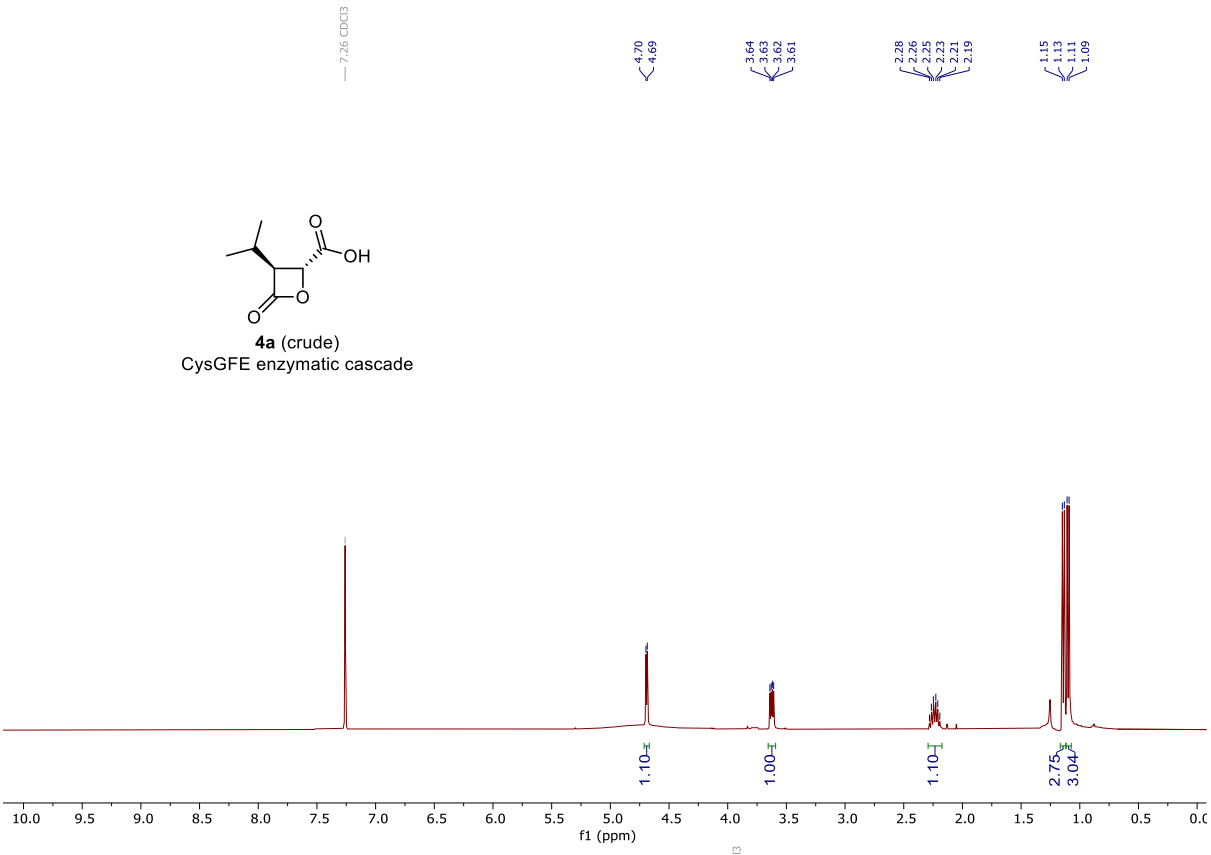


2b (crude)

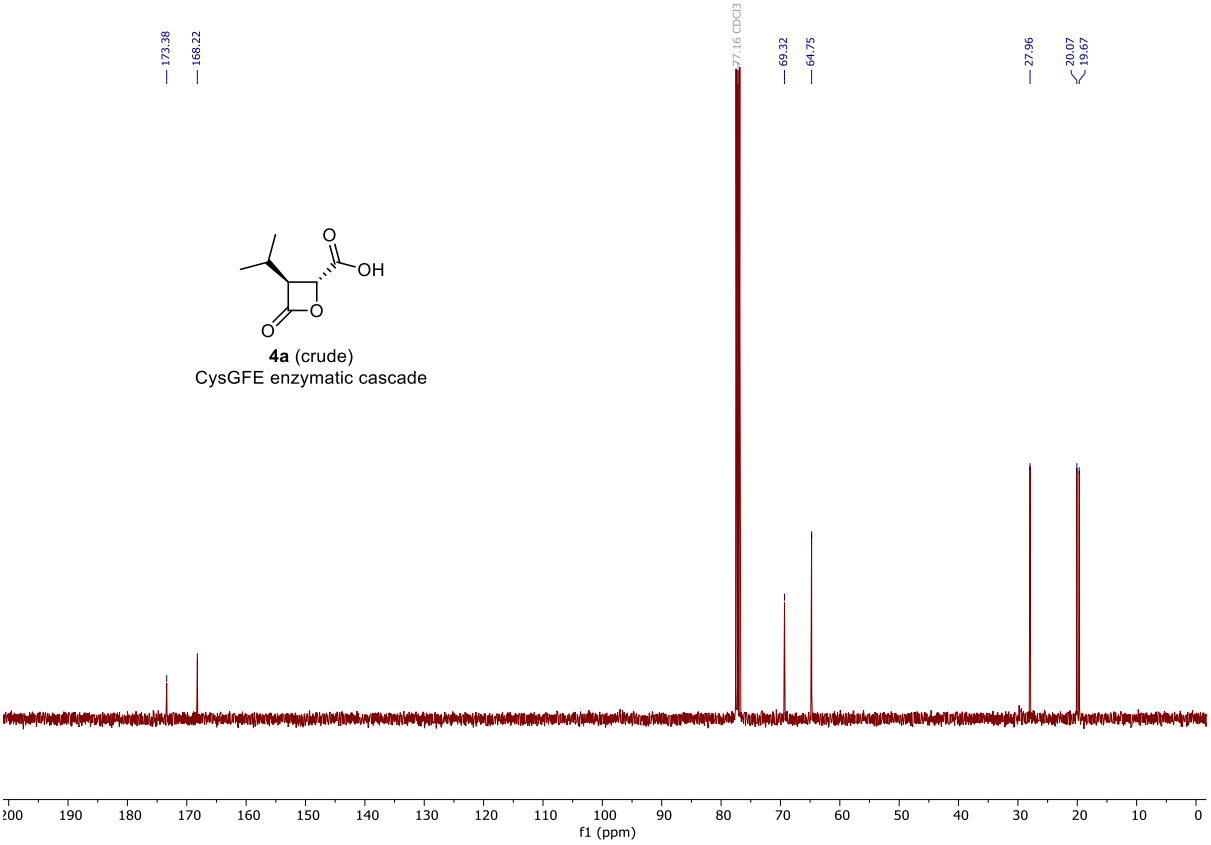
methylation of **1b**
with Bell and SAM



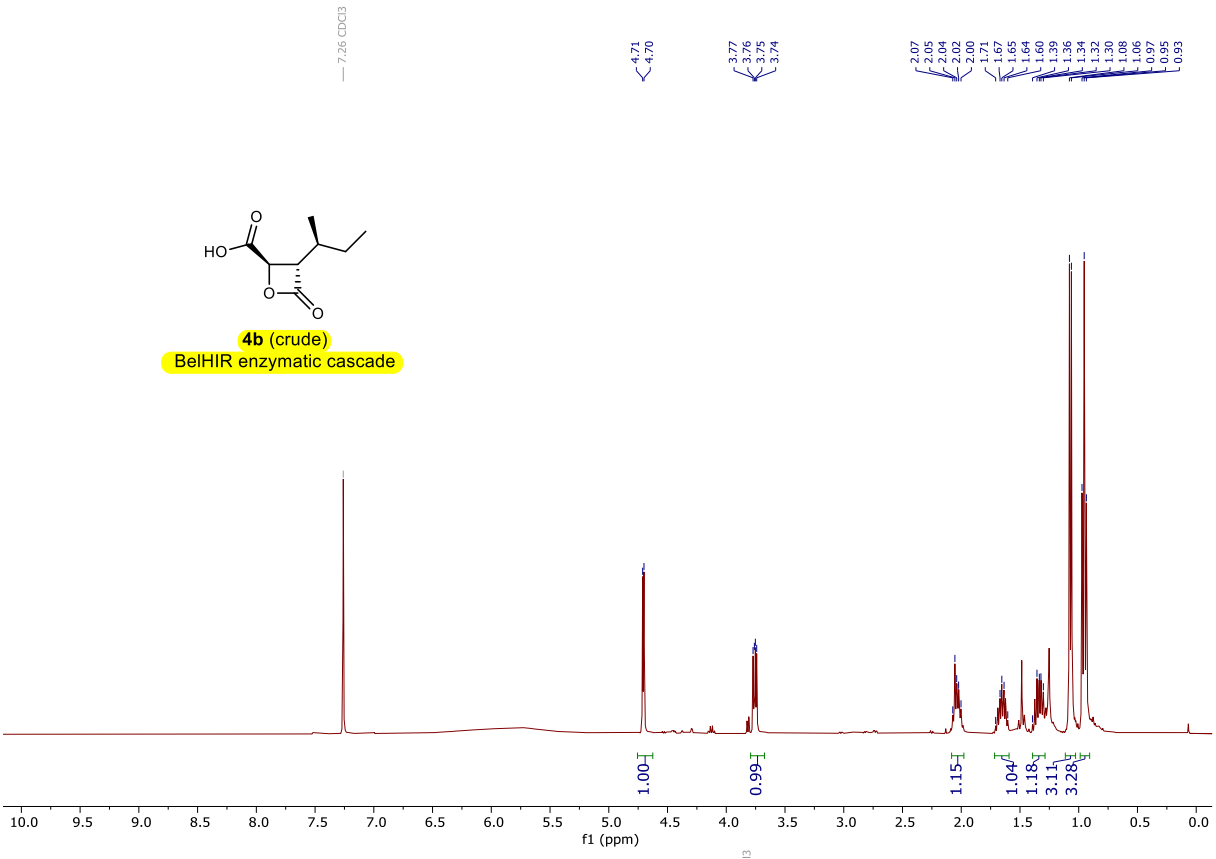
HMBC spectra of **2b** from enzymatic preparative scale reaction.

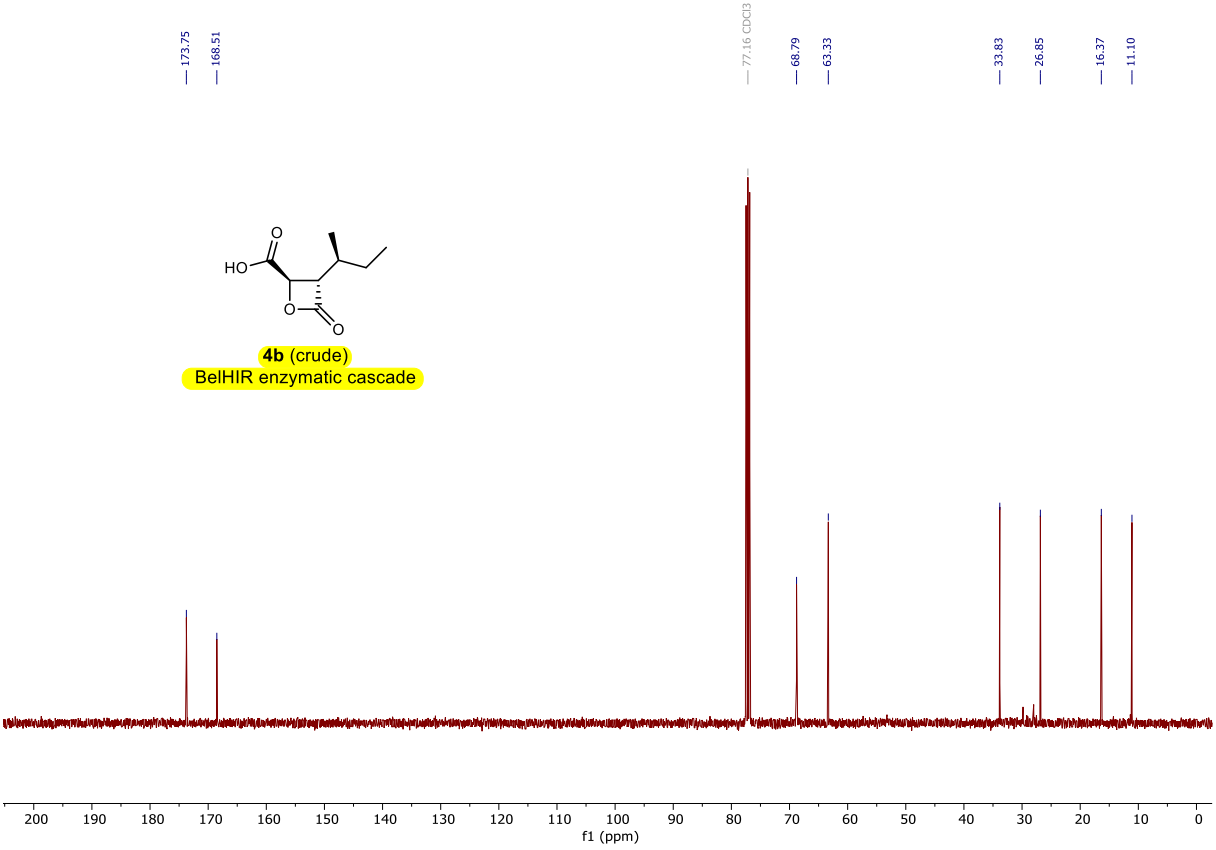


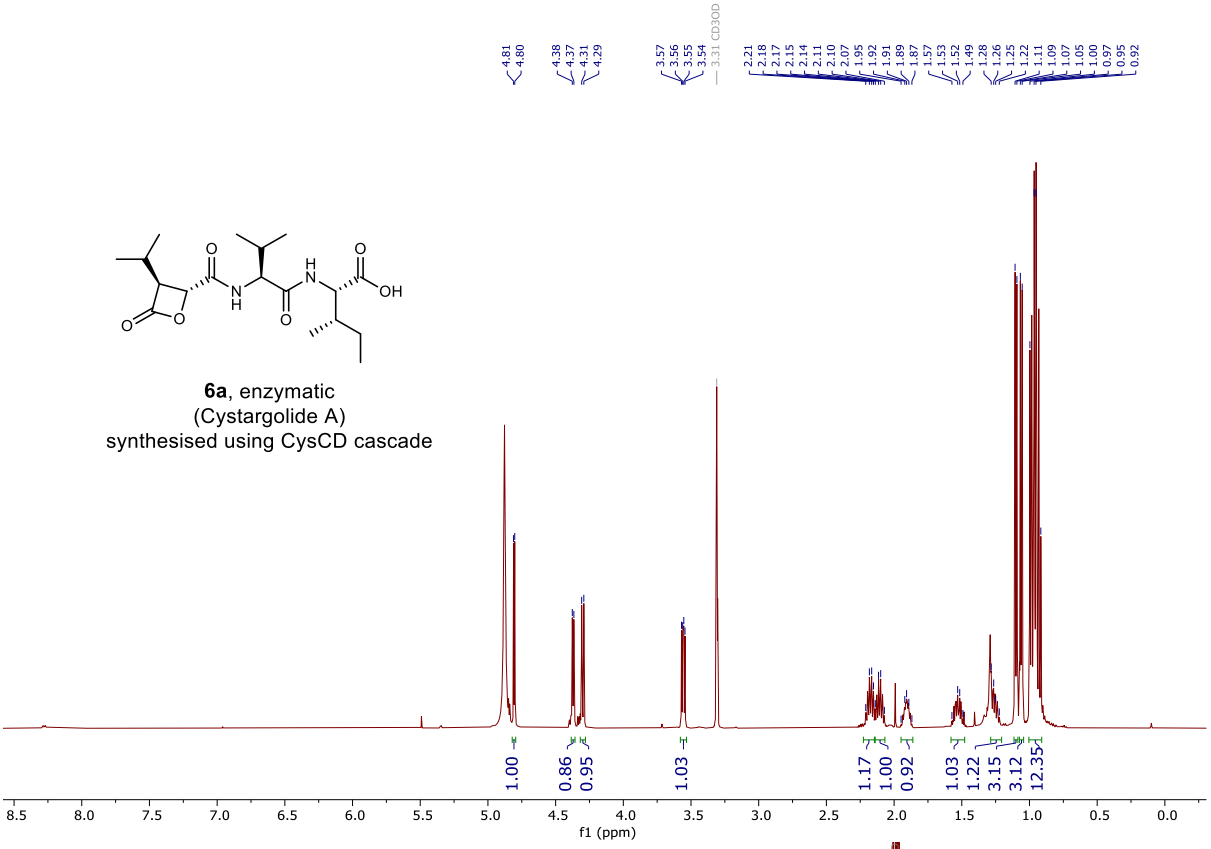
4a (crude)
CysGFE enzymatic cascade



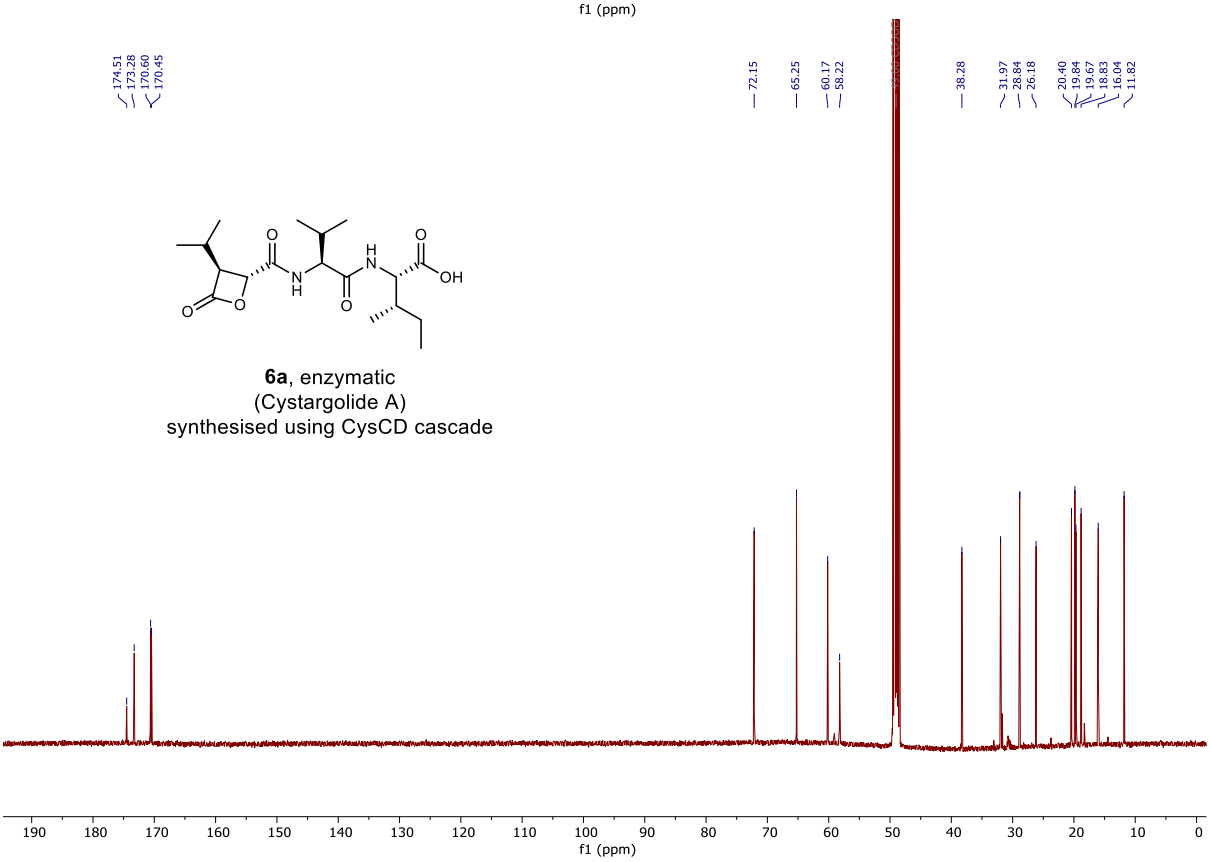
4a (crude)
CysGFE enzymatic cascade


7.26 CDCl3
4.71
4.70
3.77
3.76
3.75
3.74
2.07
2.05
2.04
2.02
2.00
1.71
1.67
1.65
1.64
1.60
1.39
1.36
1.34
1.32
1.30
1.08
1.06
0.97
0.95
0.93
HO
O
O
O
4b (crude)
BelHIR enzymatic cascade
1.00
0.99
1.15
1.04
1.18
3.11
3.28
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

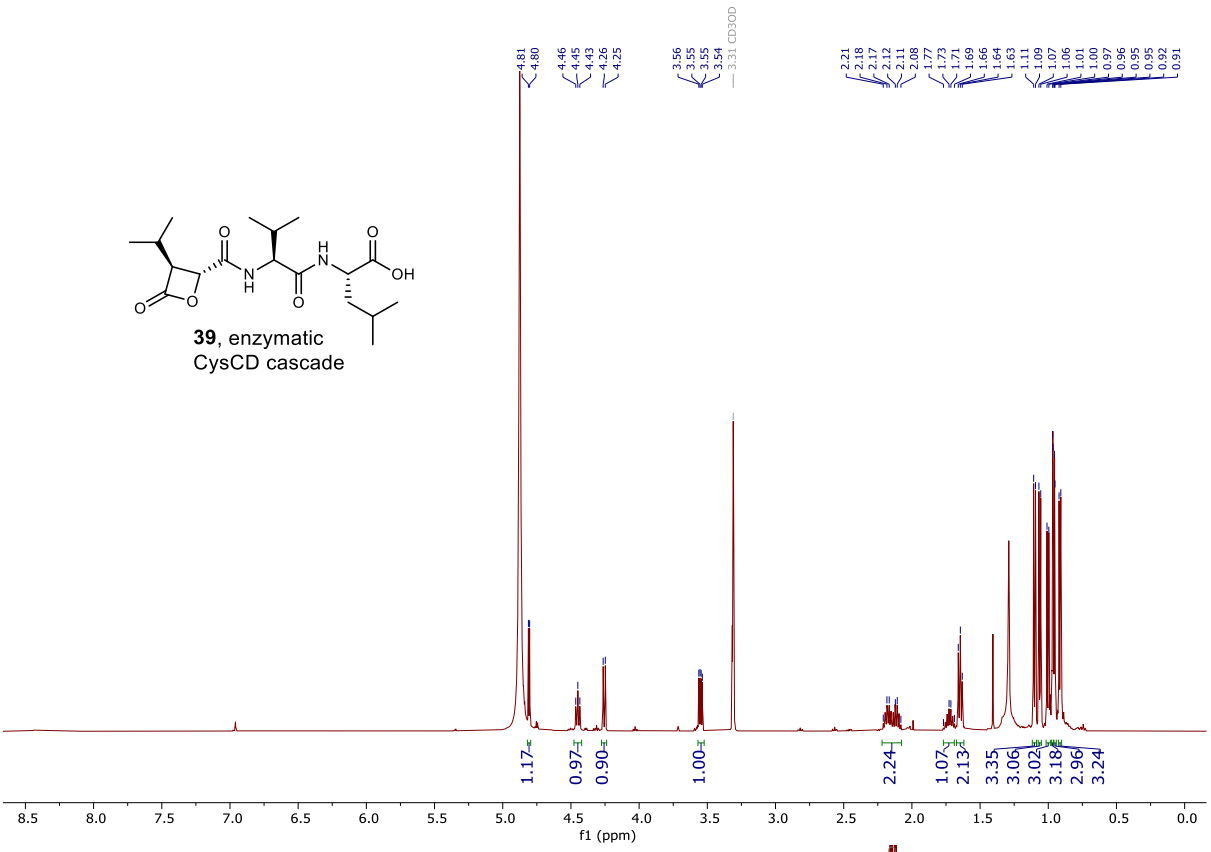

173.75
168.51
77.16 CDCl3
68.79
63.33
33.83
26.85
16.37
11.10
HO
O
O
O
4b (crude)
BelHIR enzymatic cascade
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

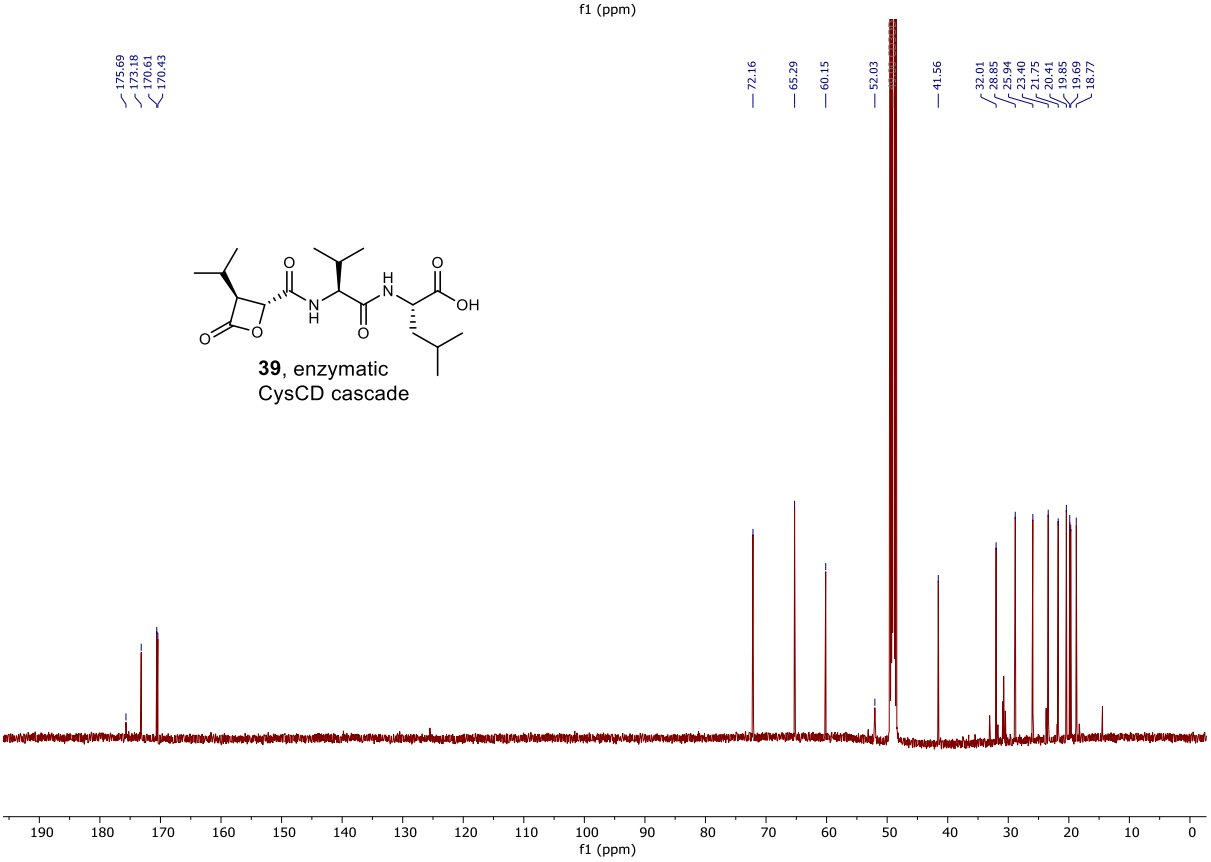


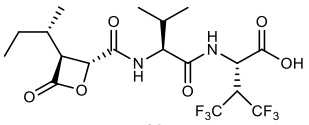
6a, enzymatic
(Cystargolide A)
synthesised using CysCD cascade



6a, enzymatic
(Cystargolide A)
synthesised using CysCD cascade

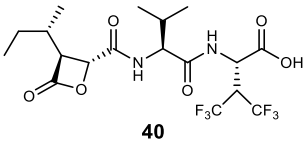

4.81
4.80
4.46
4.45
4.43
4.26
4.25
3.56
3.55
3.55
3.54
3.31 CD3OD
2.21
2.18
2.17
2.12
2.11
2.08
1.77
1.73
1.71
1.69
1.66
1.64
1.63
1.11
1.09
1.07
1.06
1.01
1.00
0.97
0.96
0.95
0.95
0.92
0.91
O
O
N
H
O
N
H
O
OH
O
O
39, enzymatic
CysCD cascade
1.17
0.97
0.90
1.00
2.24
1.07
2.13
3.35
3.06
3.02
3.18
2.96
3.24
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)


175.69
173.18
170.61
170.43
72.16
65.29
60.15
52.03
41.56
32.01
28.85
25.94
23.40
21.75
20.41
19.85
19.69
18.77
O
O
N
H
O
N
H
O
OH
O
O
39, enzymatic
CysCD cascade
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)



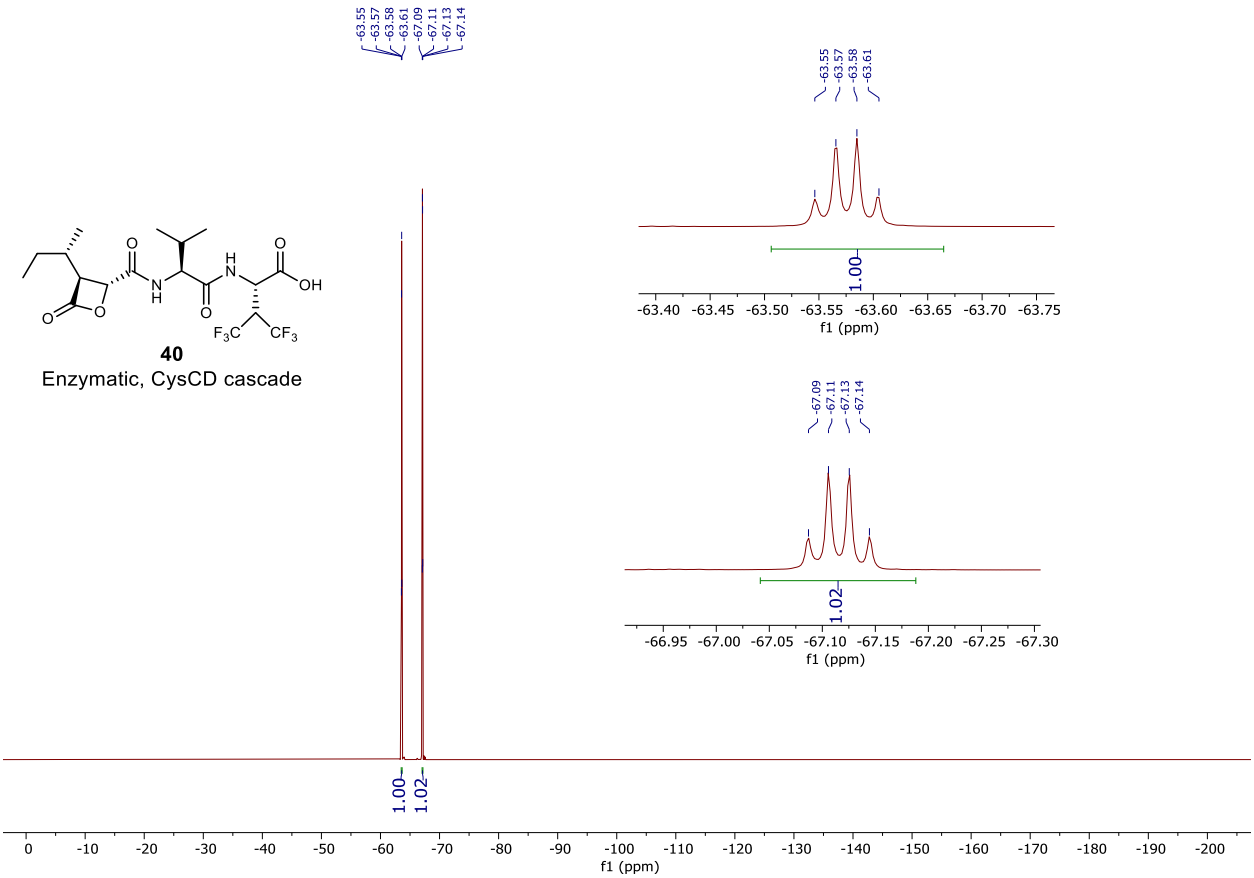
40

Enzymatic, CysCD cascade

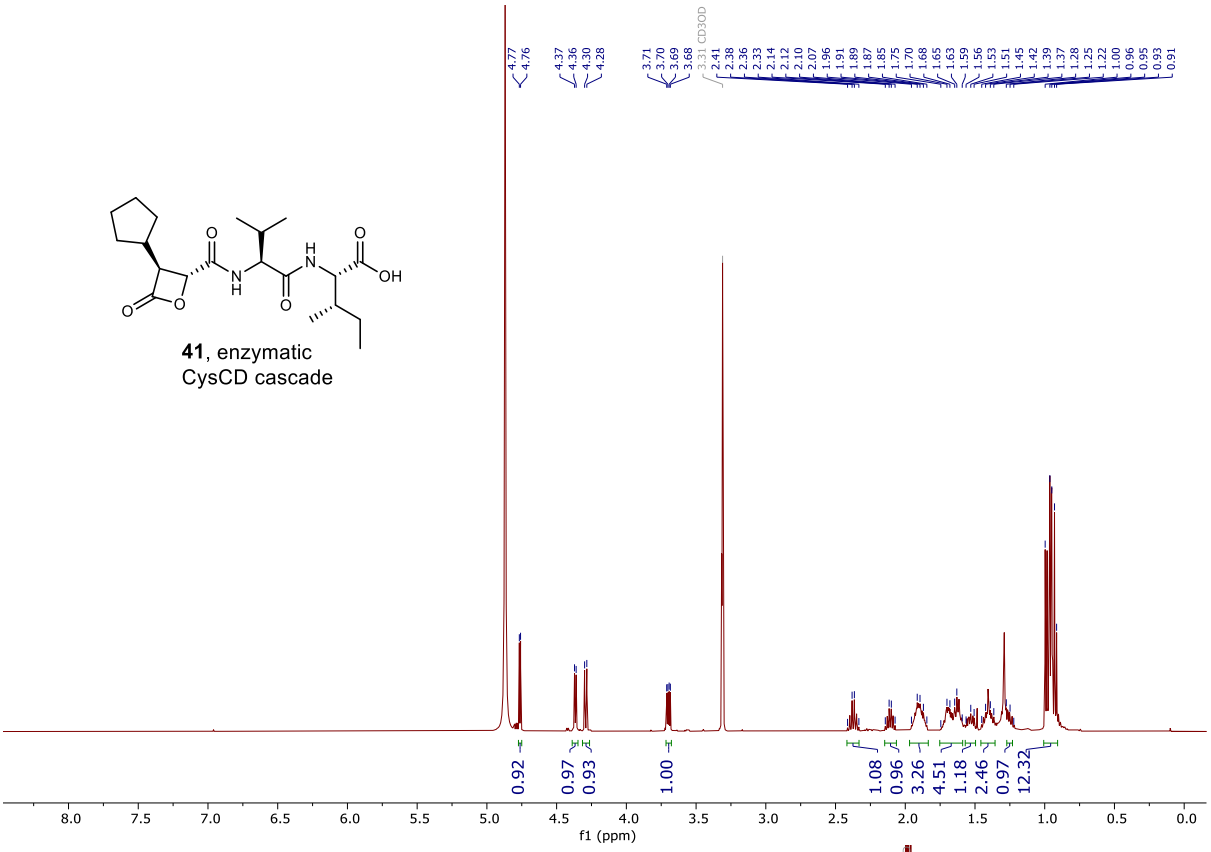


40

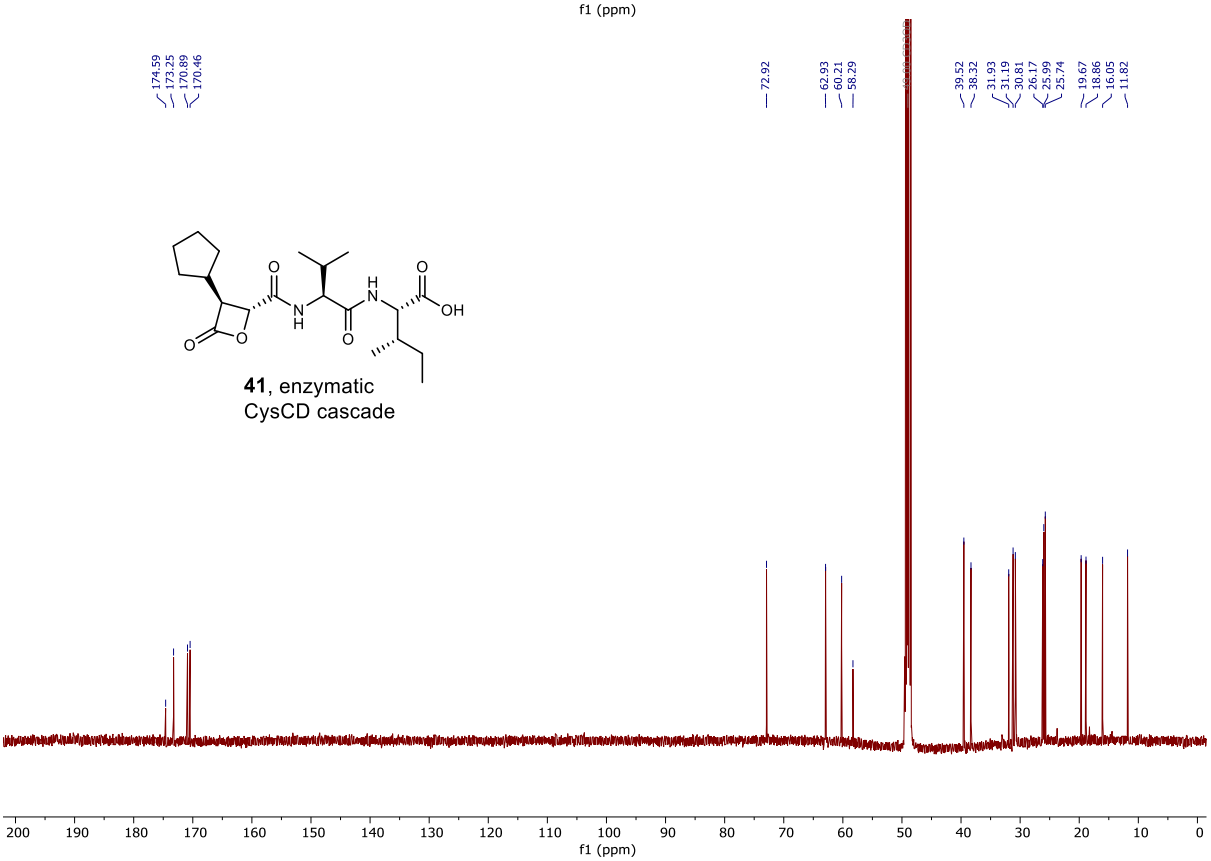
Enzymatic, CysCD cascade


-63.55
-63.57
-63.58
-63.61
-67.09
-67.11
-67.13
-67.14
1.00
1.02
f1 (ppm)
40
Enzymatic, CysCD cascade

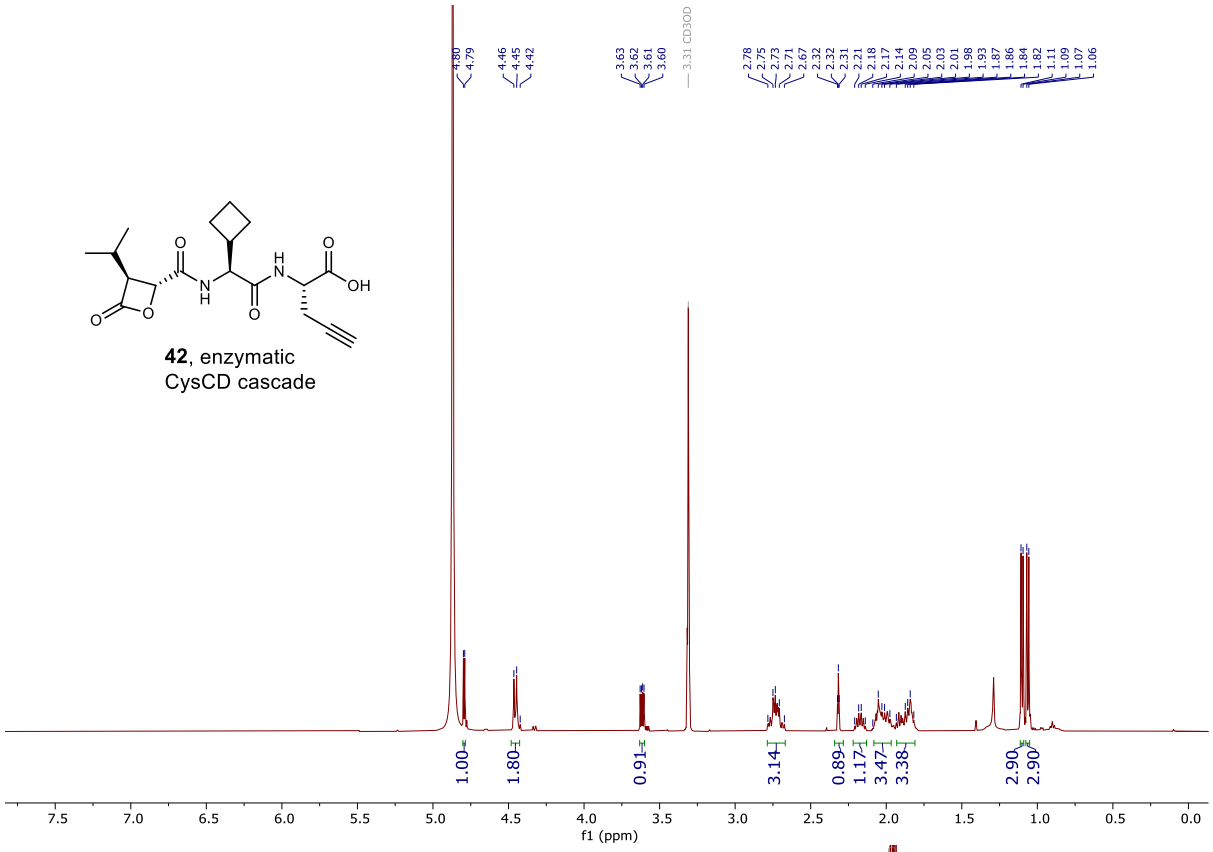
F19-NMR



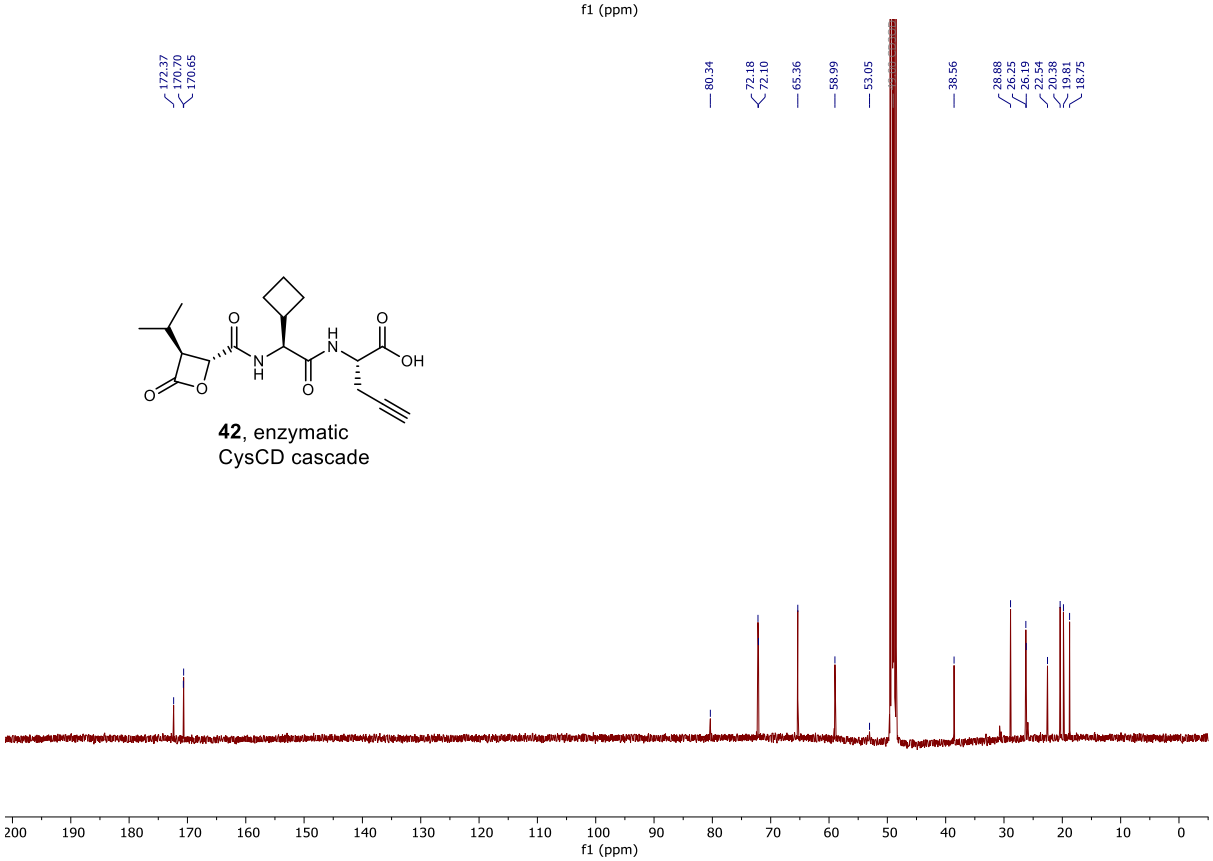
41, enzymatic CysCD cascade



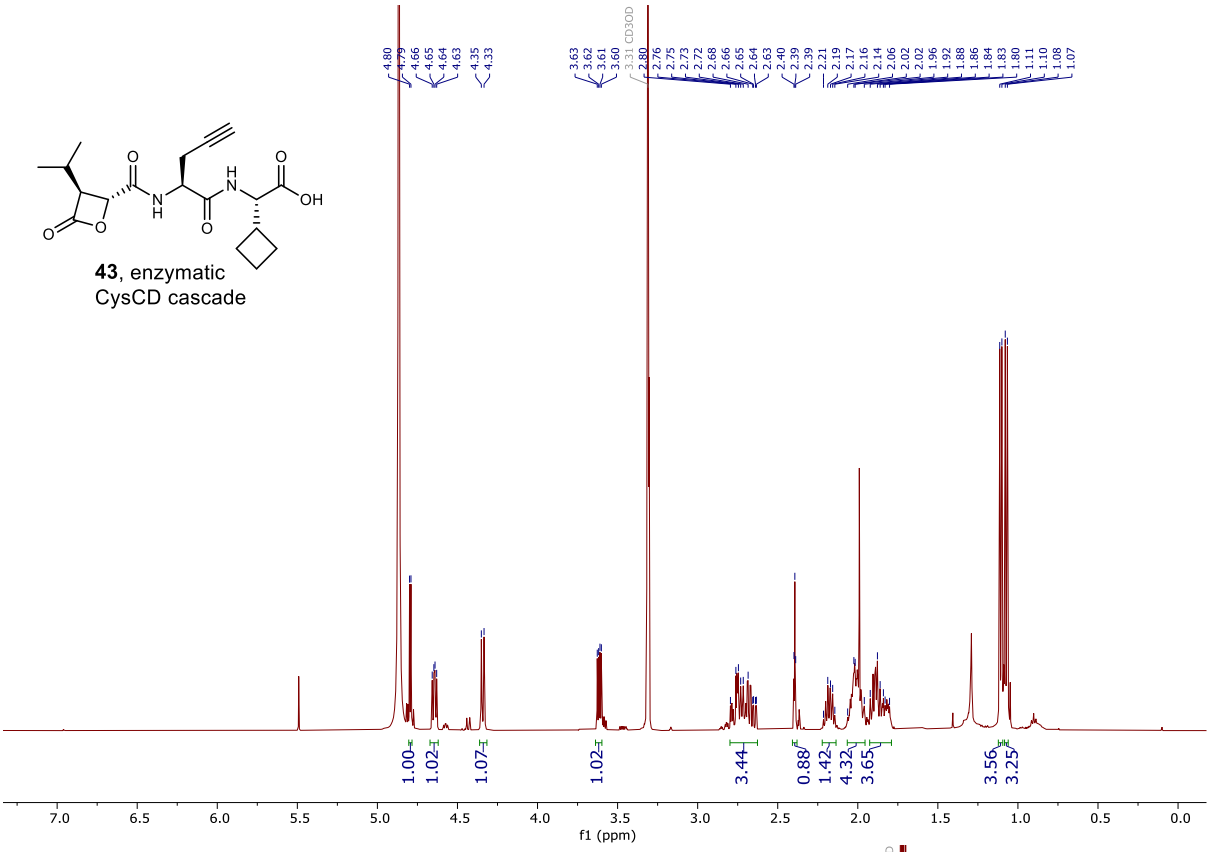
41, enzymatic CysCD cascade



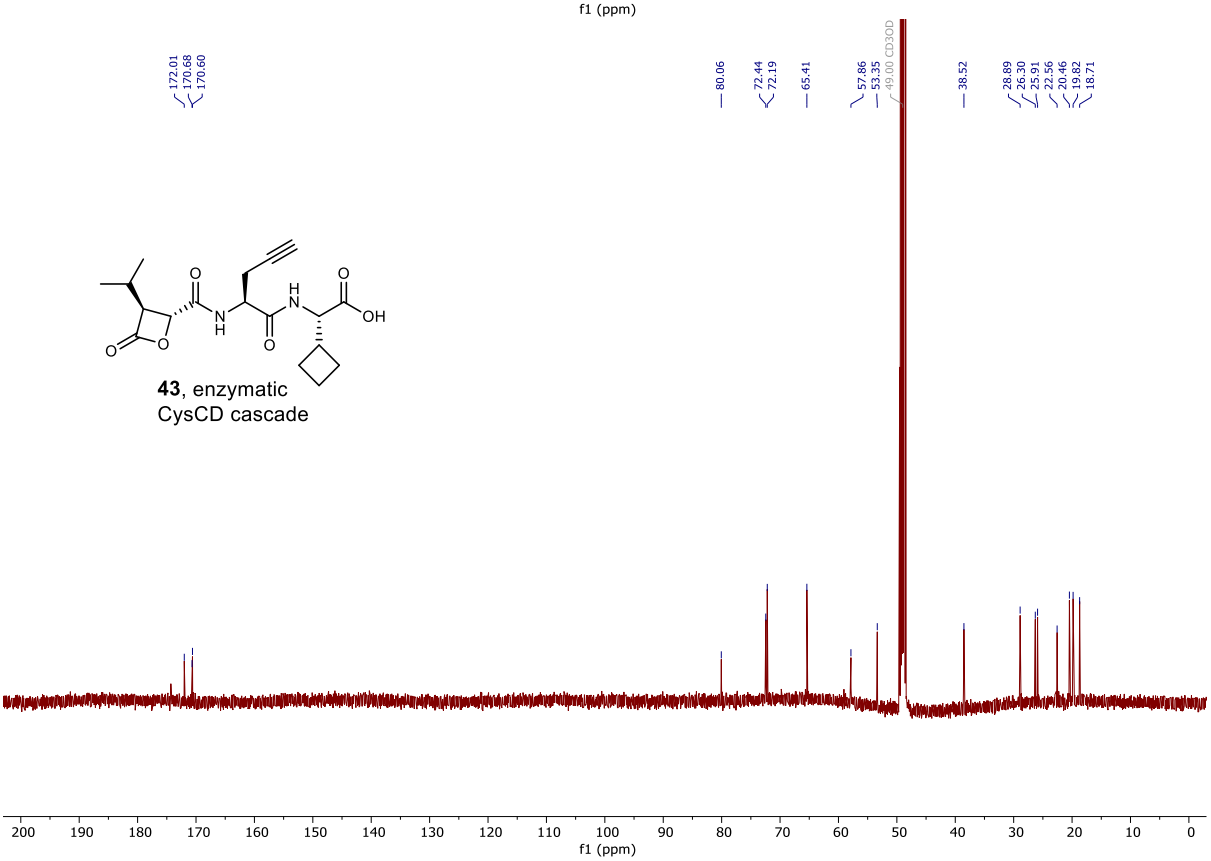
42, enzymatic
CysCD cascade



42, enzymatic
CysCD cascade

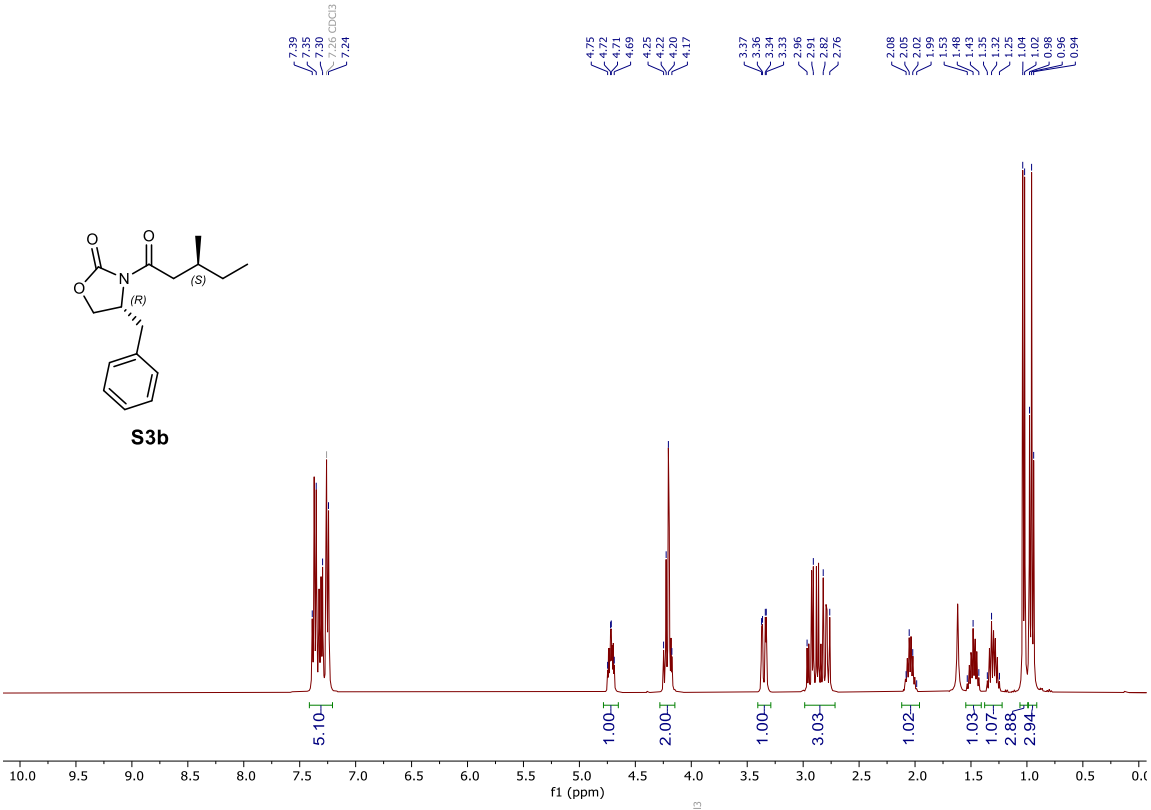


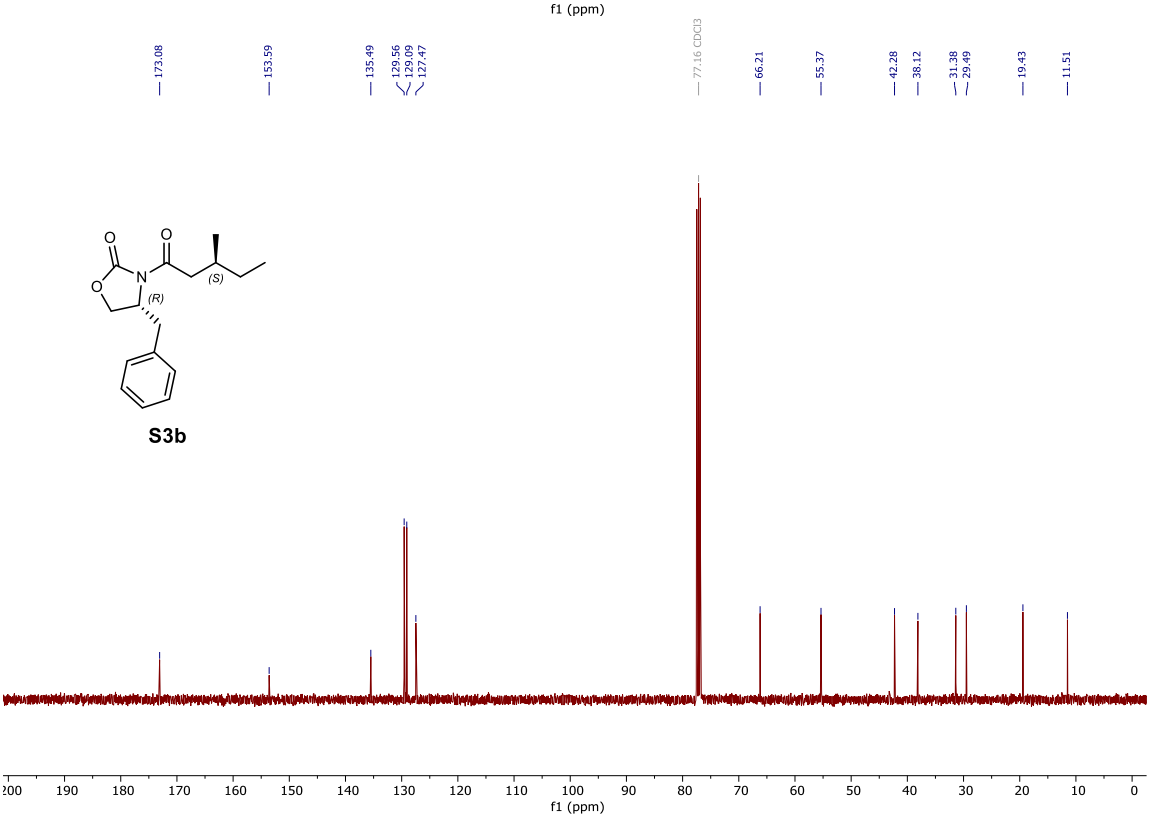
43, enzymatic
CysCD cascade

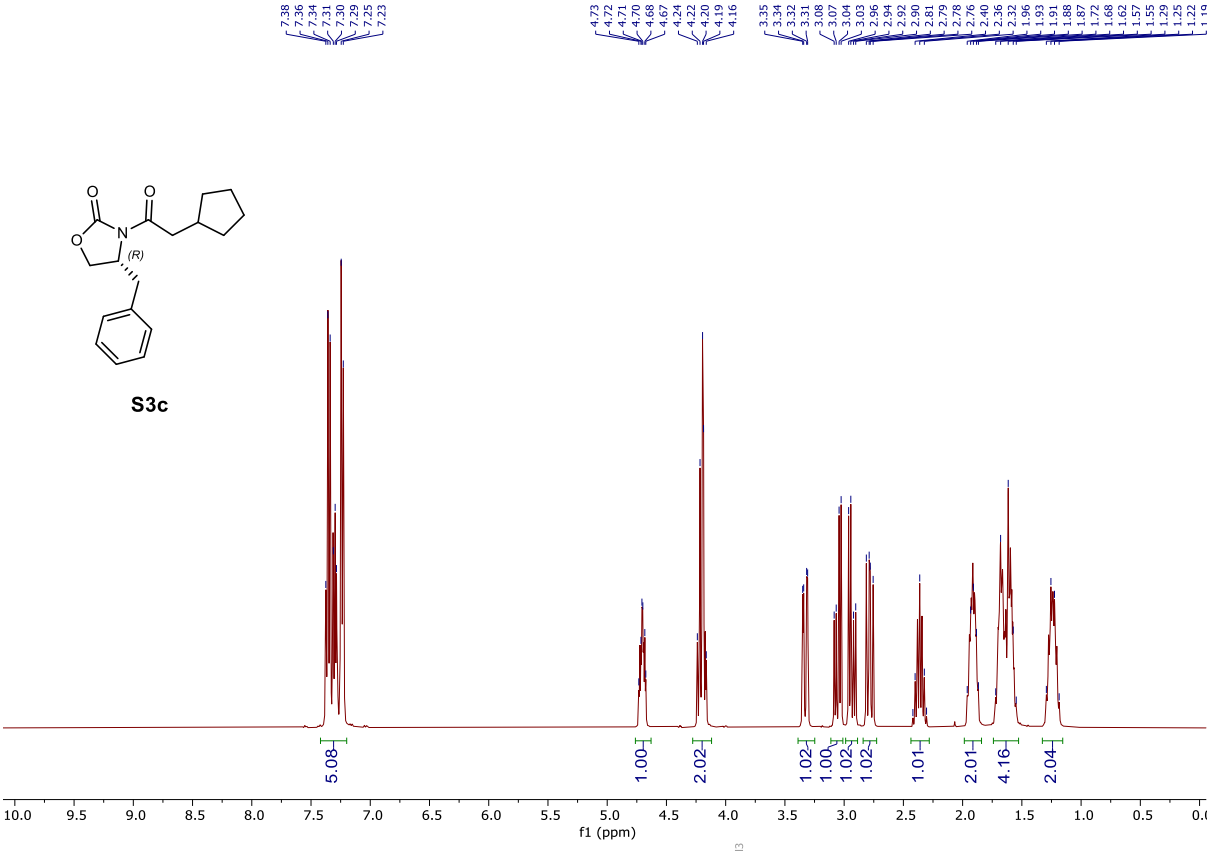


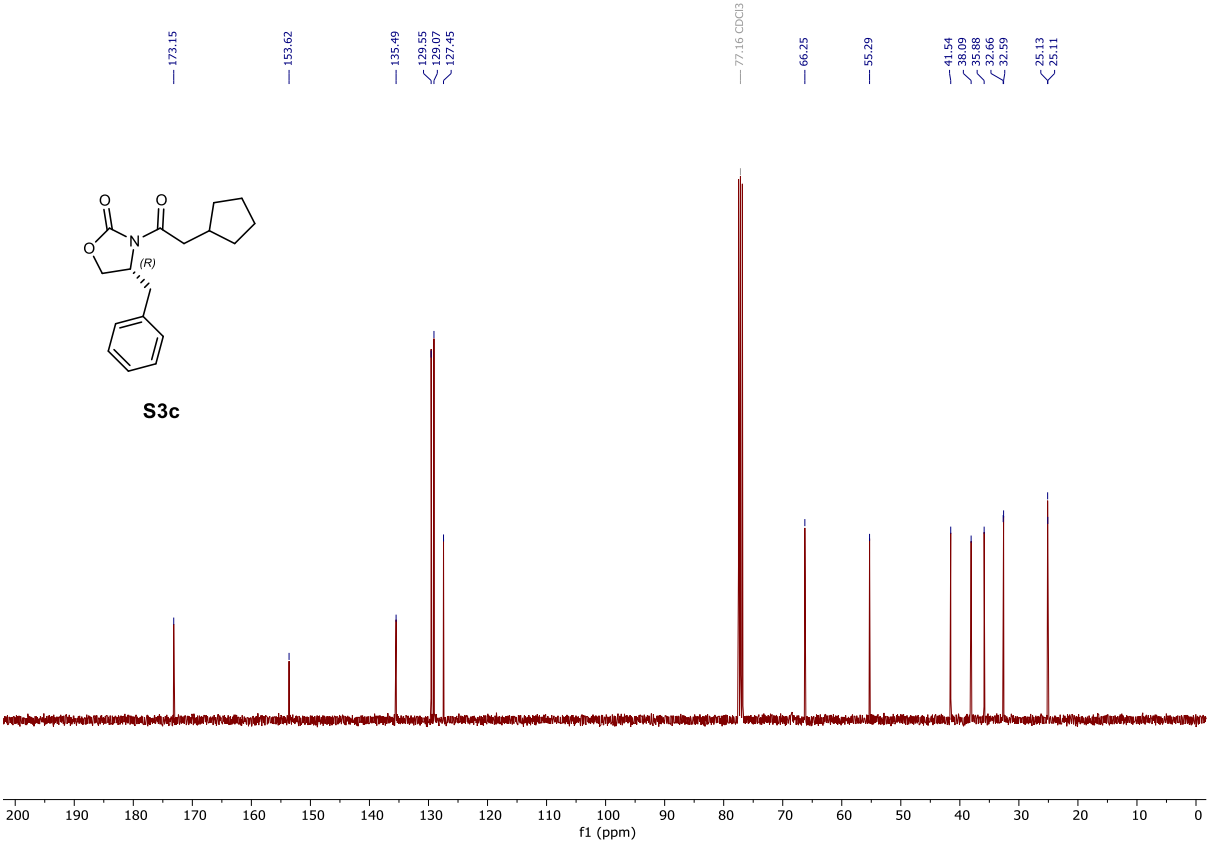
43, enzymatic
CysCD cascade

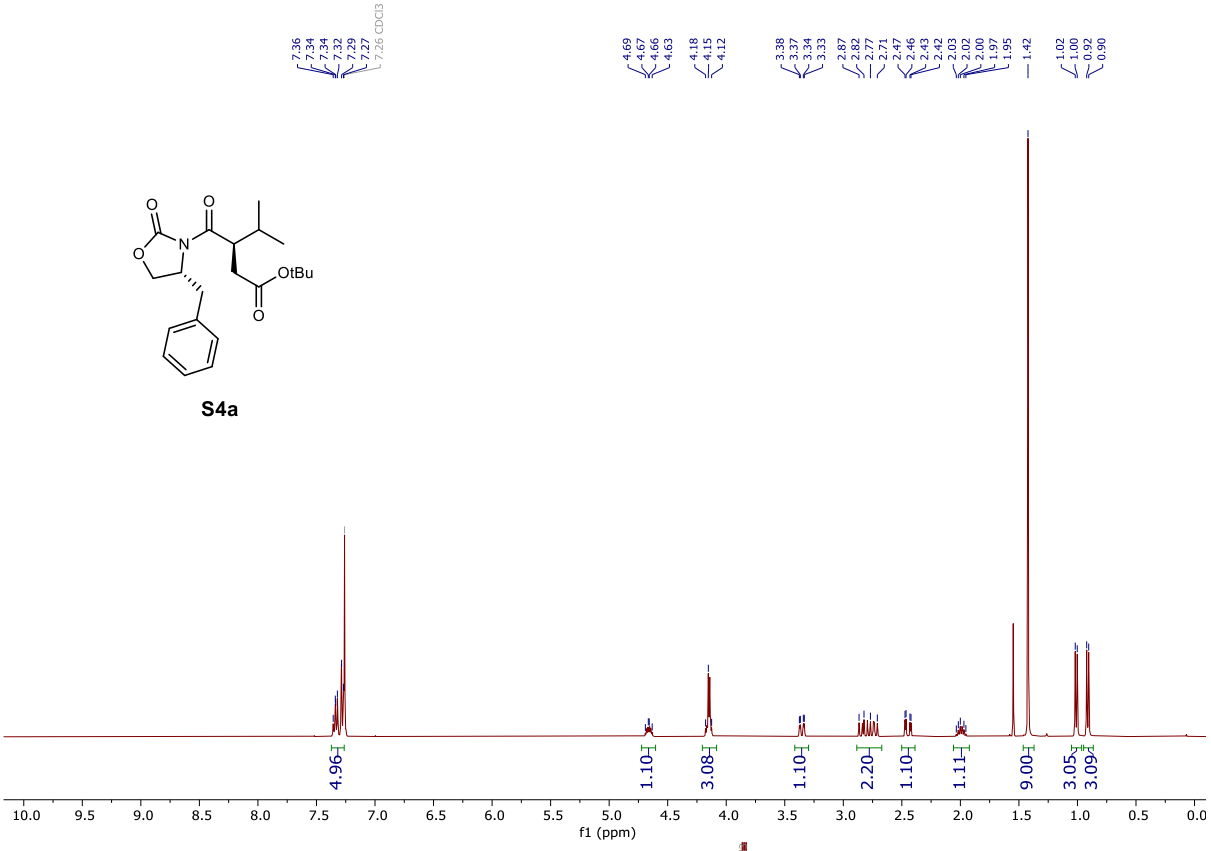
8. NMR spectra of chemically synthesised standards

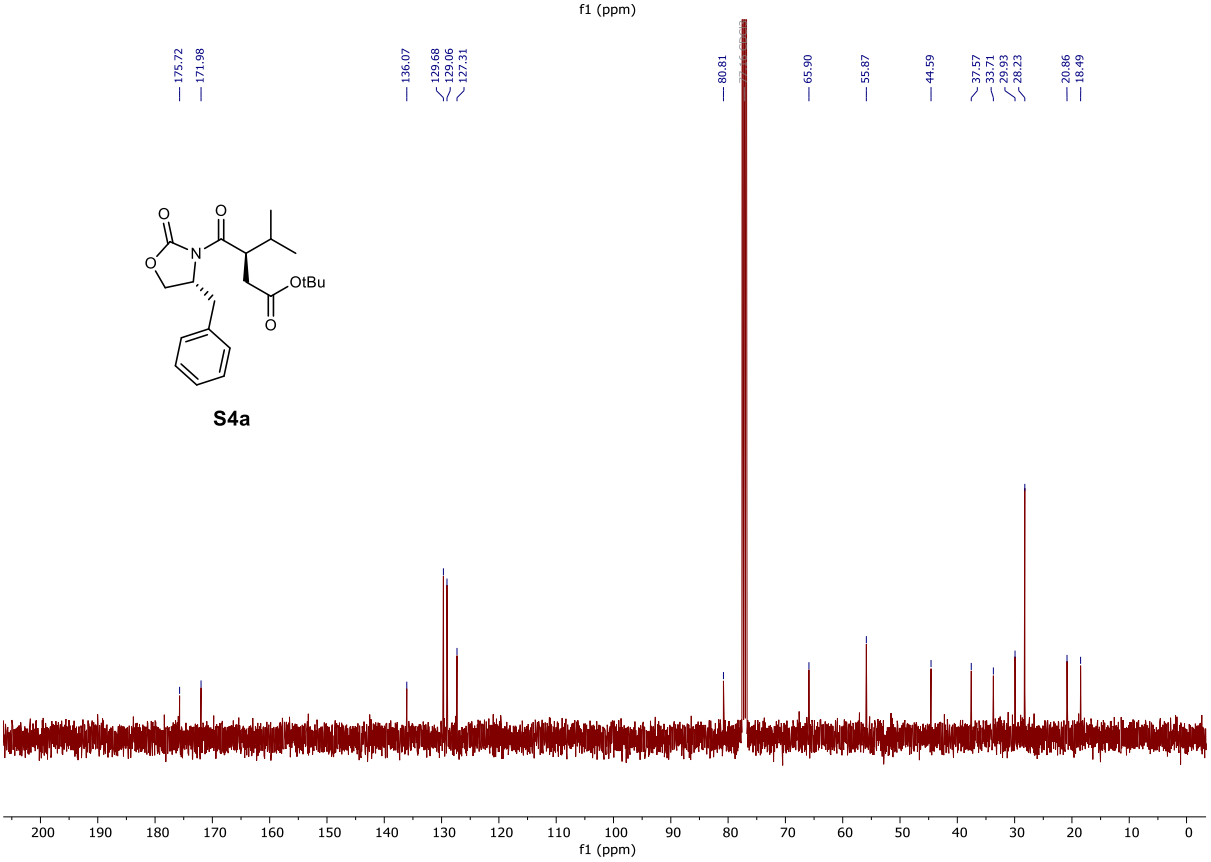


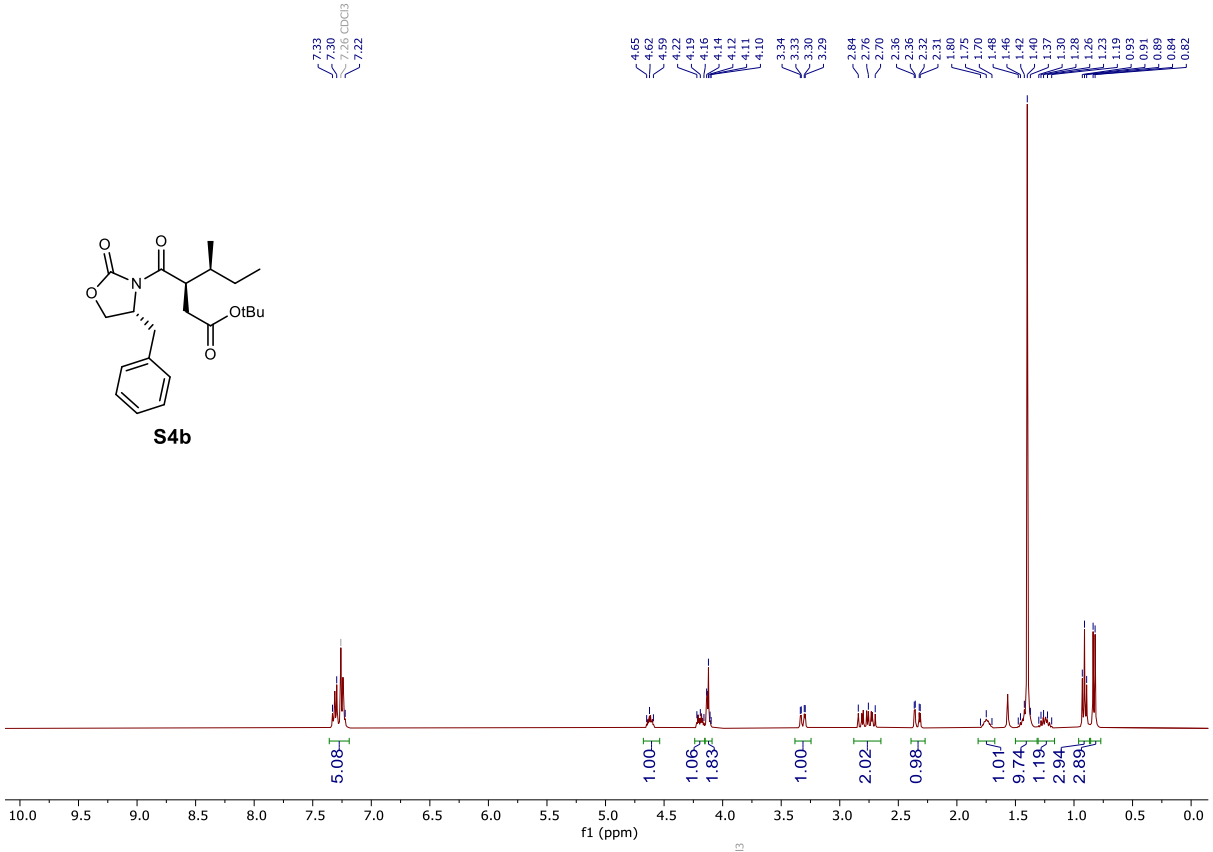


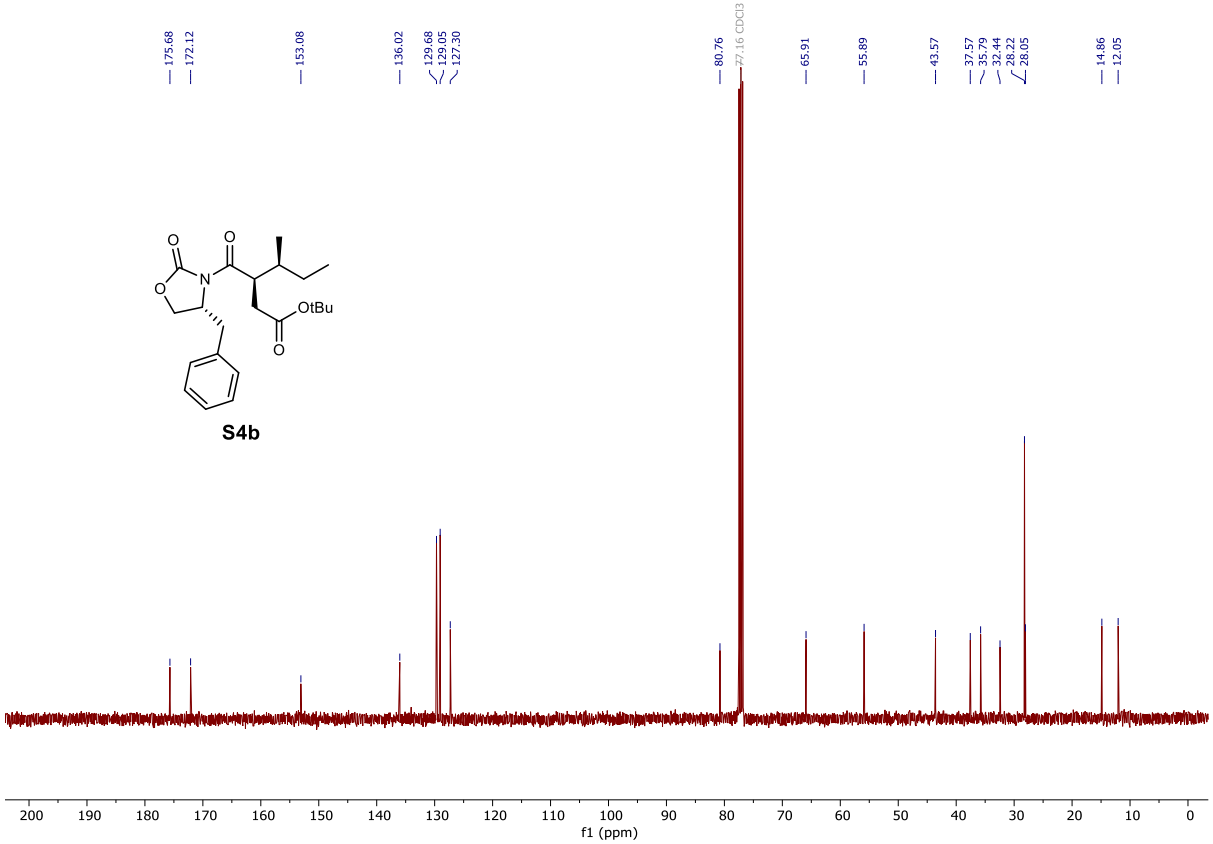

S3c
f1 (ppm)

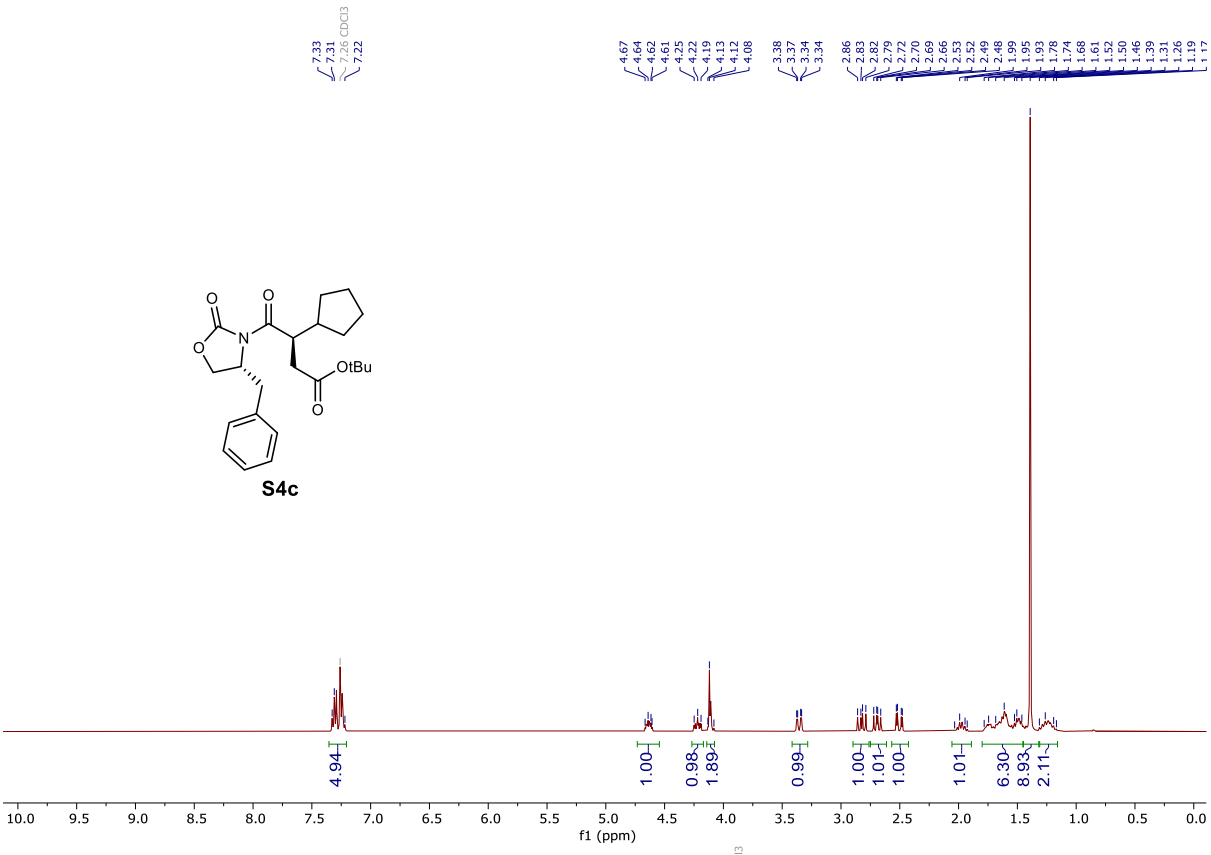

S3c
f1 (ppm)

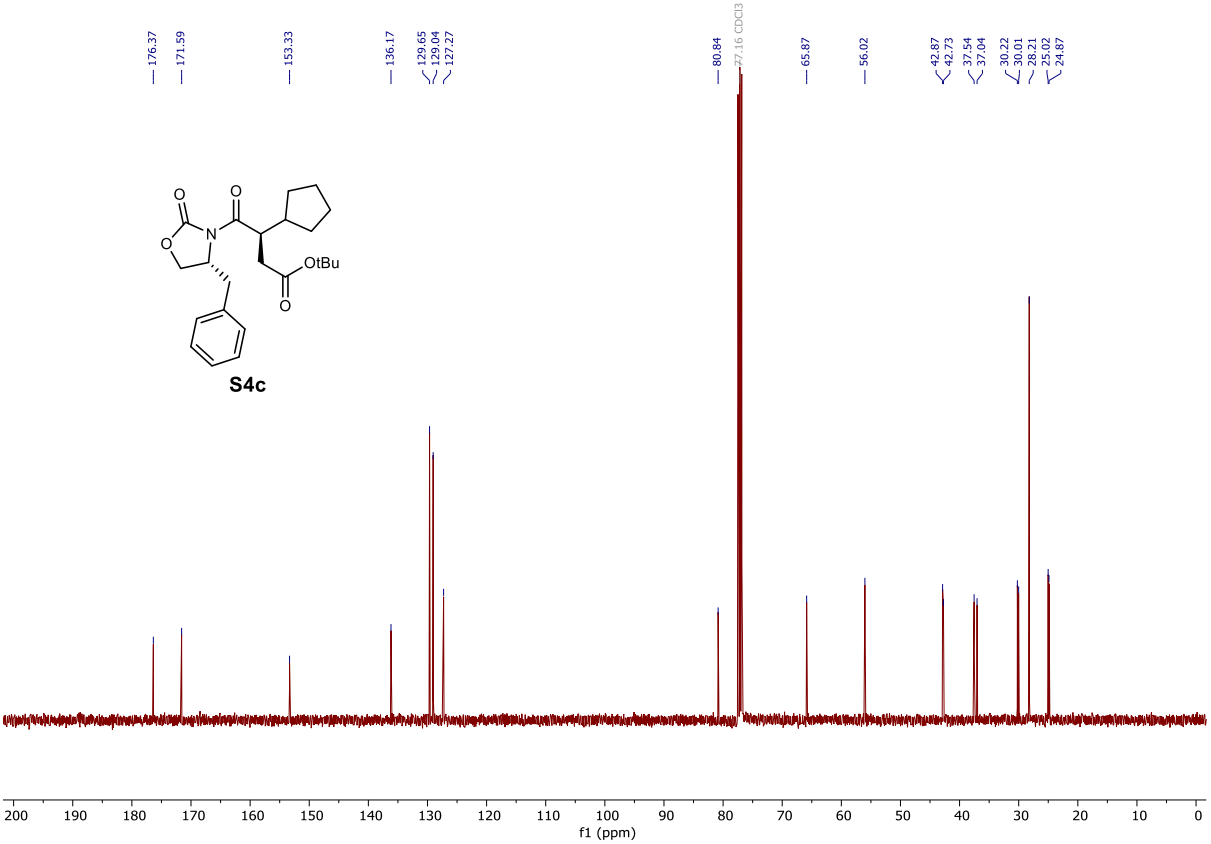

7.36
7.34
7.34
7.32
7.29
7.27
7.26 CDCl3
4.69
4.67
4.66
4.63
4.18
4.15
4.12
3.38
3.37
3.34
3.33
2.87
2.82
2.77
2.71
2.47
2.46
2.43
2.42
2.03
2.02
2.00
1.97
1.95
1.42
1.02
1.00
0.92
0.90
O
O
N
O
O
OtBu
O
S4a
4.96
1.10
3.08
1.10
2.20
1.10
1.11
9.00
3.05
3.09
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

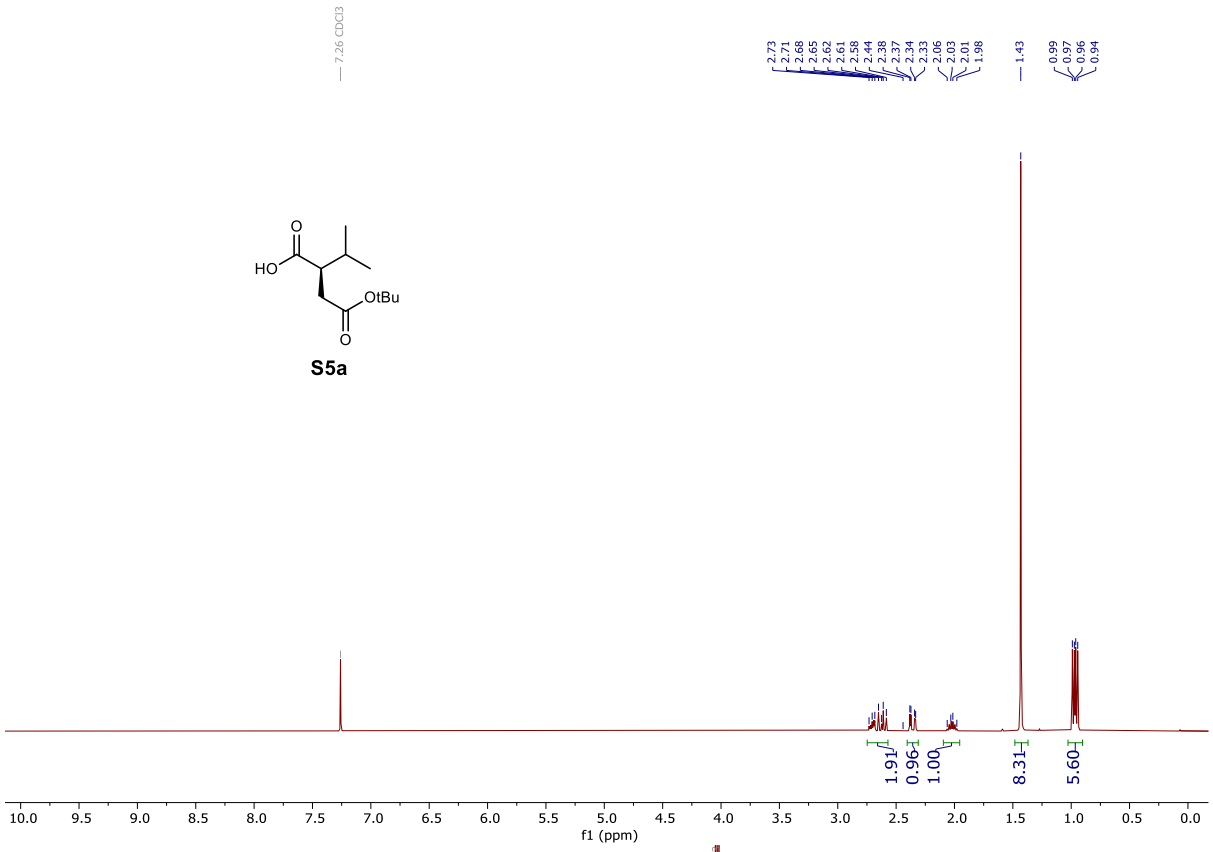

175.72
171.98
136.07
129.68
129.06
127.31
80.81
65.90
55.87
44.59
37.57
33.71
29.93
28.23
20.86
18.49
O
O
N
O
O
OtBu
O
S4a
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)


7.33
7.30
7.26 CDCl3
7.22
4.65
4.62
4.59
4.22
4.19
4.16
4.14
4.12
4.11
4.10
3.34
3.33
3.30
3.29
2.84
2.76
2.70
2.36
2.36
2.32
2.31
1.80
1.75
1.70
1.48
1.46
1.42
1.40
1.37
1.30
1.28
1.26
1.23
1.19
0.93
0.91
0.89
0.84
0.82
OtBu
S4b
5.08
1.00
1.06
1.83
1.00
2.02
0.98
1.01
9.74
1.19
2.94
2.89
f1 (ppm)

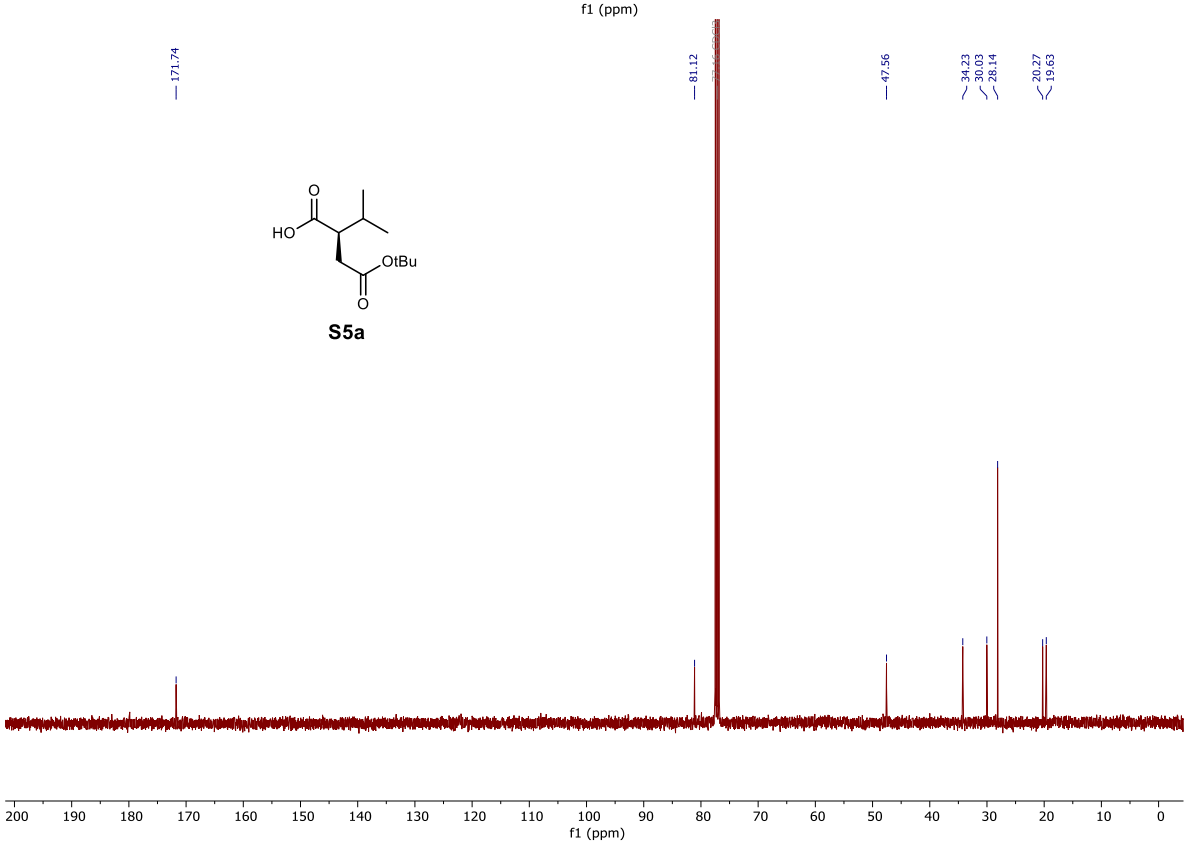

175.68
172.12
153.08
136.02
129.68
129.05
127.30
80.76
77.16 CDCl3
65.91
55.89
43.57
37.57
35.79
32.44
28.22
28.05
14.86
12.05
OtBu
S4b
f1 (ppm)


7.33
7.31
7.26 CDCl3
7.22
4.67
4.64
4.62
4.61
4.25
4.22
4.19
4.13
4.12
4.08
3.38
3.37
3.34
3.34
2.86
2.83
2.82
2.79
2.72
2.70
2.69
2.66
2.53
2.52
2.49
2.48
1.99
1.95
1.93
1.78
1.74
1.68
1.61
1.52
1.50
1.46
1.39
1.31
1.26
1.19
1.17
OtBu
S4c
4.94
1.00
0.98
1.89
0.99
1.00
1.01
1.00
1.01
6.30
8.93
2.11
f1 (ppm)

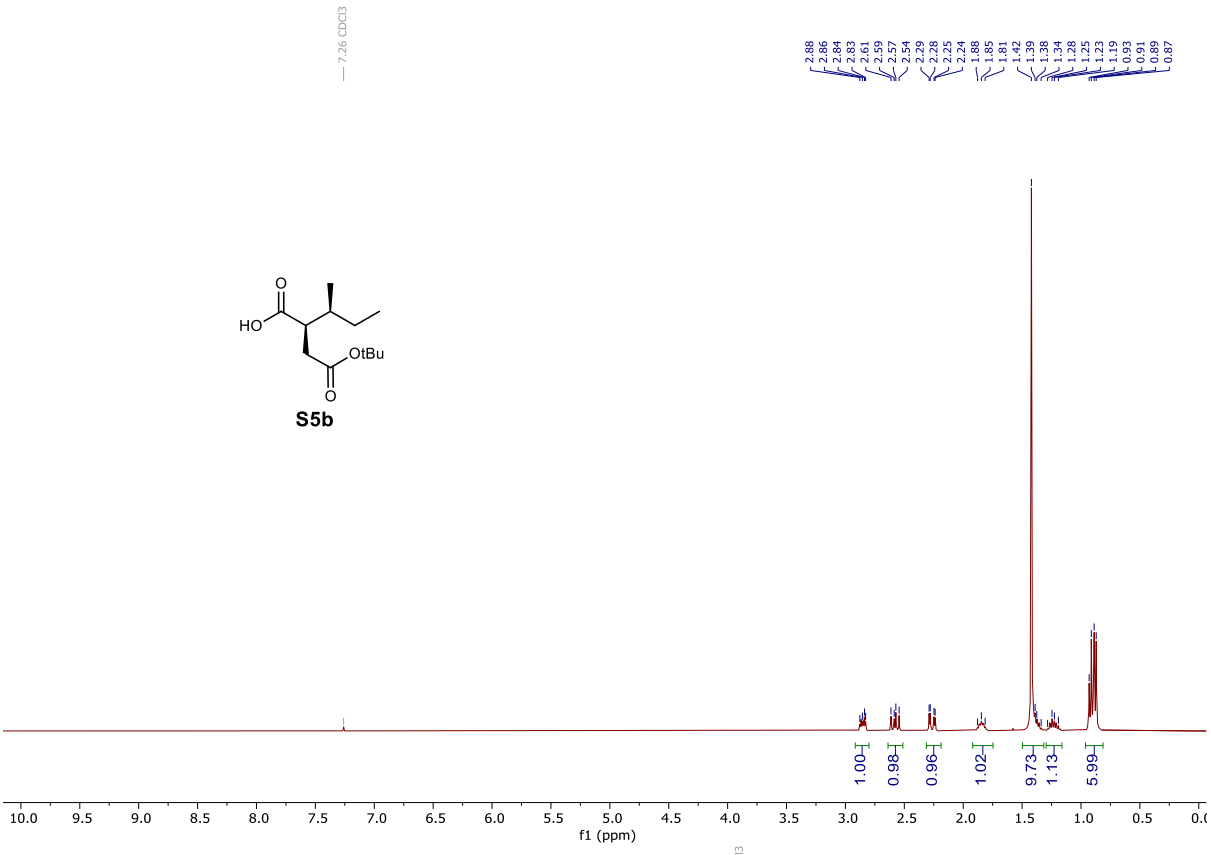

176.37
171.59
153.33
136.17
129.65
129.04
127.27
80.84
77.16 CDCl3
65.87
56.02
42.87
42.73
37.54
37.04
30.22
30.01
28.21
25.02
24.87
OtBu
S4c
f1 (ppm)

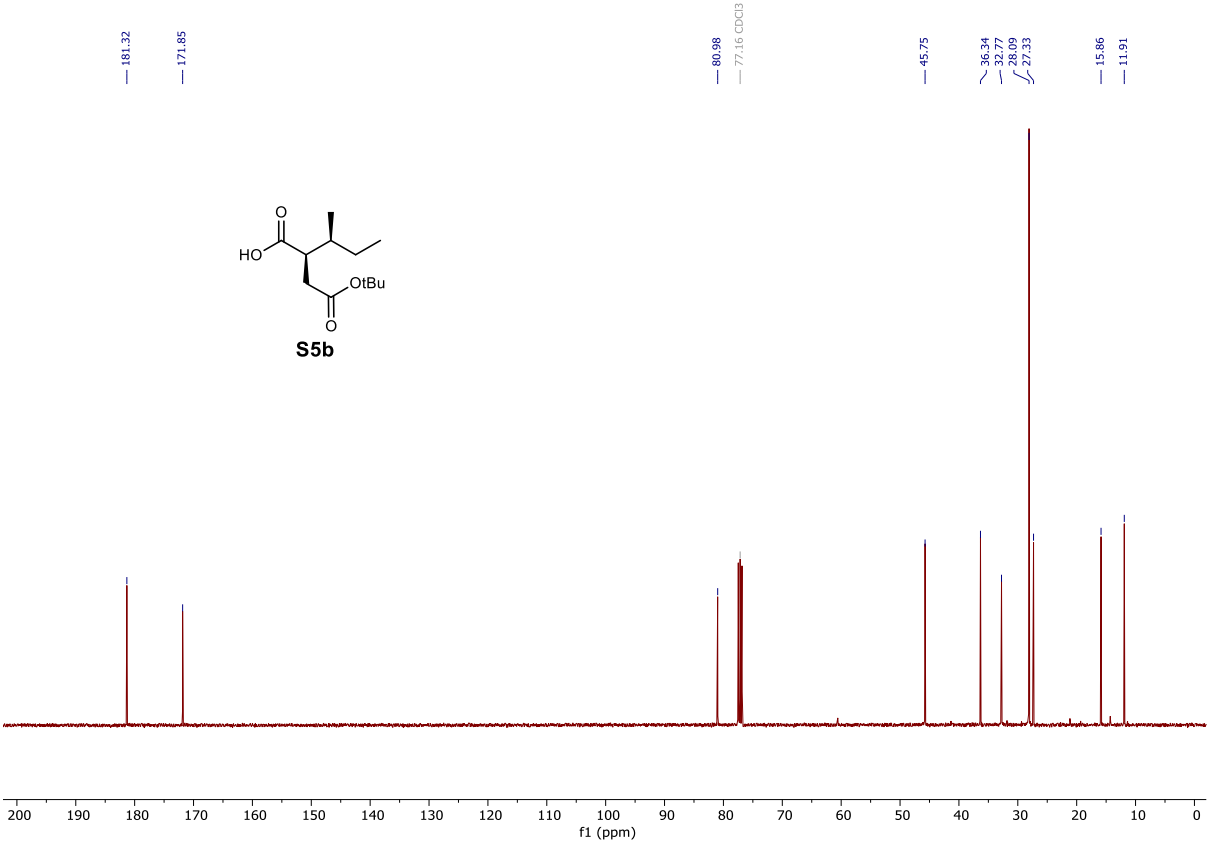

7.26 CDCl3
2.73
2.71
2.68
2.65
2.62
2.61
2.58
2.44
2.38
2.37
2.34
2.33
2.06
2.03
2.01
1.98
1.43
0.99
0.97
0.96
0.94
HO
O
OtBu
O
1.91
0.96
1.00
8.31
5.60
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

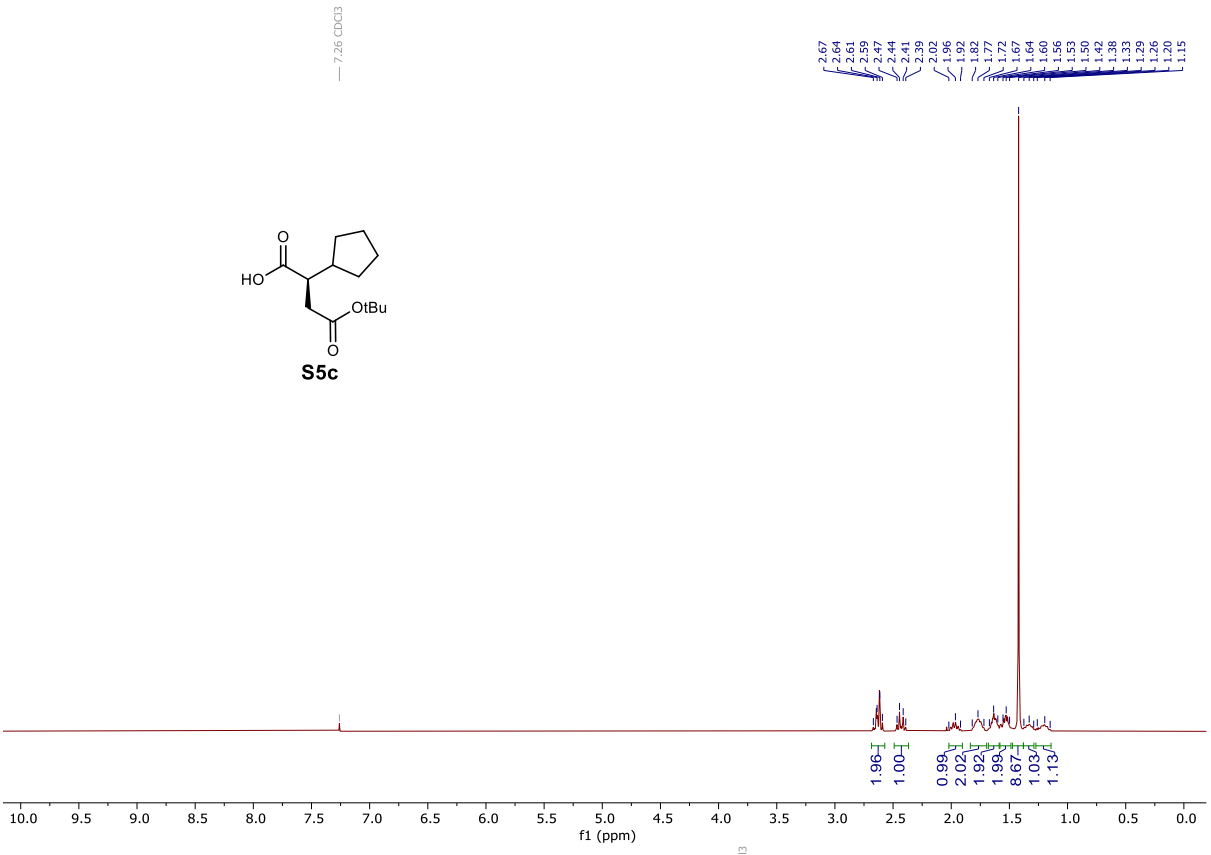
S5a

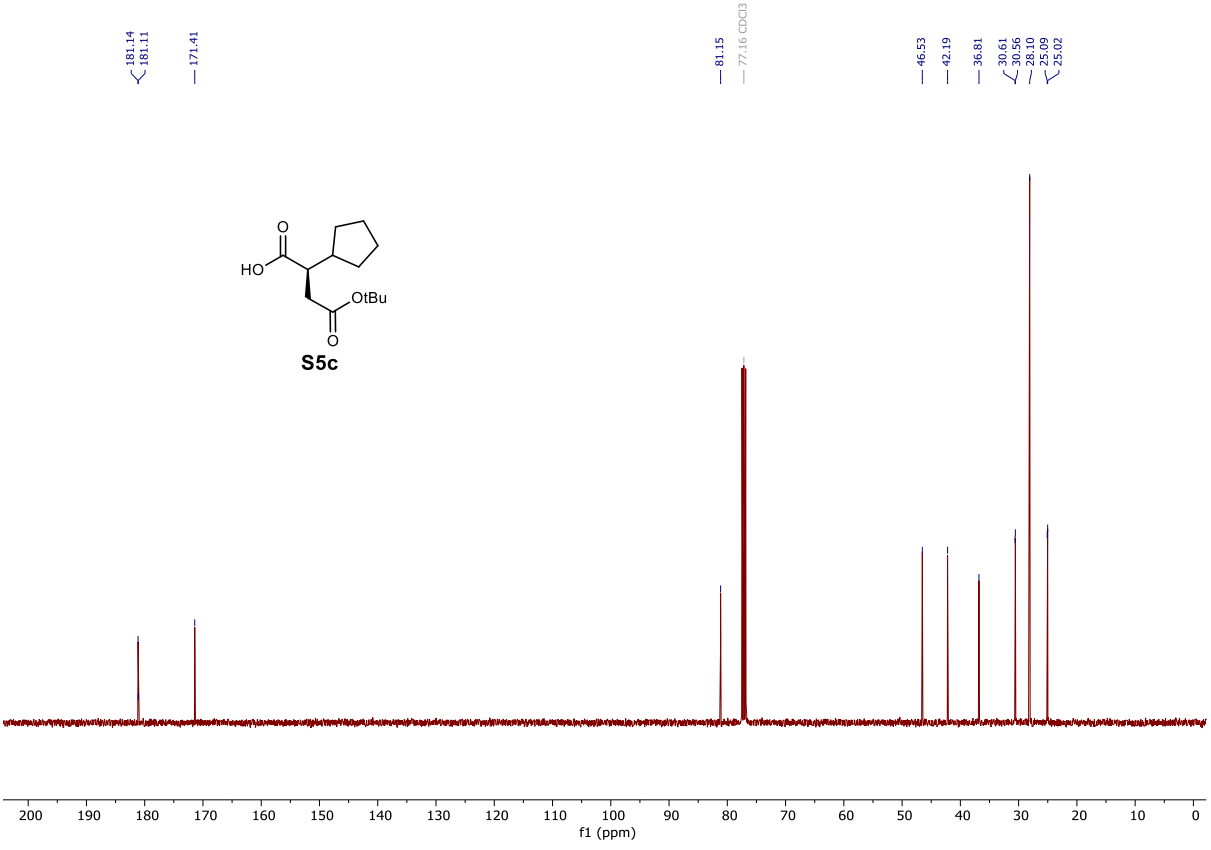

171.74
81.12
47.56
34.23
30.03
28.14
20.27
19.63
HO
O
OtBu
O
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

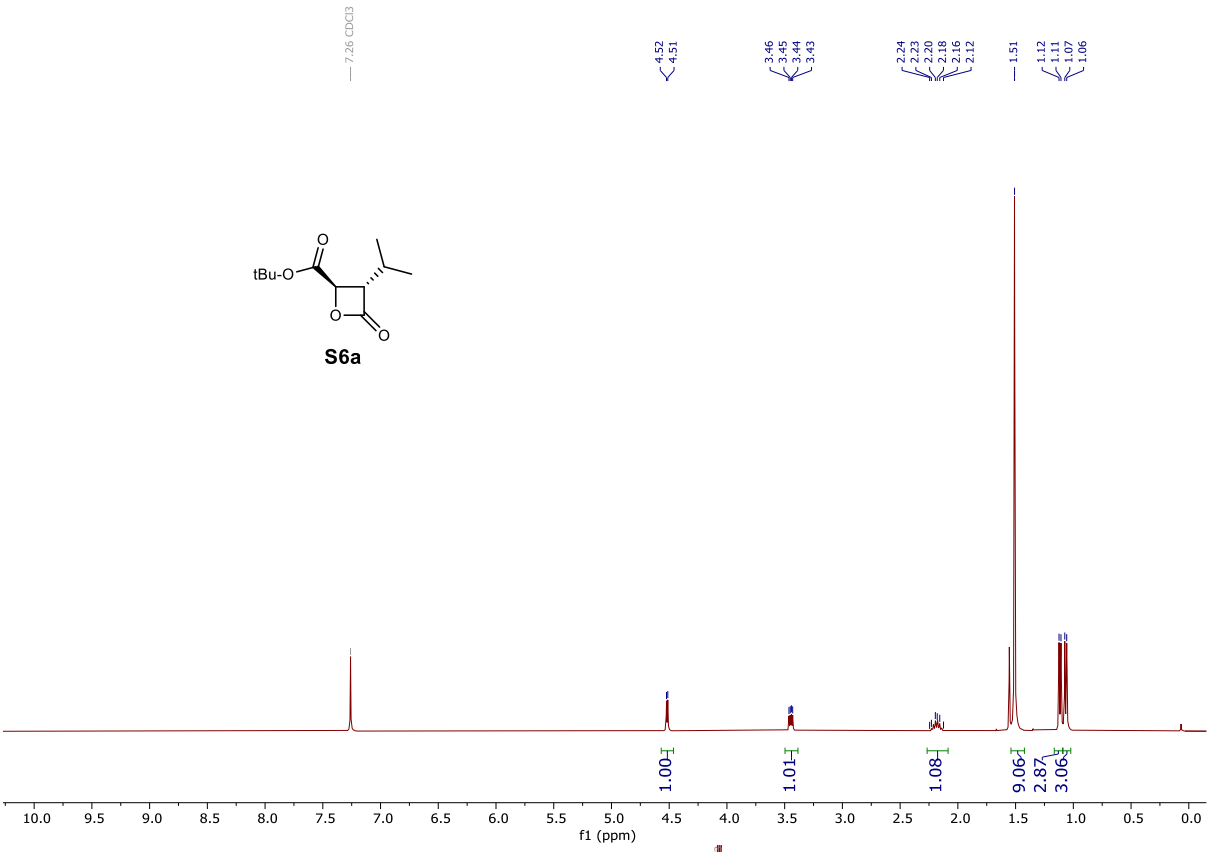
S5a

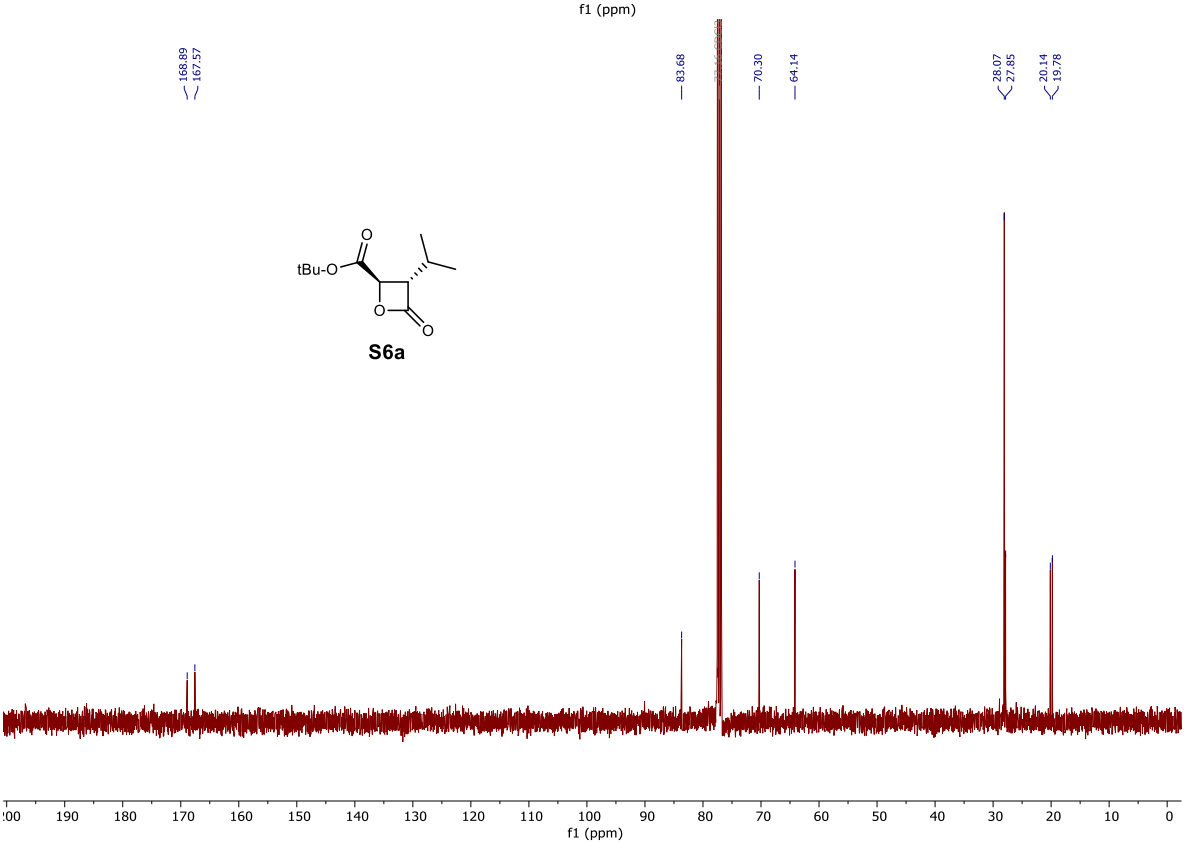

7.26 CDCl3
2.88
2.86
2.84
2.83
2.61
2.59
2.57
2.54
2.29
2.28
2.25
2.24
1.88
1.85
1.81
1.42
1.39
1.38
1.34
1.28
1.25
1.23
1.19
0.93
0.91
0.89
0.87
HO
OtBu
O
O
S5b
1.00
0.98
0.96
1.02
9.73
1.13
5.99
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
f1 (ppm)

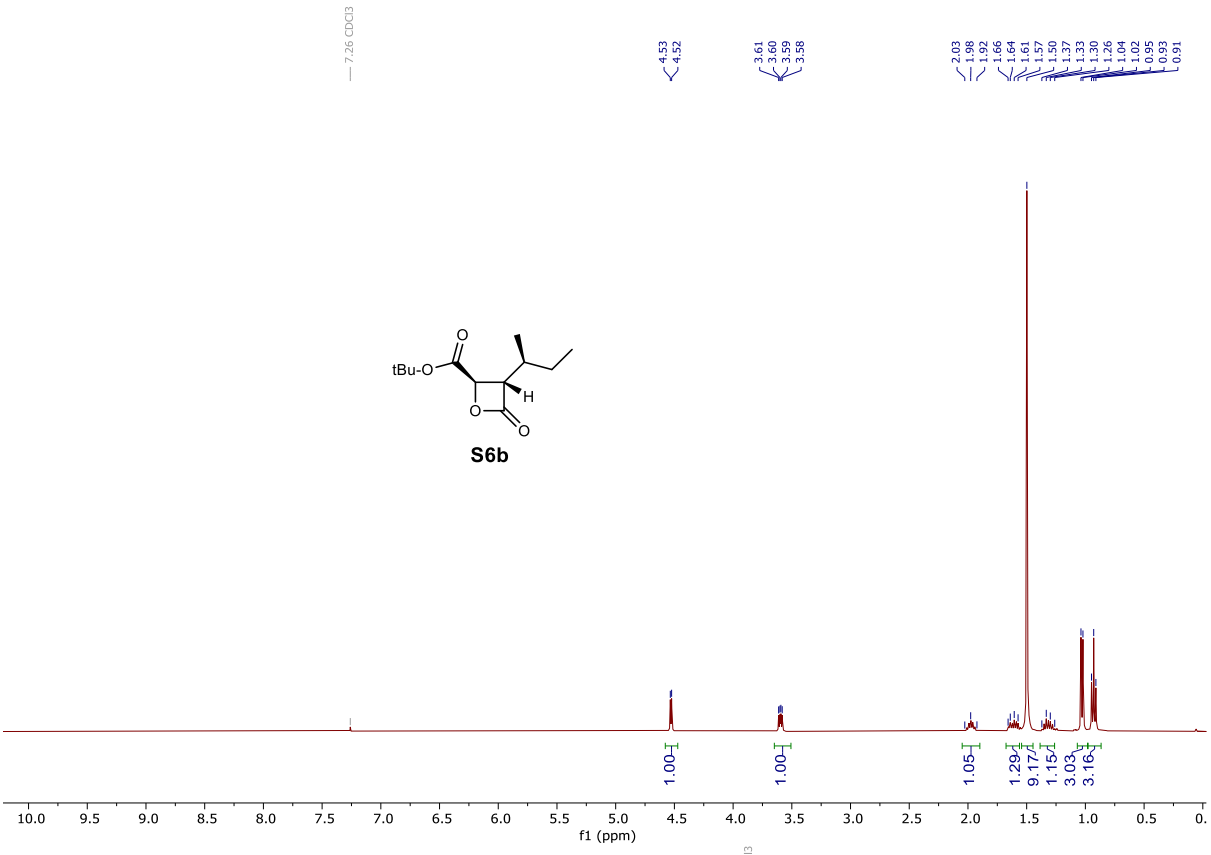

181.32
171.85
80.98
77.16 CDCl3
45.75
36.34
32.77
28.09
27.33
15.86
11.91
HO
OtBu
O
O
S5b
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

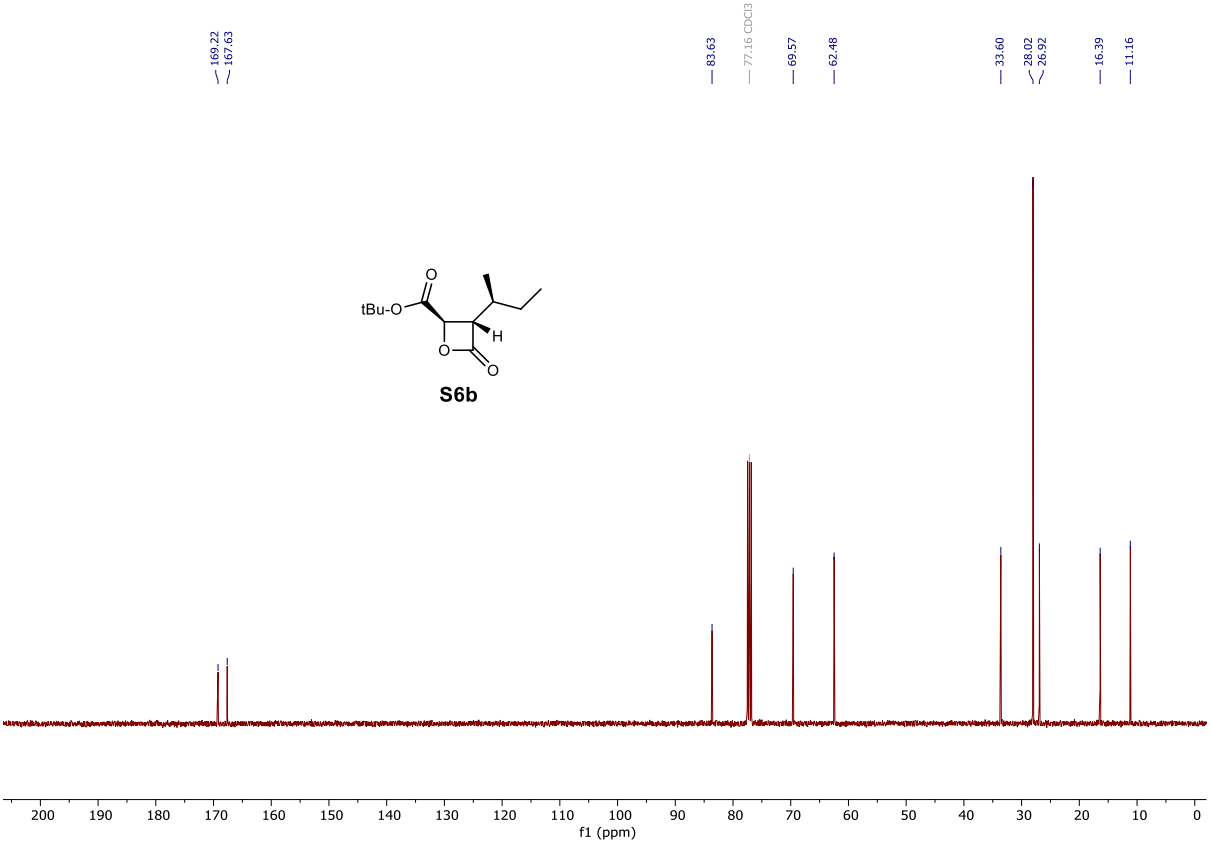

7.26 CDCl3
2.67
2.64
2.61
2.59
2.47
2.44
2.41
2.39
2.02
1.96
1.92
1.82
1.77
1.72
1.67
1.64
1.60
1.56
1.53
1.50
1.42
1.38
1.33
1.29
1.26
1.20
1.15
HO
O
OtBu
O
S5c
1.96
1.00
0.99
2.02
1.92
1.99
8.67
1.03
1.13
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

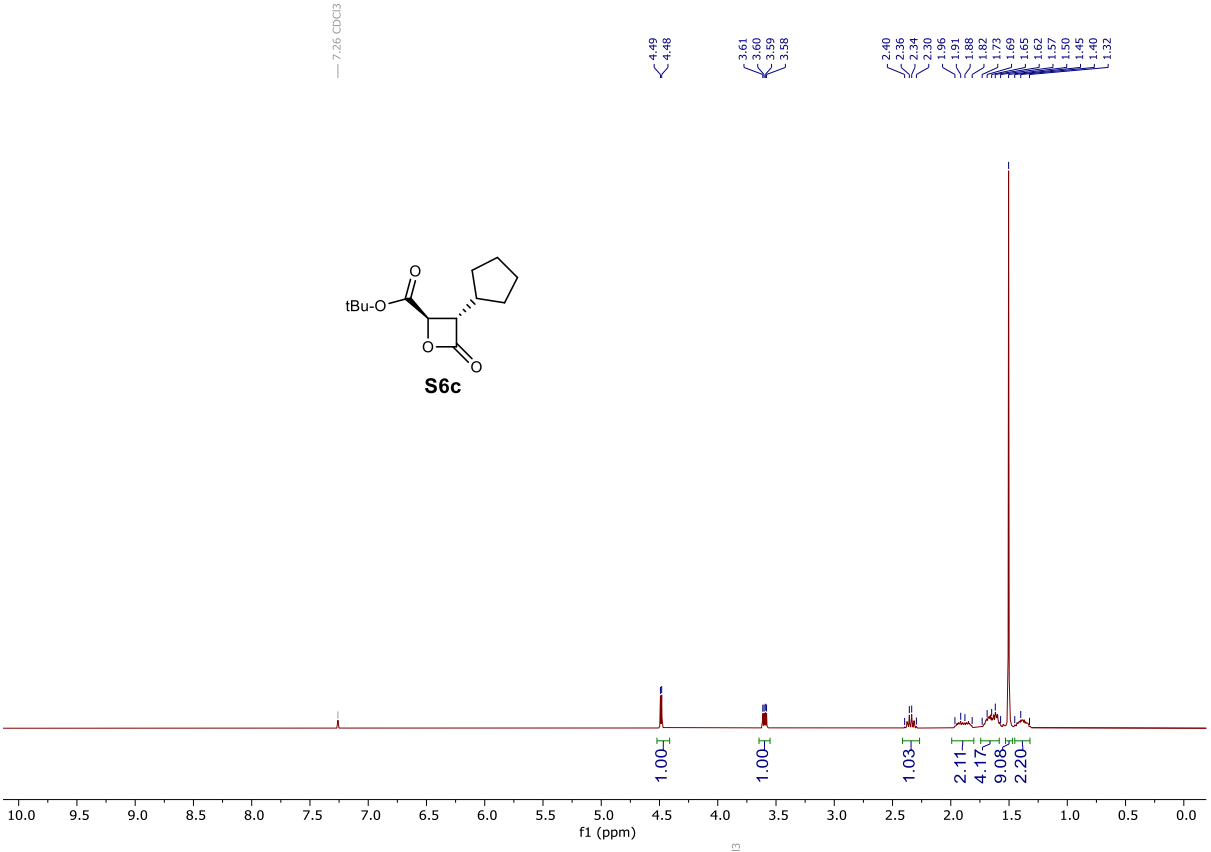

181.14
181.11
171.41
81.15
77.16 CDCl3
46.53
42.19
36.81
30.61
30.56
28.10
25.09
25.02
HO
O
OtBu
O
S5c
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)


7.26 CDCl3
4.52
4.51
3.46
3.45
3.44
3.43
2.24
2.23
2.20
2.18
2.16
2.12
1.51
1.12
1.11
1.07
1.06
tBu-O
O
O
O
S6a
1.00
1.01
1.08
9.06
2.87
3.06
f1 (ppm)

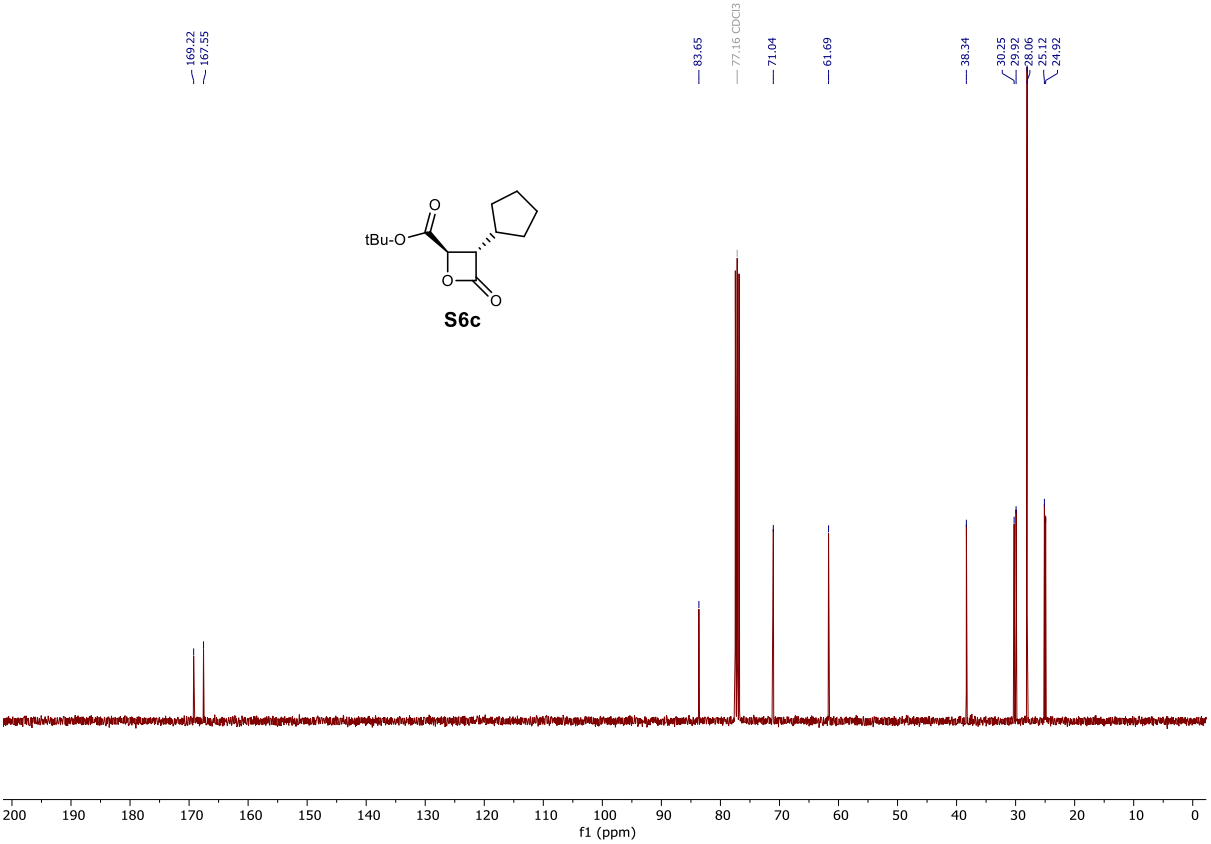

168.89
167.57
83.68
70.30
64.14
28.07
27.85
20.14
19.78
tBu-O
O
O
O
S6a
f1 (ppm)


7.26 CDCl3
4.53
4.52
3.61
3.60
3.59
3.58
2.03
1.98
1.92
1.66
1.64
1.61
1.57
1.50
1.37
1.33
1.30
1.26
1.04
1.02
0.95
0.93
0.91
tBu-O
O
O
H
O
S6b
1.00
1.00
1.05
1.29
9.17
1.15
3.03
3.16
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.
f1 (ppm)

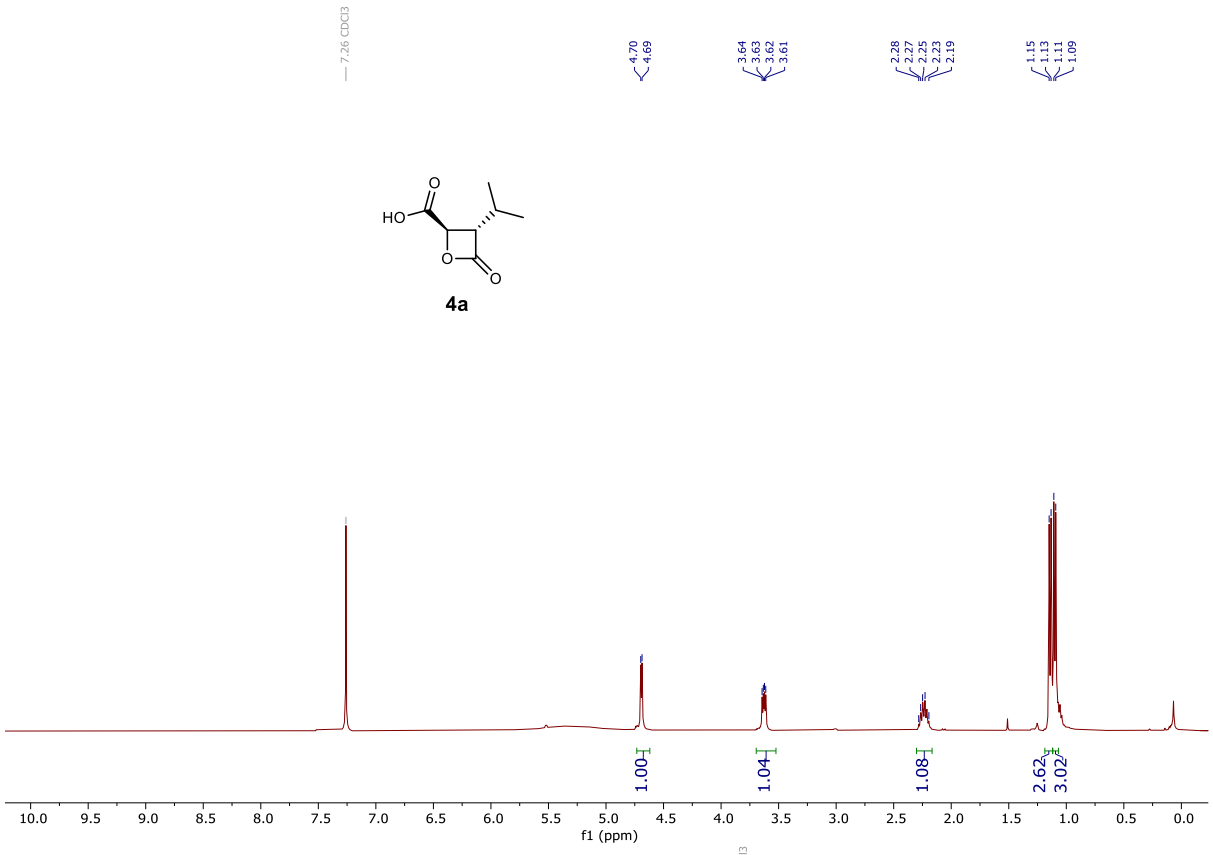

169.22
167.63
83.63
77.16 CDCl3
69.57
62.48
33.60
28.02
26.92
16.39
11.16
tBu-O
O
O
H
O
S6b
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

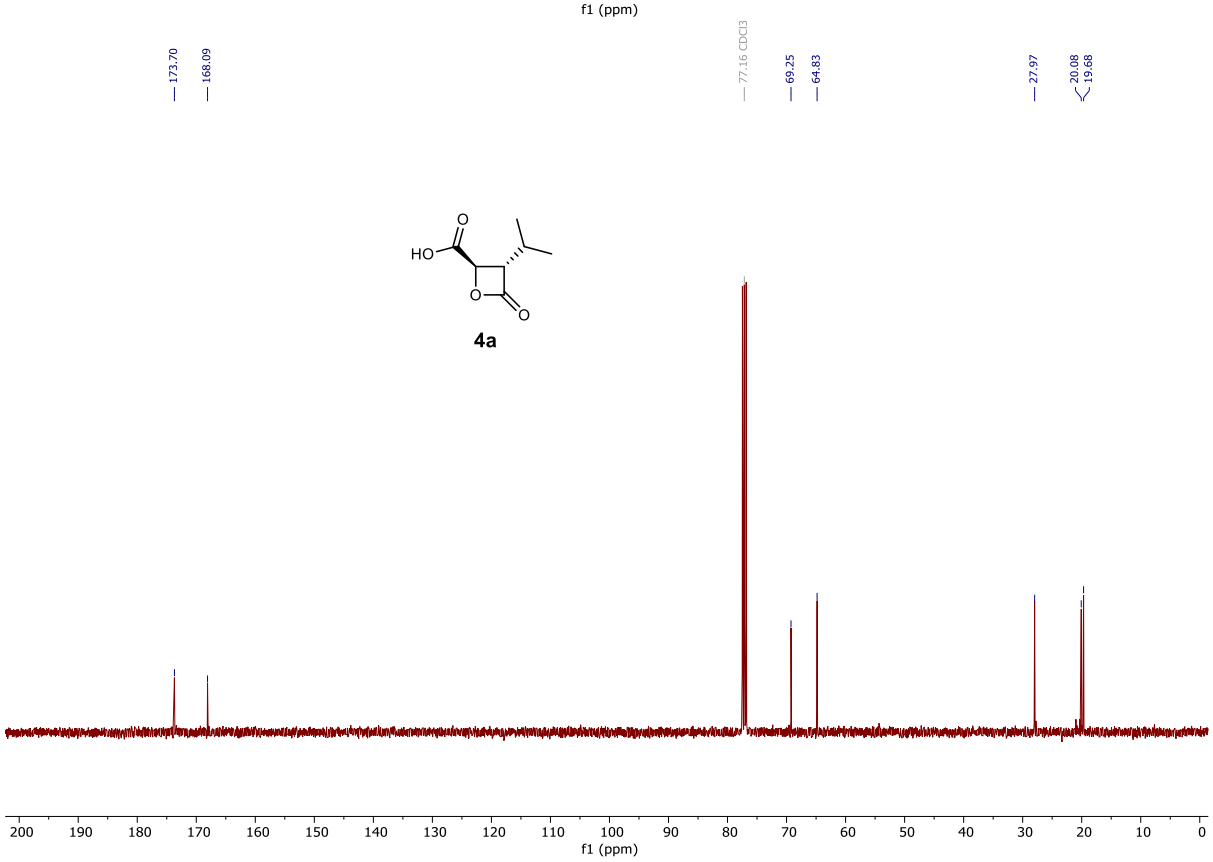

7.26 CDCl3
4.49
4.48
3.61
3.60
3.59
3.58
2.40
2.36
2.34
2.30
1.96
1.91
1.88
1.82
1.73
1.69
1.65
1.62
1.57
1.50
1.45
1.40
1.32
tBu-O
O
O
O
1.00
1.00
1.03
2.11
4.17
9.08
2.20
f1 (ppm)

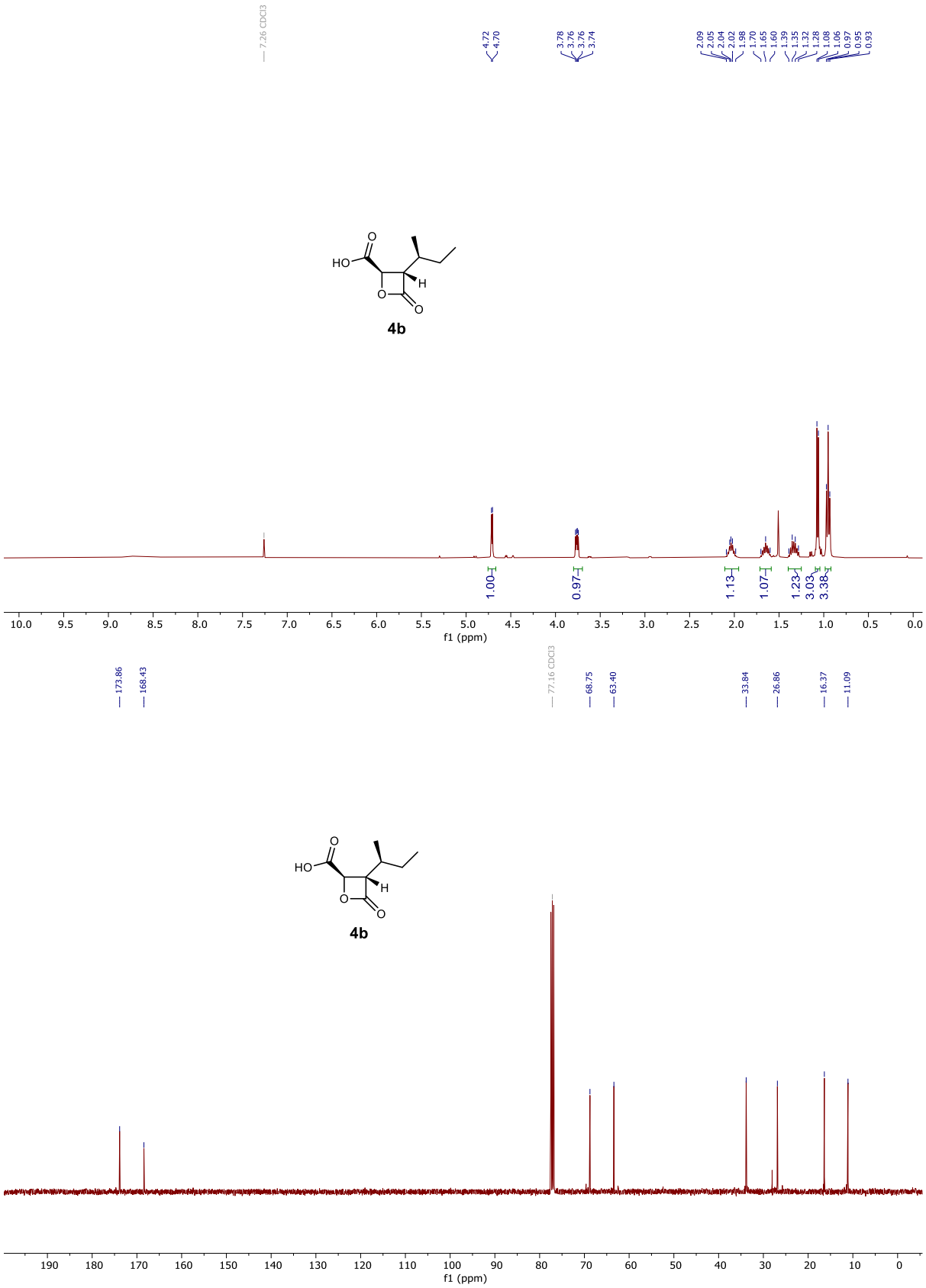
S6c

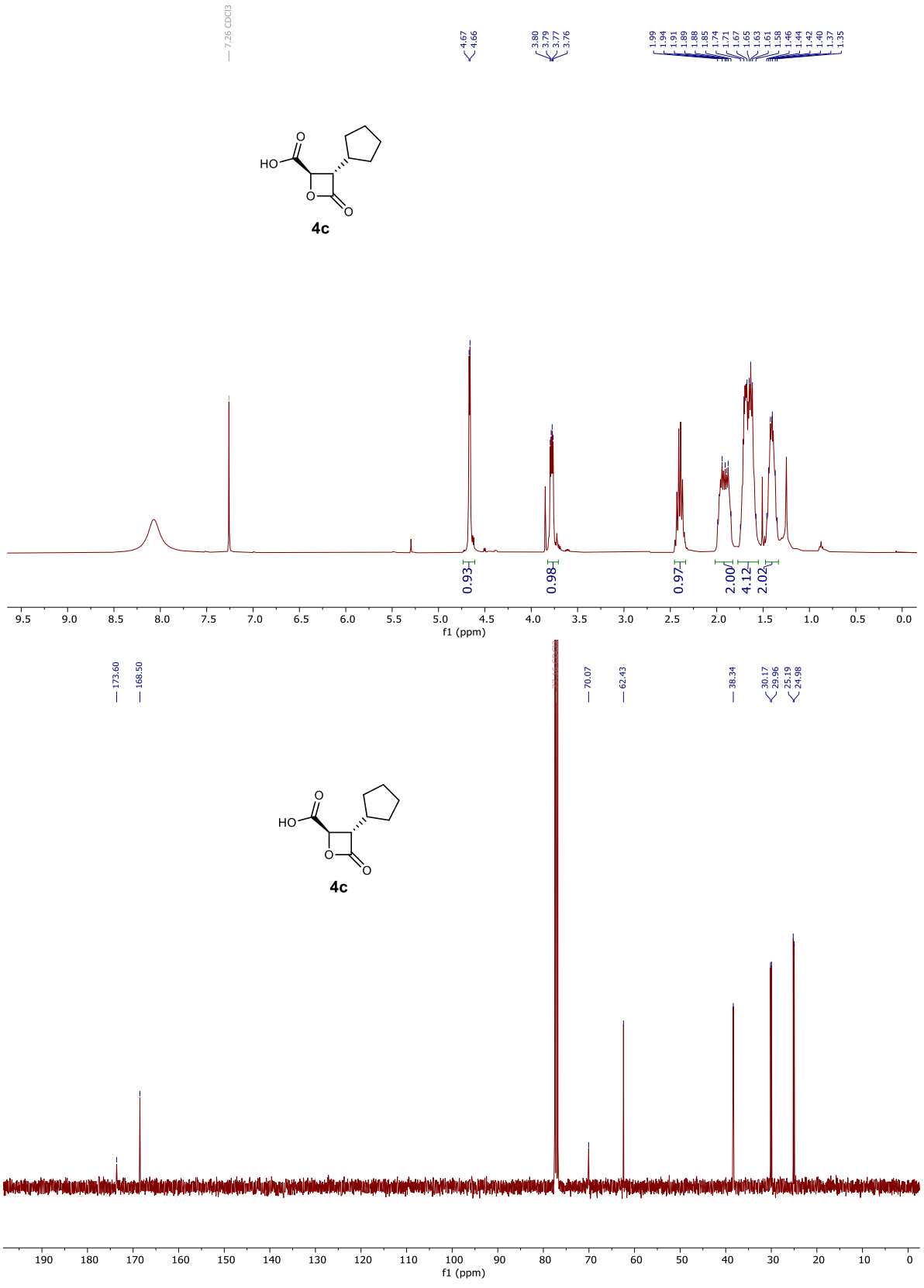

169.22
167.55
83.65
77.16 CDCl3
71.04
61.69
38.34
30.25
29.92
28.06
25.12
24.92
tBu-O
O
O
O
f1 (ppm)

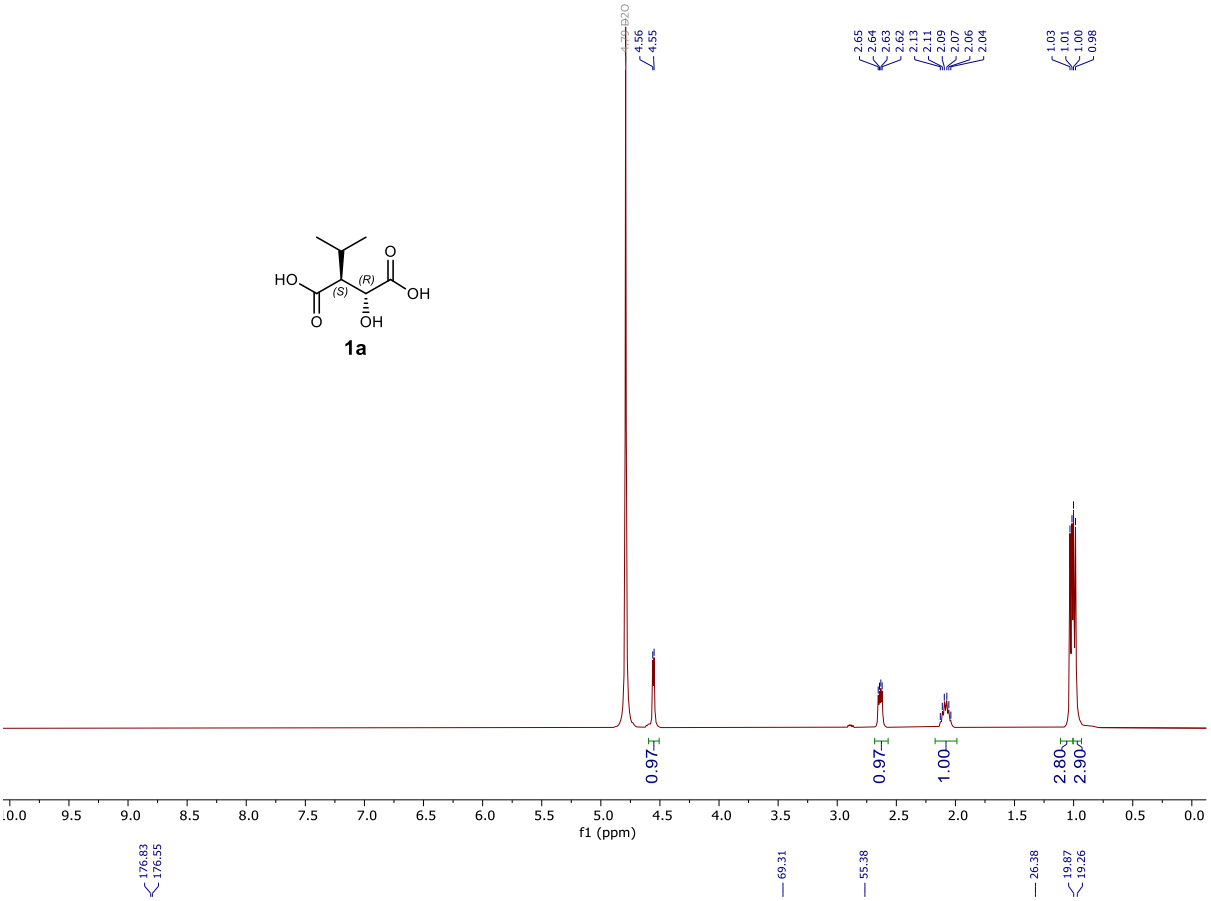
S6c

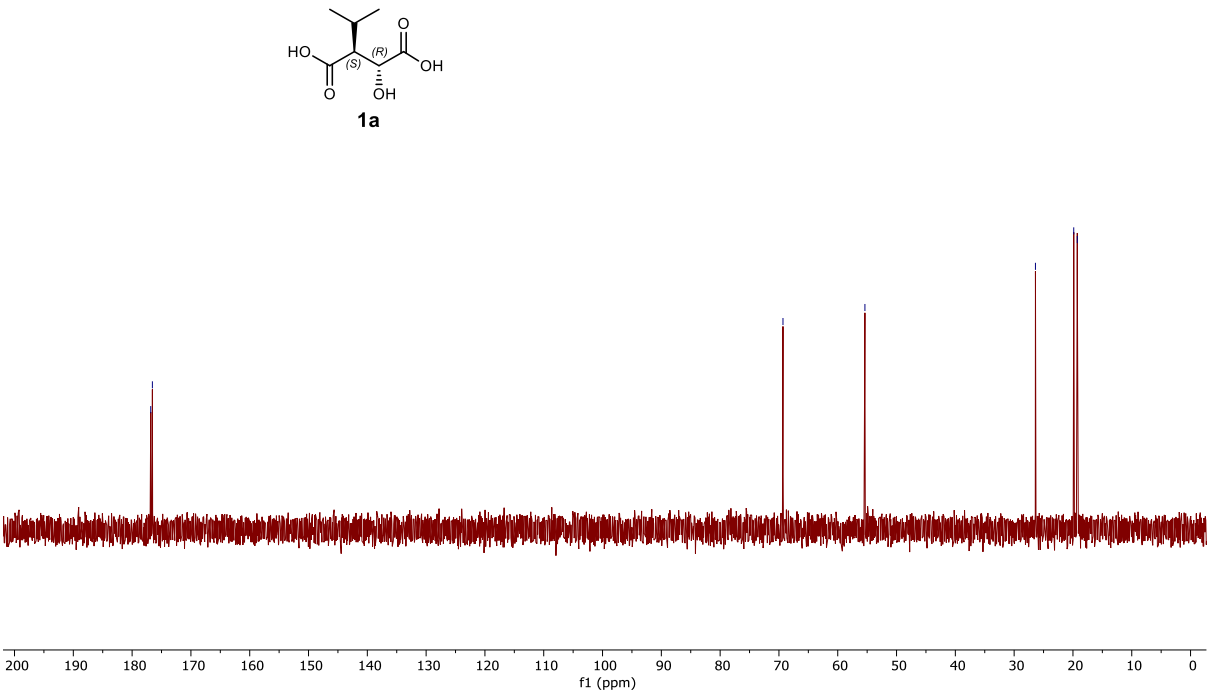

7.26 CDCl3
4.70
4.69
3.64
3.63
3.62
3.61
2.28
2.27
2.25
2.23
2.19
1.15
1.13
1.11
1.09
HO
O
O
O
4a
1.00
1.04
1.08
2.62
3.02
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

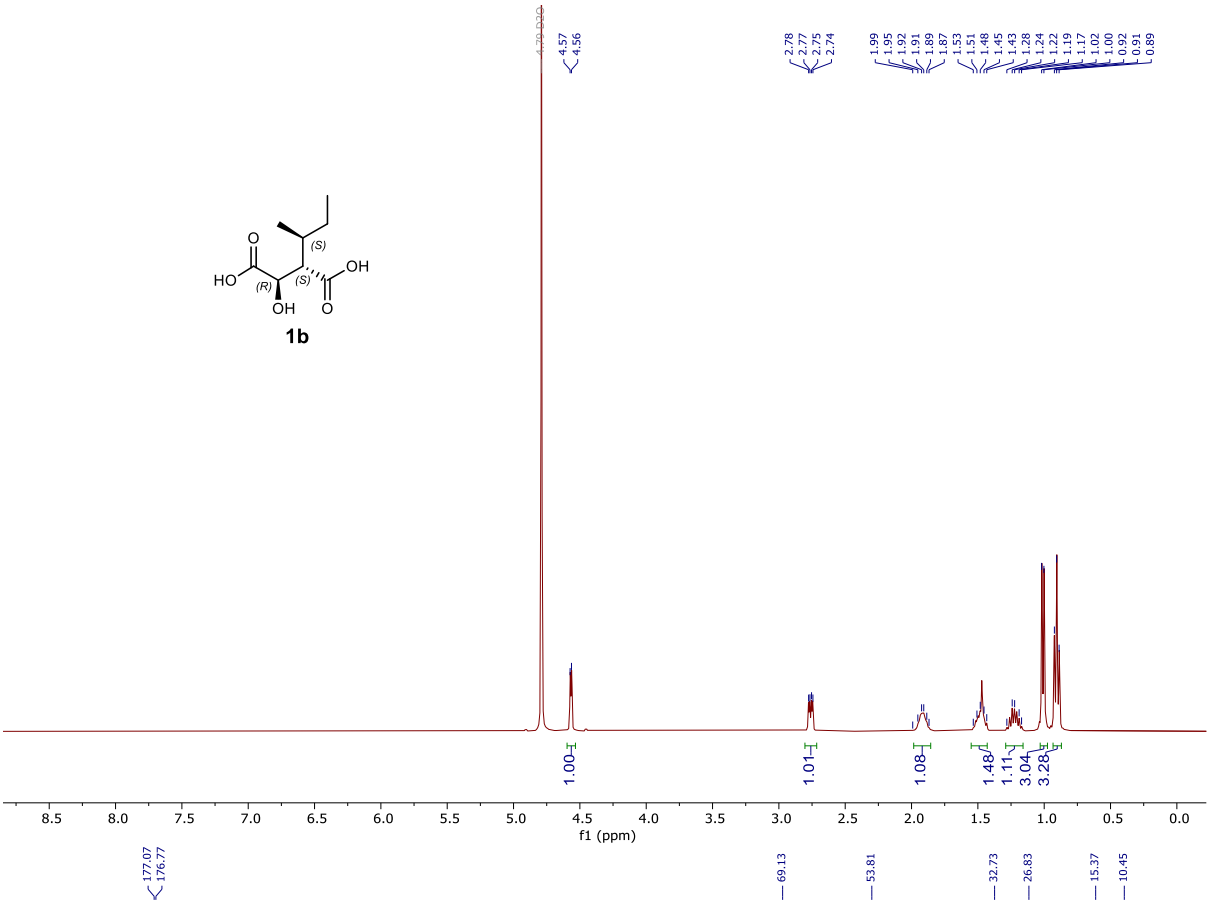

173.70
168.09
77.16 CDCl3
69.25
64.83
27.97
20.08
19.68
HO
O
O
O
4a
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)


7.26 CDCl3
4.72
4.70
3.78
3.76
3.76
3.74
2.09
2.05
2.04
2.02
1.98
1.70
1.65
1.60
1.39
1.35
1.32
1.28
1.08
1.06
0.97
0.95
0.93
HO
O
H
O
O
4b
1.00
0.97
1.13
1.07
1.23
3.03
3.38
f1 (ppm)
173.86
168.43
77.16 CDCl3
68.75
63.40
33.84
26.86
16.37
11.09
HO
O
H
O
O
4b
f1 (ppm)


7.26 CDCl3
4.67
4.66
3.80
3.79
3.77
3.76
1.99
1.94
1.91
1.89
1.88
1.85
1.74
1.71
1.67
1.65
1.63
1.61
1.58
1.46
1.44
1.42
1.40
1.37
1.35
HO
O
O
O
4c
0.93
0.98
0.97
2.00
4.12
2.02
f1 (ppm)
173.60
168.50
70.07
62.43
38.34
30.17
29.96
25.19
24.98
HO
O
O
O
4c
f1 (ppm)


HO
O
(S)
(R)
O
OH
OH
1a
4.56
4.55
2.65
2.64
2.63
2.62
2.13
2.11
2.09
2.07
2.06
2.04
1.03
1.01
1.00
0.98
0.97
0.97
1.00
2.80
2.90
f1 (ppm)

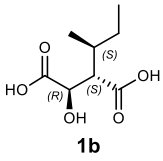

HO
O
(S)
(R)
O
OH
OH
1a
176.83
176.55
69.31
55.38
26.38
19.87
19.26
f1 (ppm)


4.57
4.56
2.78
2.77
2.75
2.74
1.99
1.95
1.92
1.91
1.89
1.87
1.53
1.51
1.48
1.45
1.43
1.28
1.24
1.22
1.19
1.17
1.02
1.00
0.92
0.91
0.89
1.00
1.01
1.08
1.48
1.11
3.04
3.28
f1 (ppm)

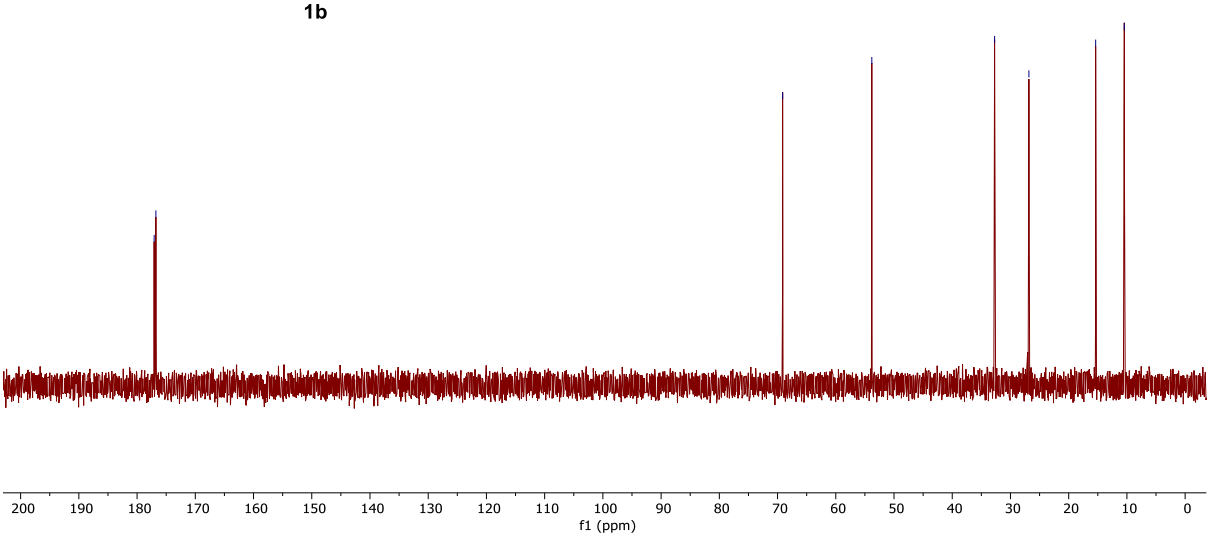
HO
O
(R)
OH
(S)
(S)
OH
O

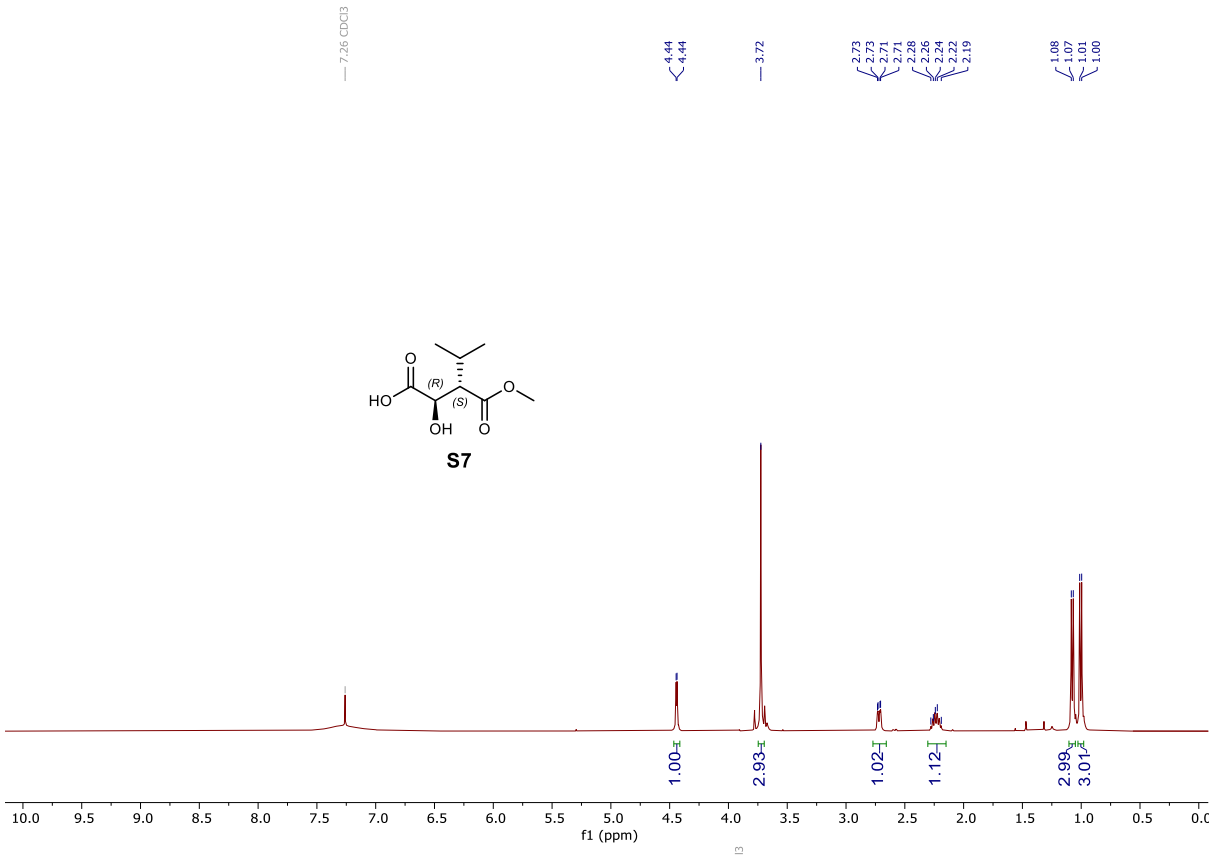
1b

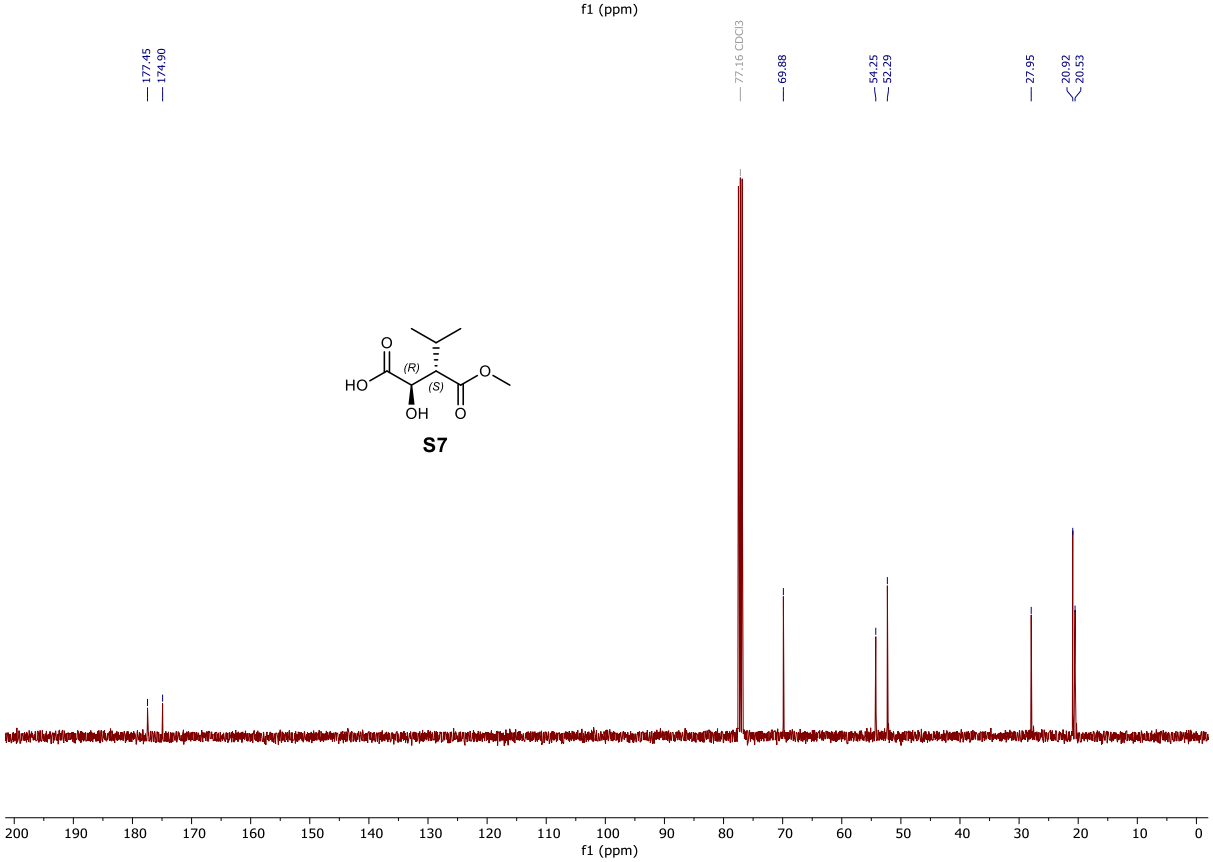
177.07
176.77
69.13
53.81
32.73
26.83
15.37
10.45

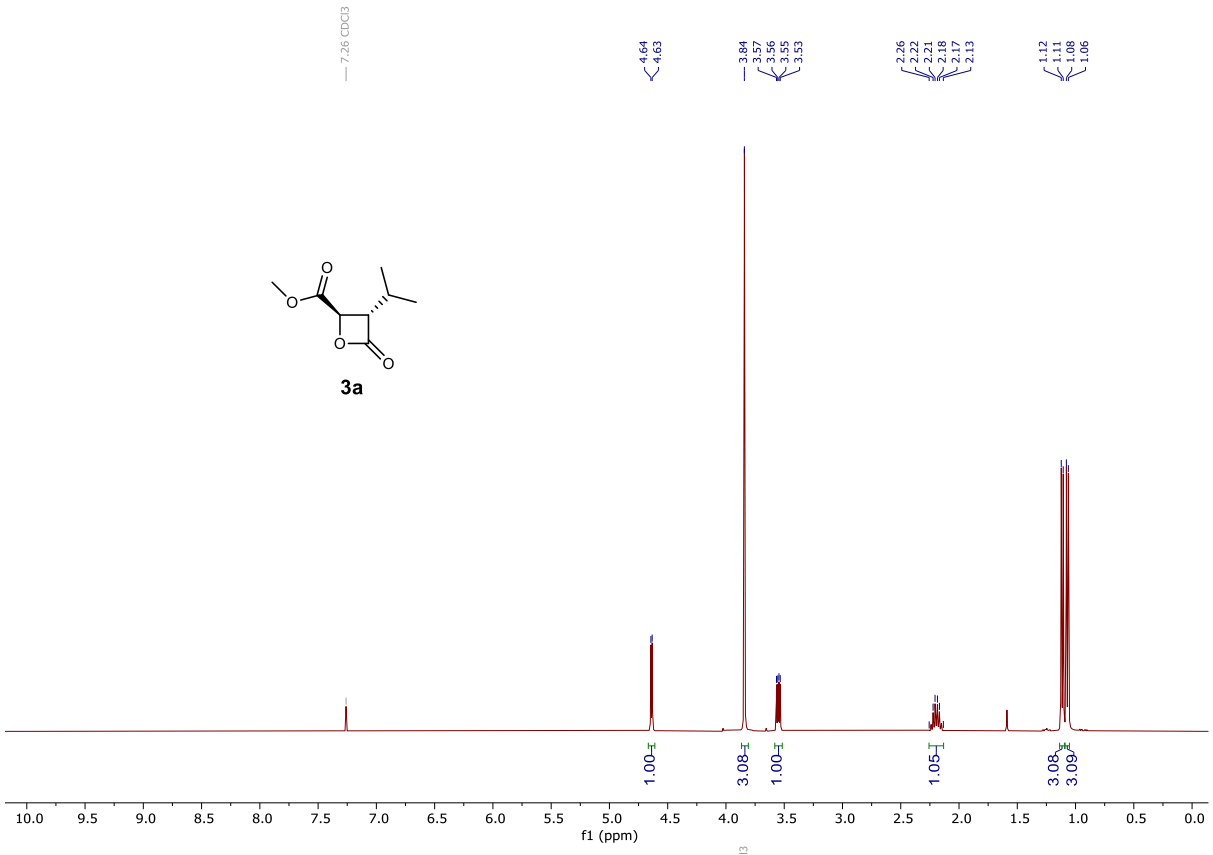

HO
O
(R)
OH
(S)
(S)
OH
O

1b

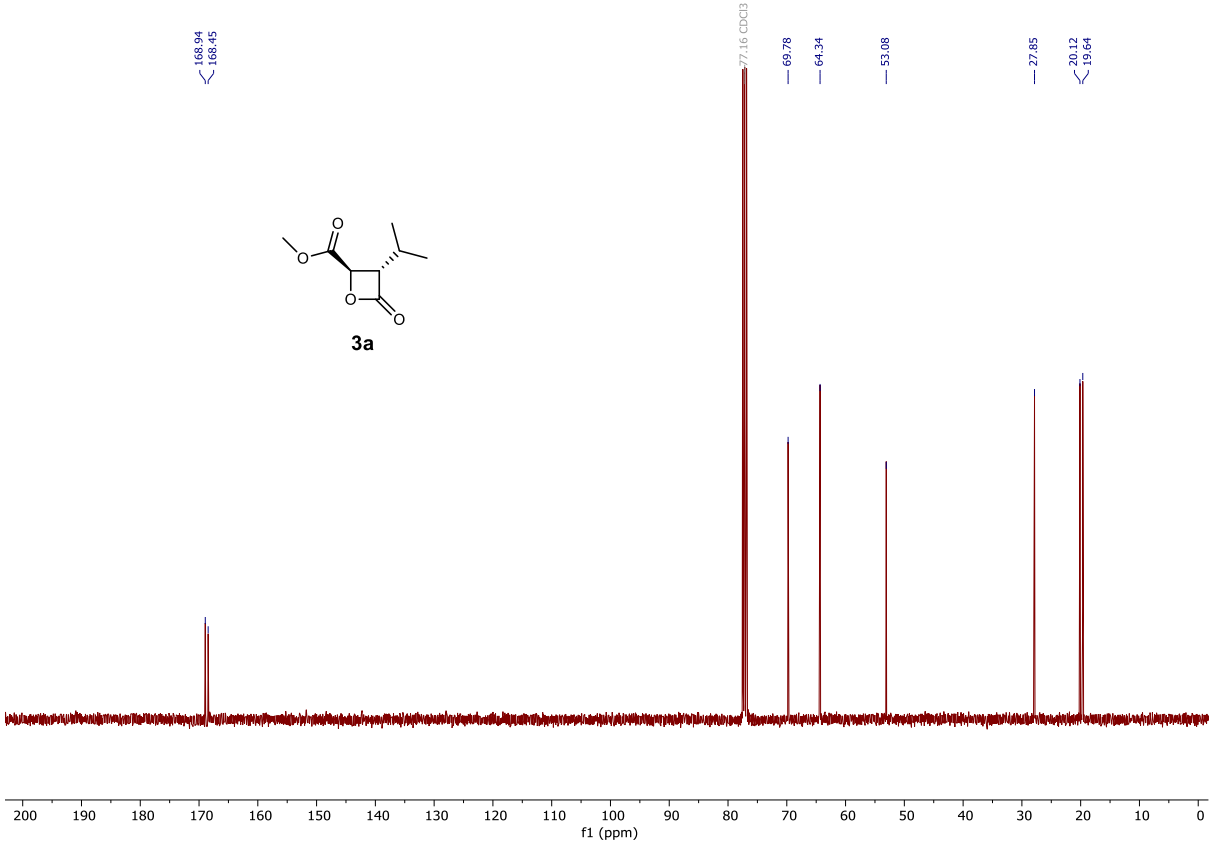

f1 (ppm)


7.26 CDCl3
4.44
4.44
3.72
2.73
2.73
2.71
2.71
2.28
2.26
2.24
2.22
2.19
1.08
1.07
1.01
1.00
HO
(R)
(S)
OH
O
O
O
S7
1.00
2.93
1.02
1.12
2.99
3.01
f1 (ppm)

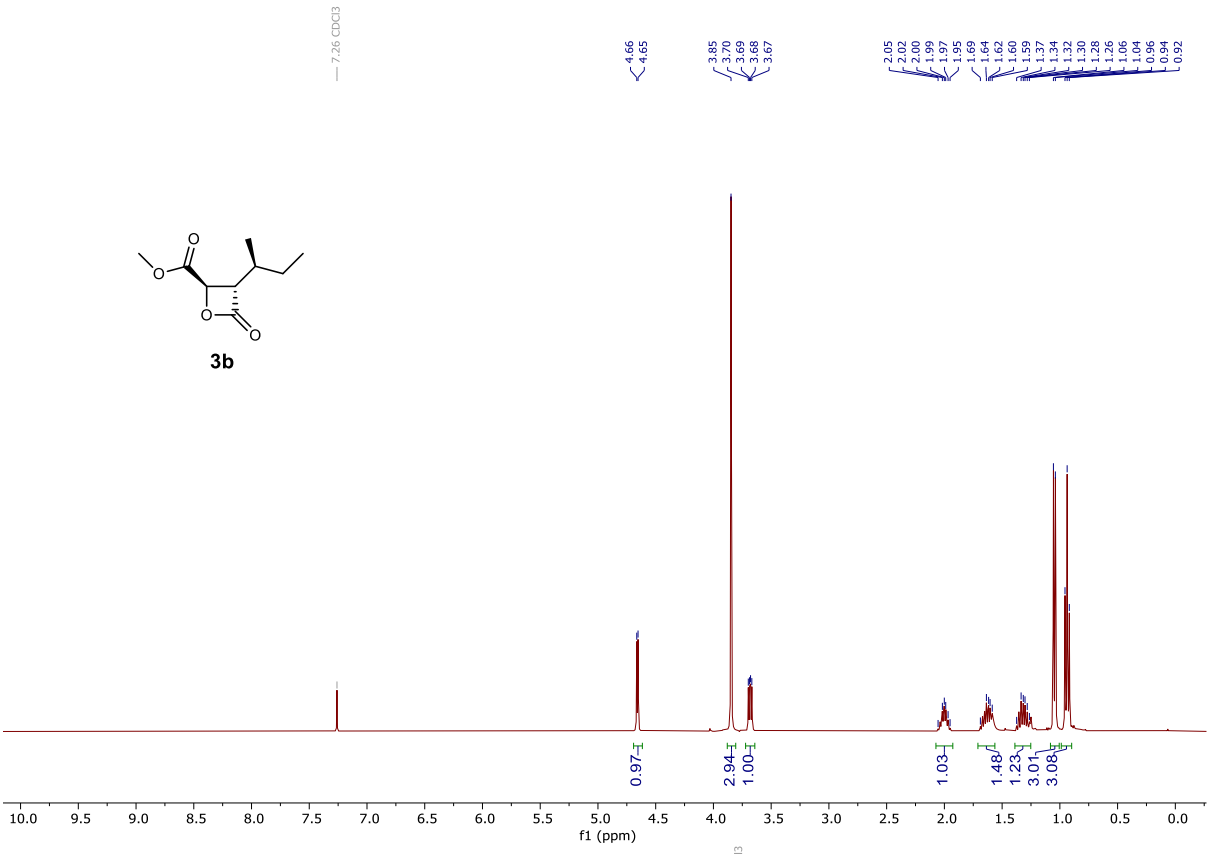

177.45
174.90
77.16 CDCl3
69.88
54.25
52.29
27.95
20.92
20.53
HO
(R)
(S)
OH
O
O
O
S7
f1 (ppm)


7.26 CDCl3
4.64
4.63
3.84
3.57
3.56
3.55
3.53
2.26
2.22
2.21
2.18
2.17
2.13
1.12
1.11
1.08
1.06
1.00
3.08
1.00
1.05
3.08
3.09
f1 (ppm)

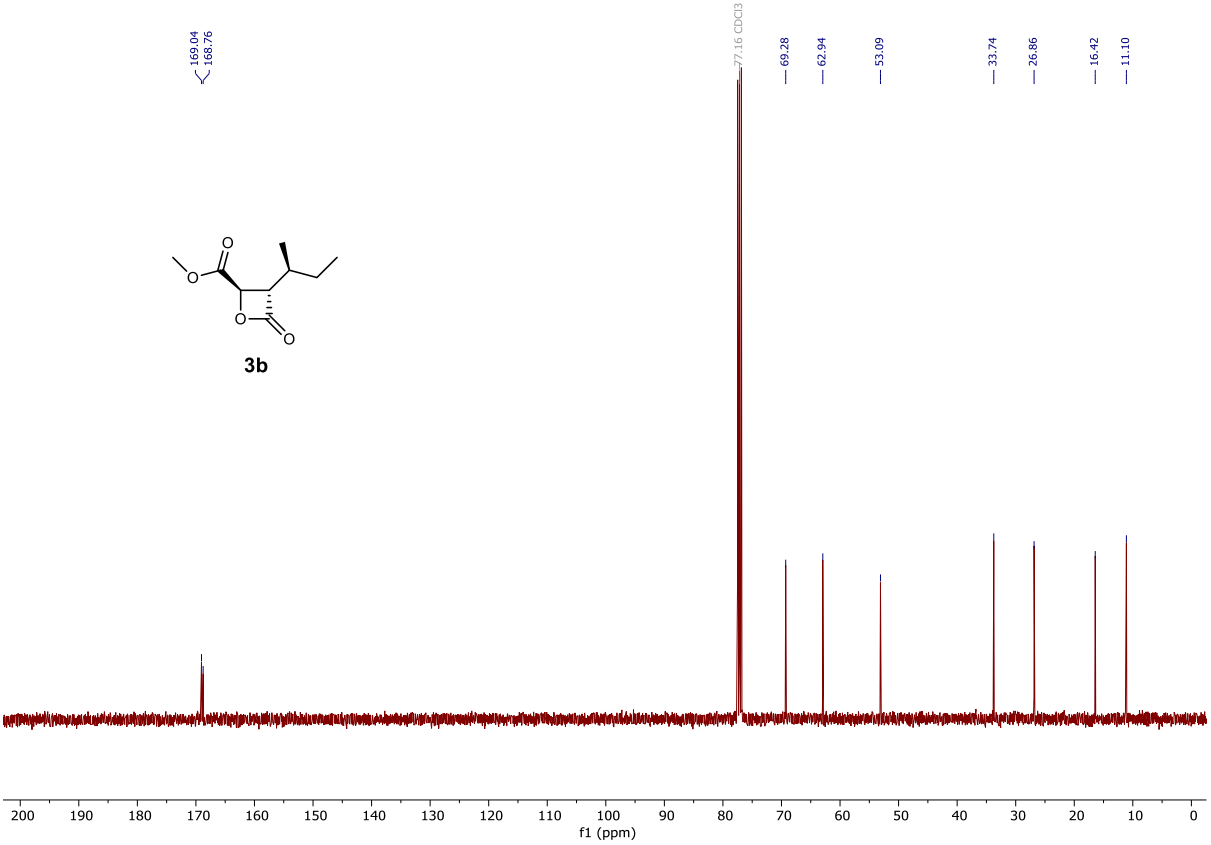
3a


168.94
168.45
77.16 CDCl3
69.78
64.34
53.08
27.85
20.12
19.64
f1 (ppm)

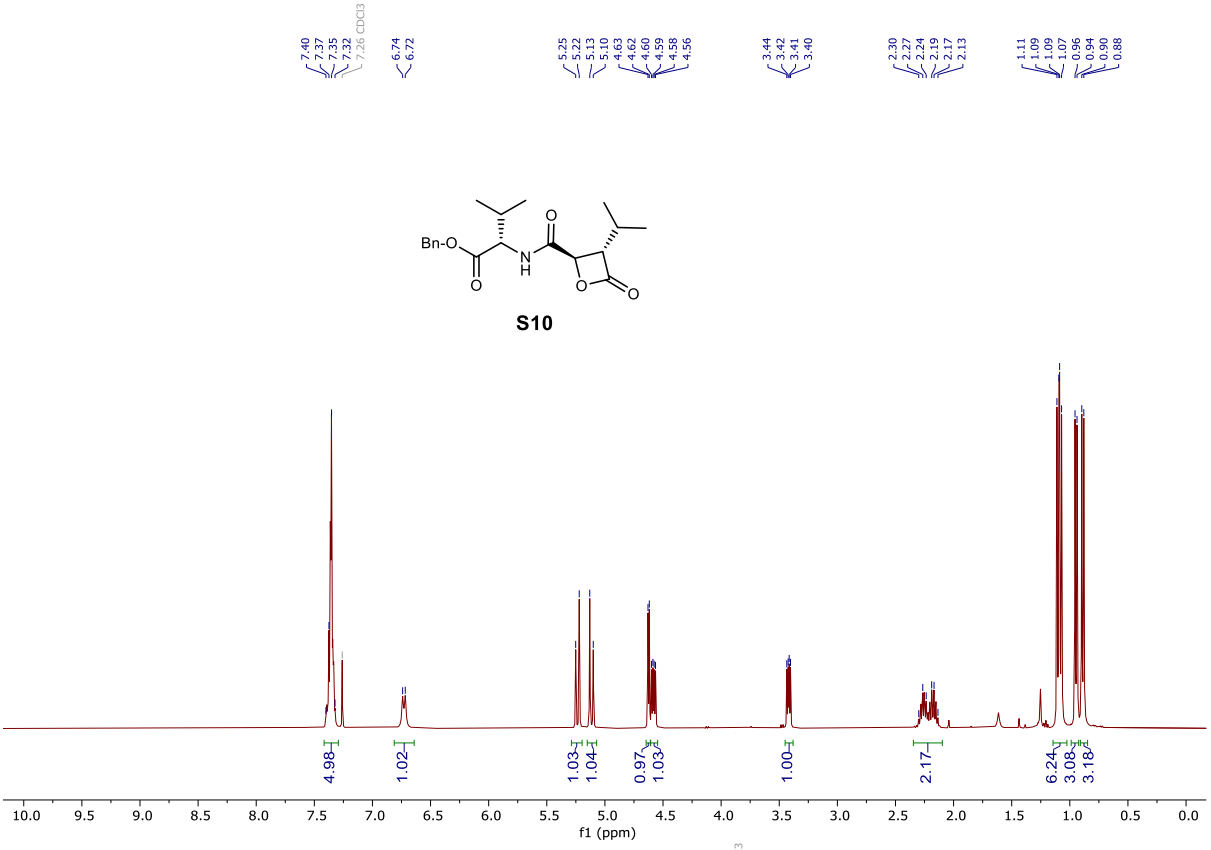
3a

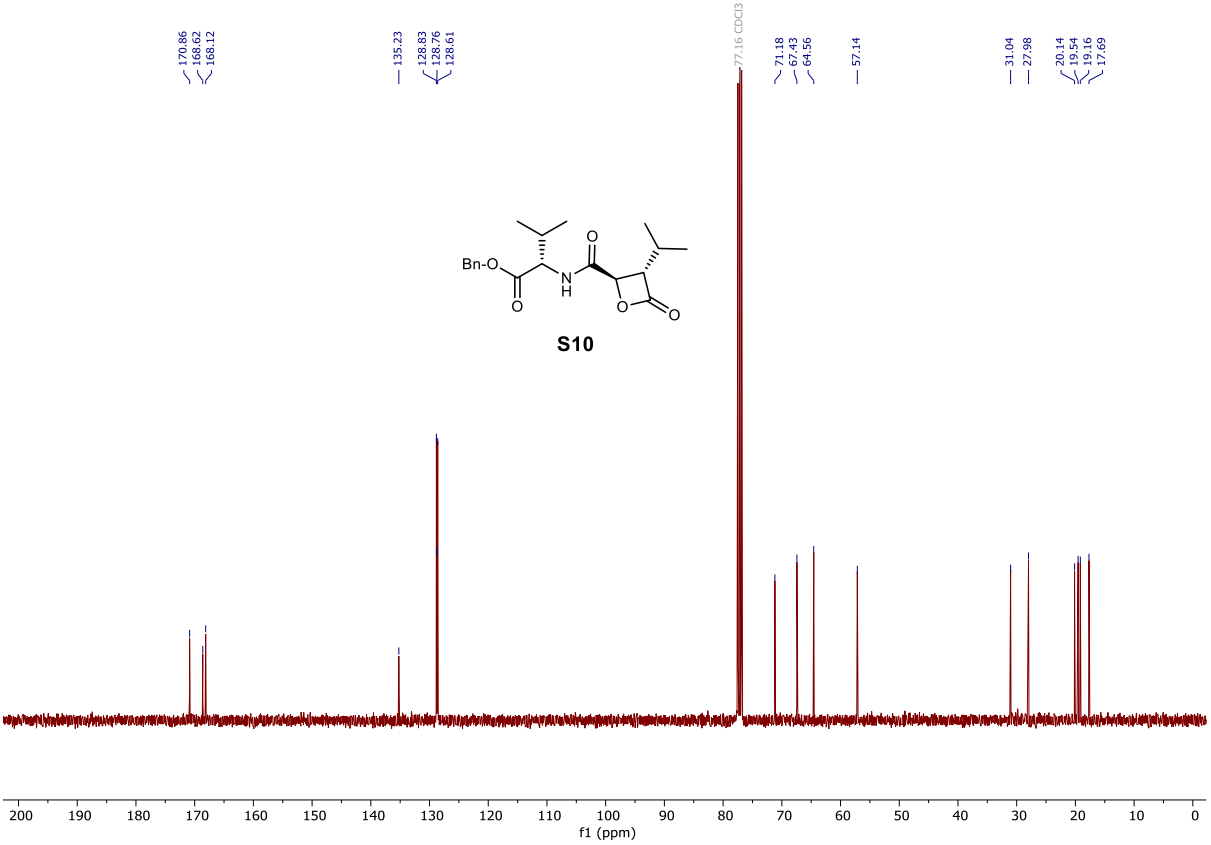

7.26 CDCl3
4.66
4.65
3.85
3.70
3.69
3.68
3.67
2.05
2.02
2.00
1.99
1.97
1.95
1.69
1.64
1.62
1.60
1.59
1.37
1.34
1.32
1.30
1.28
1.26
1.06
1.04
0.96
0.94
0.92
0.97
2.94
1.00
1.03
1.48
1.23
3.01
3.08
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

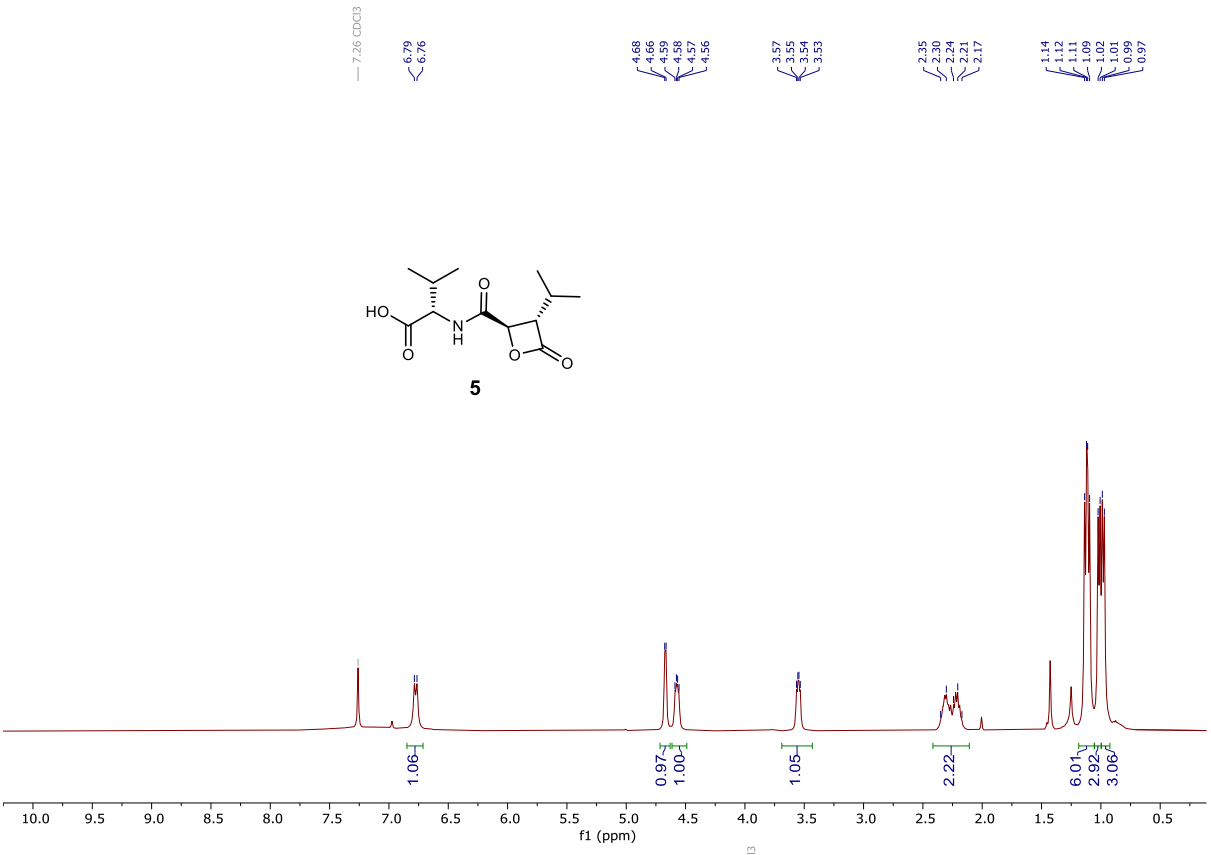
3b

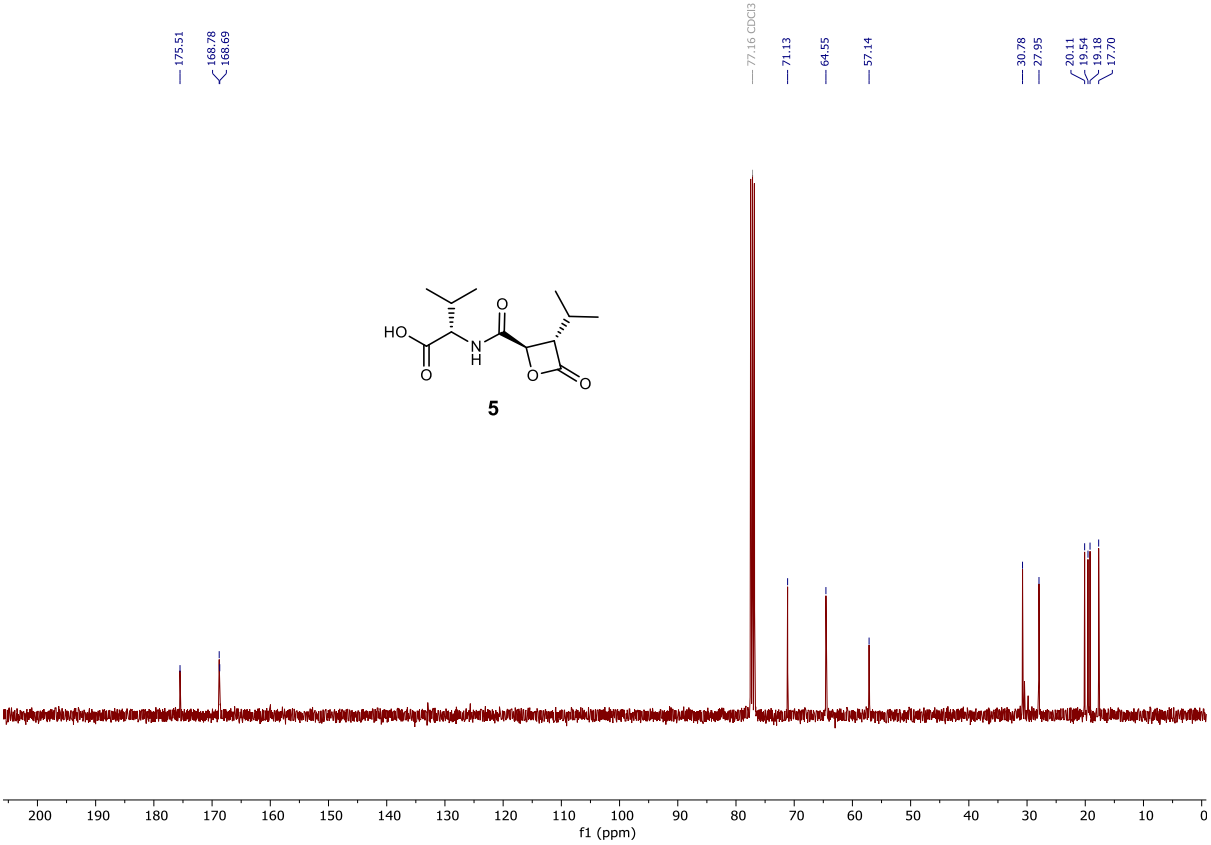

169.04
168.76
77.16 CDCl3
69.28
62.94
53.09
33.74
26.86
16.42
11.10
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

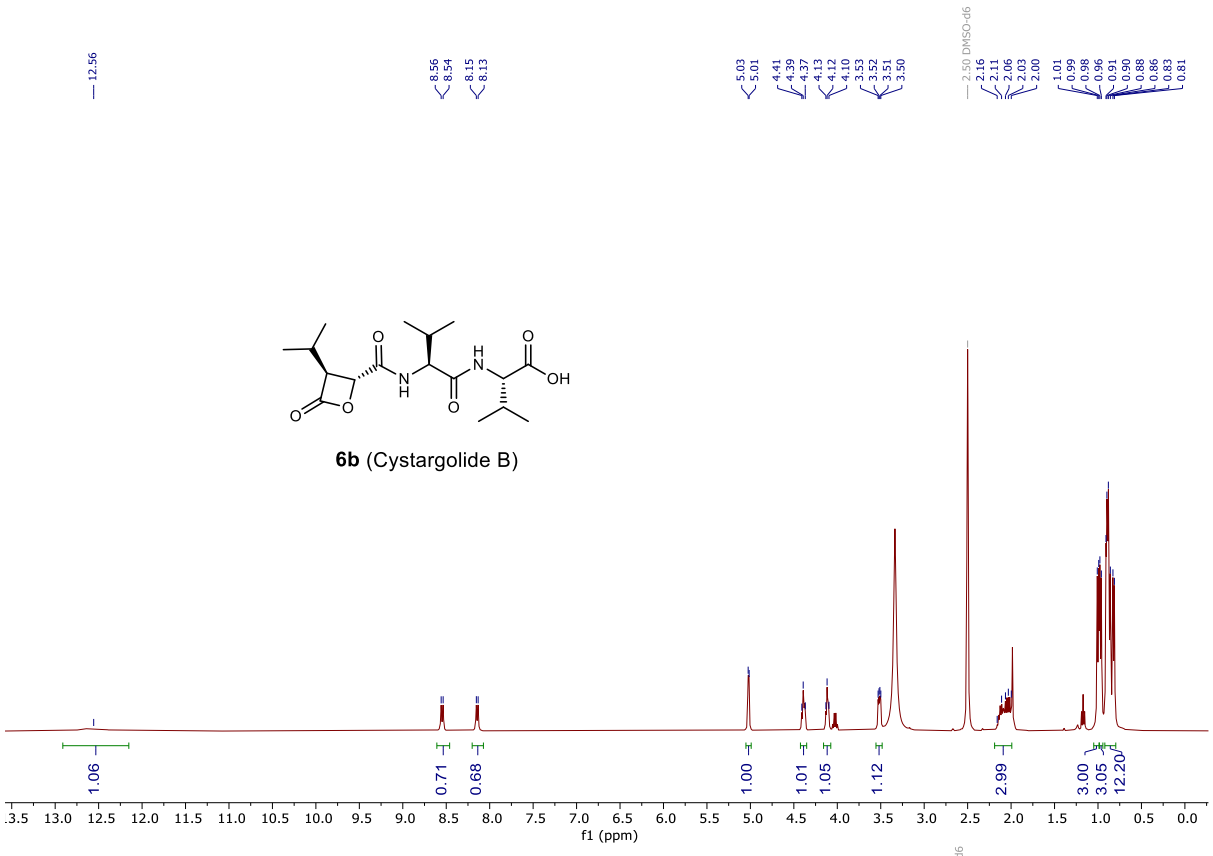
3b


7.40
7.37
7.35
7.32
7.26 CDCl3
6.74
6.72
5.25
5.22
5.13
5.10
4.63
4.62
4.60
4.59
4.58
4.56
3.44
3.42
3.41
3.40
2.30
2.27
2.24
2.19
2.17
2.13
1.11
1.09
1.09
1.07
0.96
0.94
0.90
0.88
Bn-O
O
N
H
O
O
O
S10
4.98
1.02
1.03
1.04
0.97
1.03
1.00
2.17
6.24
3.08
3.18
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

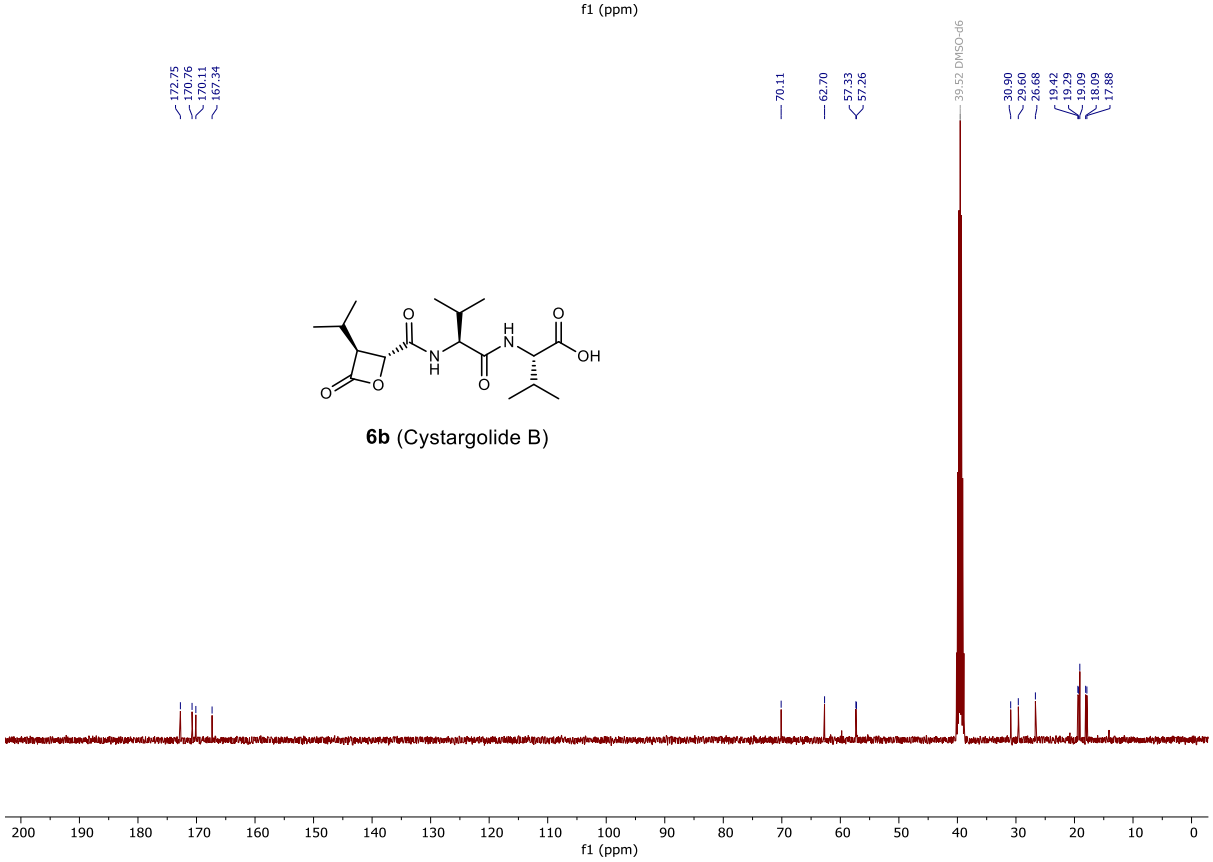

170.86
168.62
168.12
135.23
128.83
128.76
128.61
77.16 CDCl3
71.18
67.43
64.56
57.14
31.04
27.98
20.14
19.54
19.16
17.69
Bn-O
O
N
H
O
O
O
S10
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)


7.26 CDCl3
6.79
6.76
4.68
4.66
4.59
4.58
4.57
4.56
3.57
3.55
3.54
3.53
2.35
2.30
2.24
2.21
2.17
1.14
1.12
1.11
1.09
1.02
1.01
0.99
0.97
HO
O
N
H
O
O
O
O
5
1.06
0.97
1.00
1.05
2.22
6.01
2.92
3.06
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
f1 (ppm)

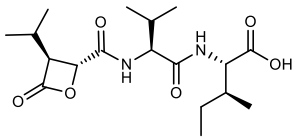

175.51
168.78
168.69
77.16 CDCl3
71.13
64.55
57.14
30.78
27.95
20.11
19.54
19.18
17.70
HO
O
N
H
O
O
O
O
5
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)



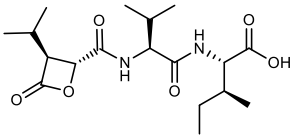
6b (Cystargolide B)



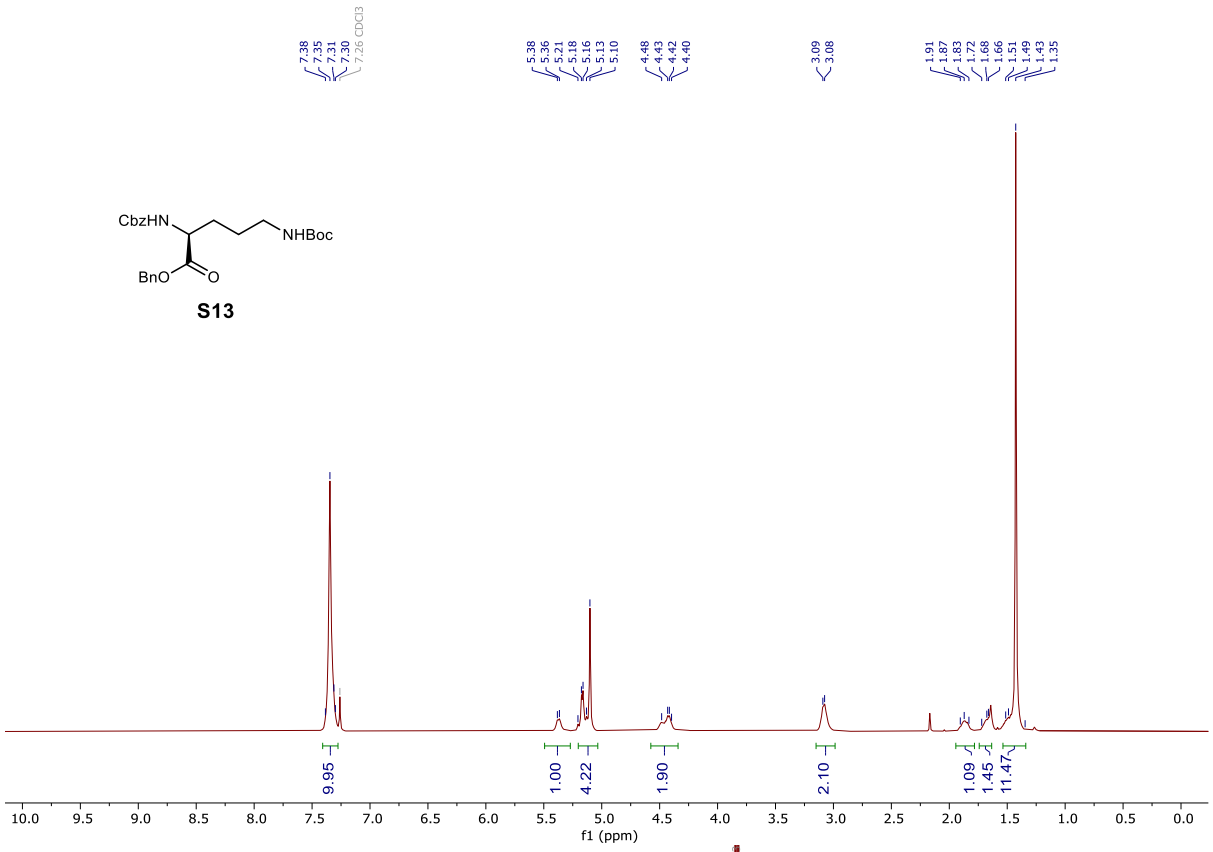
6b (Cystargolide B)

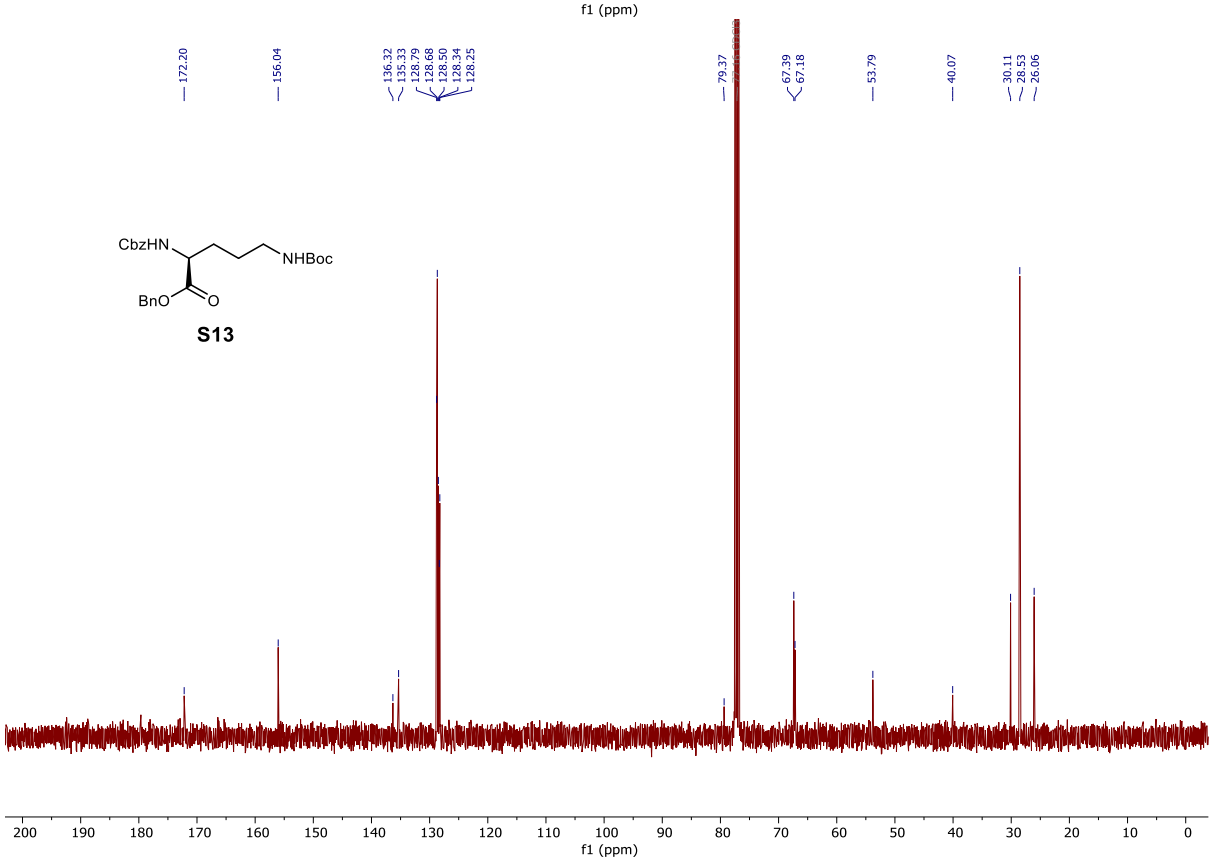


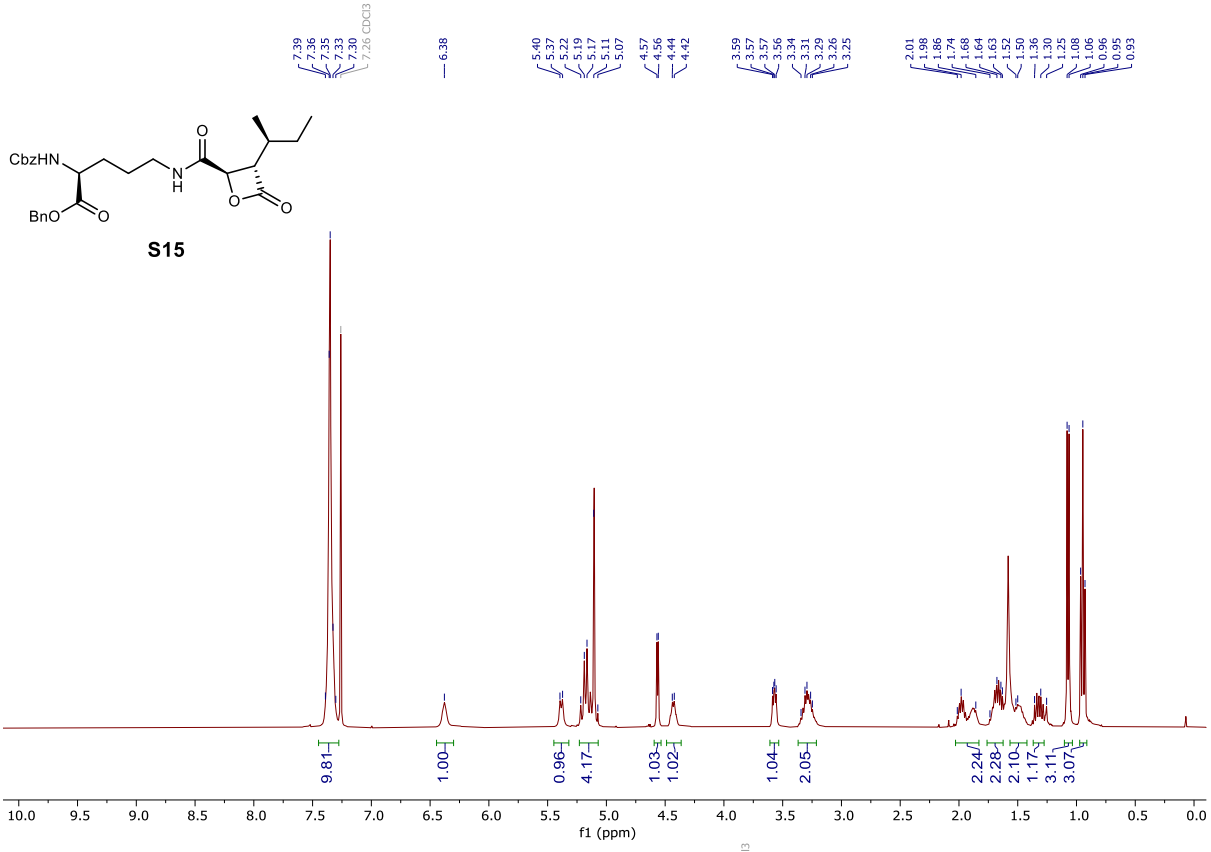
6a (Cystargolide A)

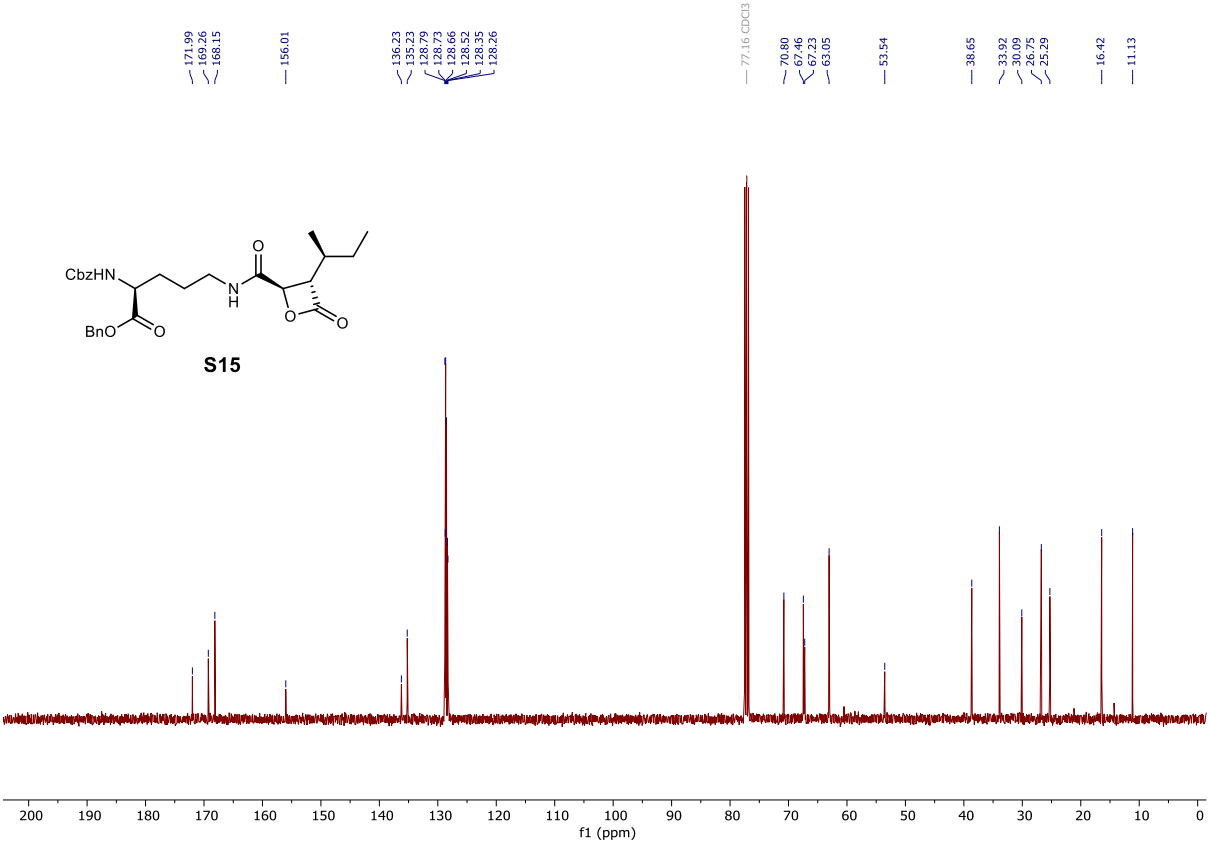


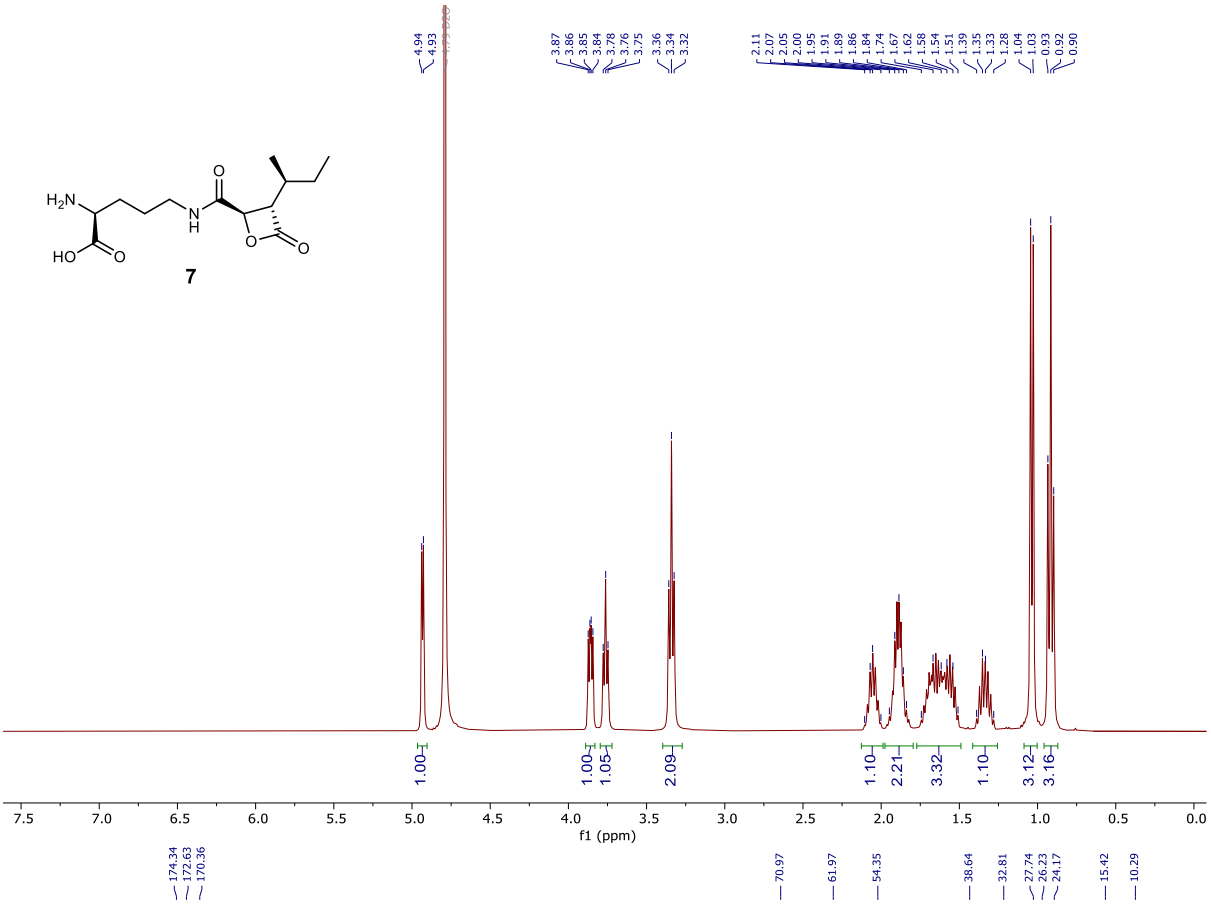
6a (Cystargolide A)

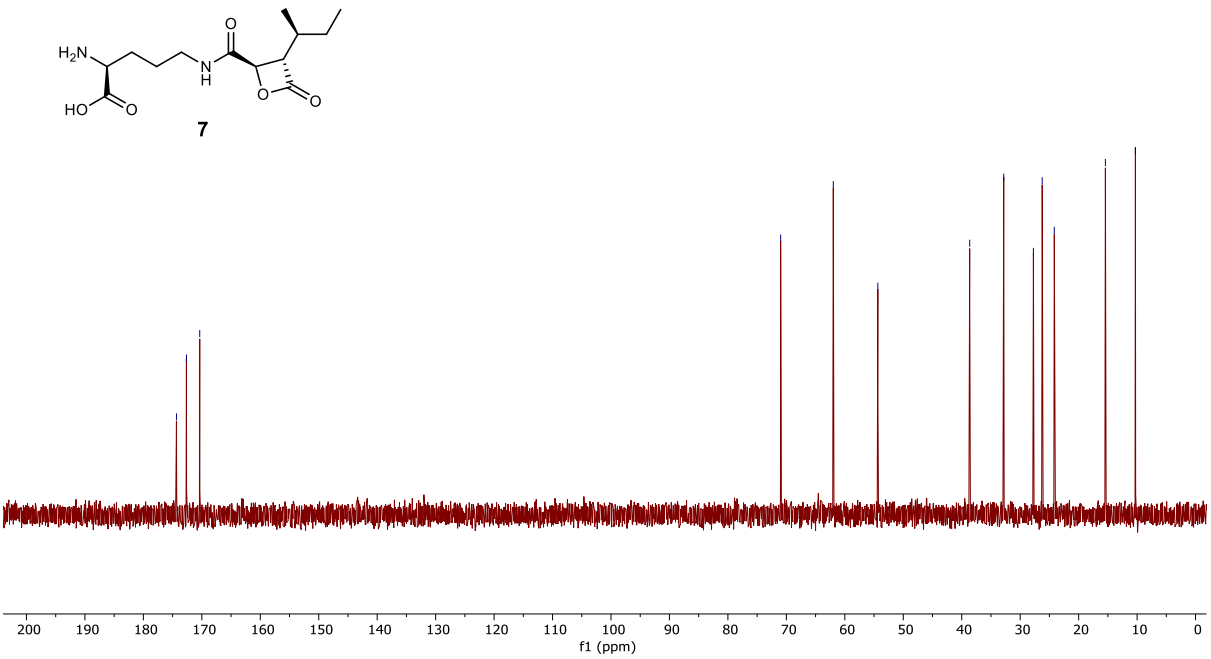

7.38
7.35
7.31
7.30
7.26 CDCl3
5.38
5.36
5.21
5.18
5.16
5.13
5.10
4.48
4.43
4.42
4.40
3.09
3.08
1.91
1.87
1.83
1.72
1.68
1.66
1.51
1.49
1.43
1.35
CbzHN
NHBoc
BnO
O
S13
9.95
1.00
4.22
1.90
2.10
1.09
1.45
11.47
f1 (ppm)

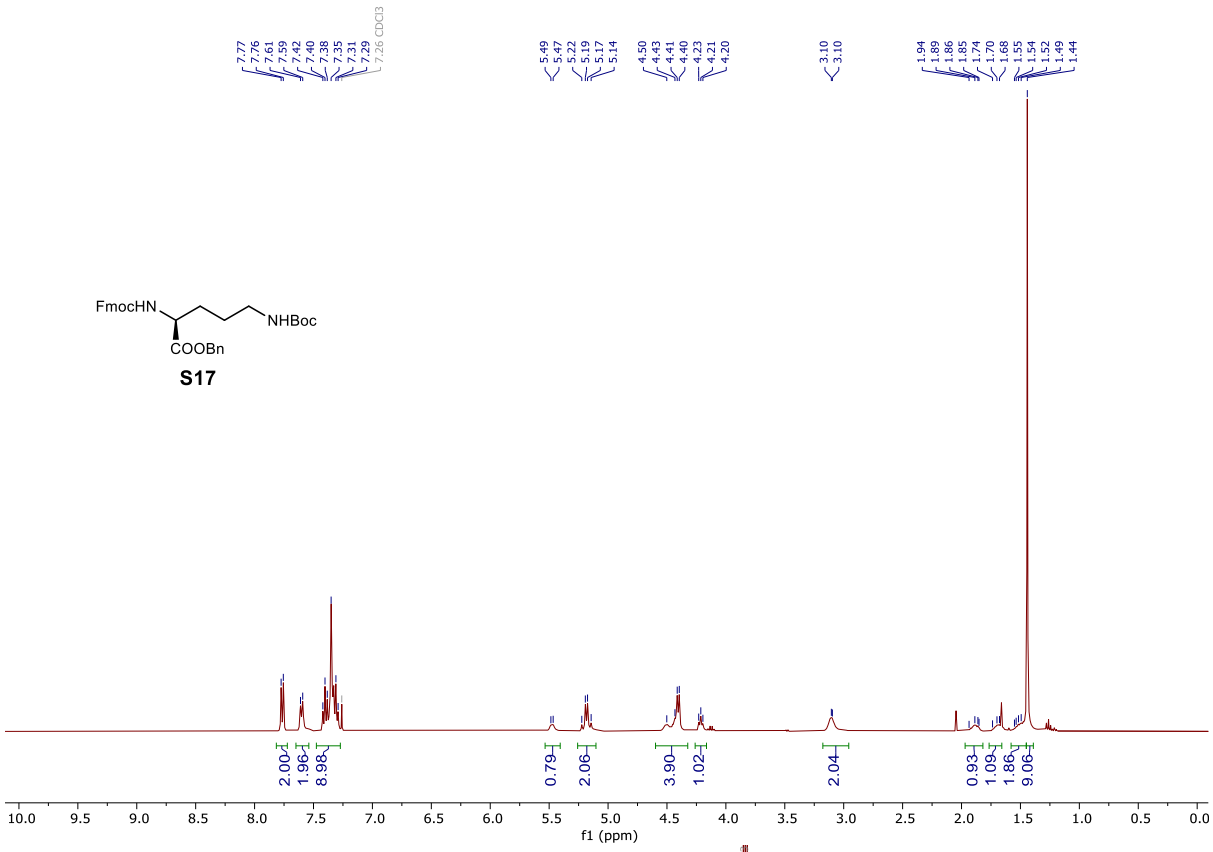

172.20
156.04
136.32
135.33
128.79
128.68
128.50
128.34
128.25
79.37
67.39
67.18
53.79
40.07
30.11
28.53
26.06
CbzHN
NHBoc
BnO
O
S13
f1 (ppm)

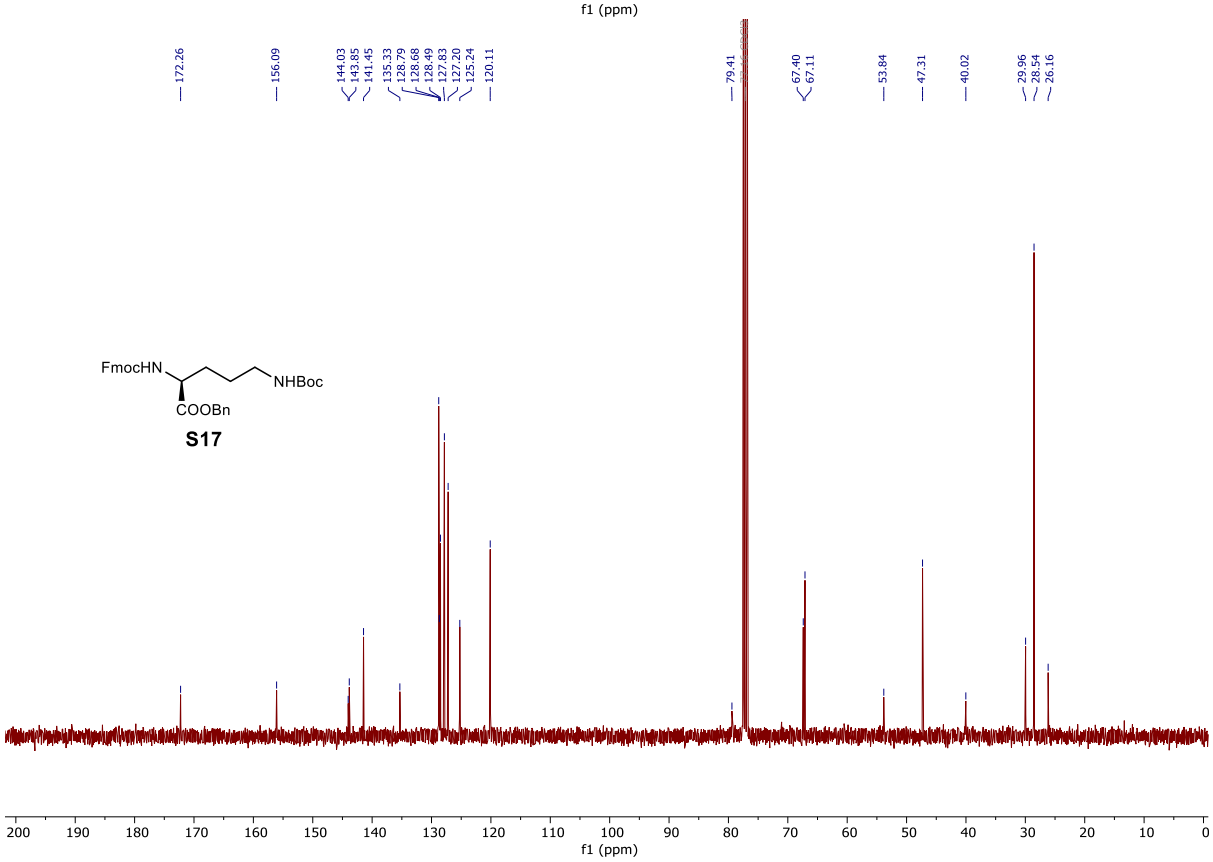

7.39
7.36
7.35
7.33
7.30
7.26 CDCl3
6.38
5.40
5.37
5.22
5.19
5.17
5.11
5.07
4.57
4.56
4.44
4.42
3.59
3.57
3.57
3.56
3.34
3.31
3.29
3.26
3.25
2.01
1.98
1.86
1.74
1.68
1.64
1.63
1.52
1.50
1.36
1.30
1.25
1.08
1.06
0.96
0.95
0.93
CbzHN
BnO
S15
9.81
1.00
0.96
4.17
1.03
1.02
1.04
2.05
2.24
2.28
2.10
1.17
3.11
3.07
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

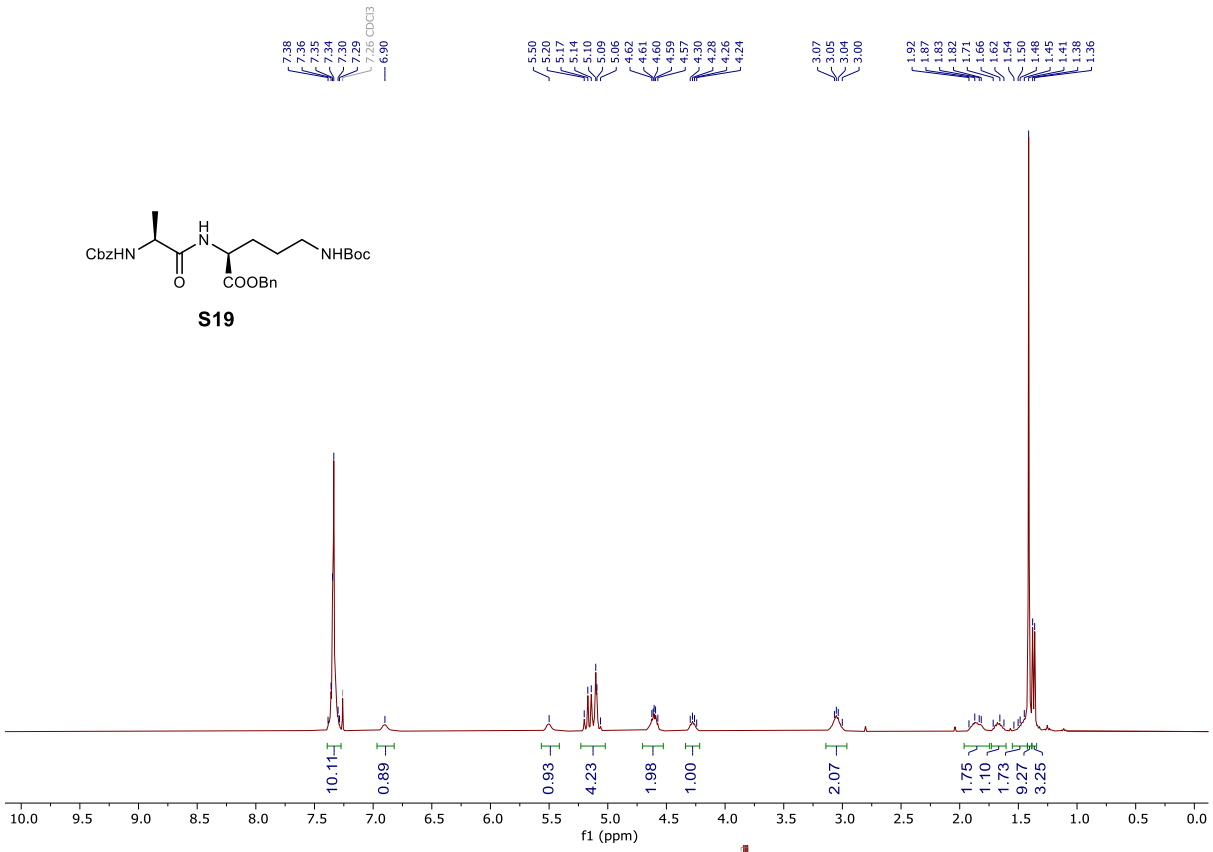

171.99
169.26
168.15
156.01
136.23
135.23
128.79
128.73
128.66
128.52
128.35
128.26
77.16 CDCl3
70.80
67.46
67.23
63.05
53.54
38.65
33.92
30.09
26.75
25.29
16.42
11.13
CbzHN
BnO
S15
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

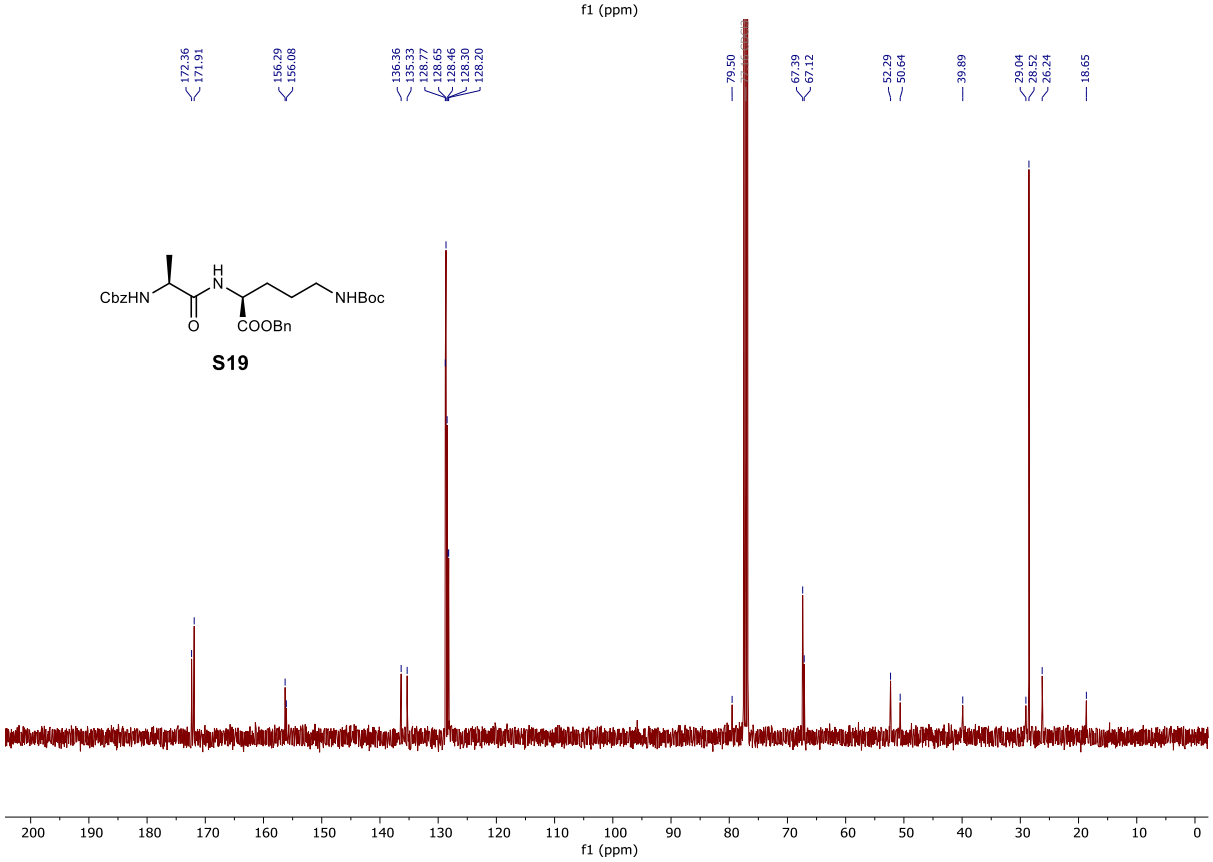

7
4.94
4.93
3.87
3.86
3.85
3.84
3.78
3.76
3.75
3.36
3.34
3.32
2.11
2.07
2.05
2.00
1.95
1.91
1.89
1.86
1.84
1.74
1.67
1.62
1.58
1.54
1.51
1.39
1.35
1.33
1.28
1.04
1.03
0.93
0.92
0.90
1.00
1.00
1.05
2.09
1.10
2.21
3.32
1.10
3.12
3.16
f1 (ppm)

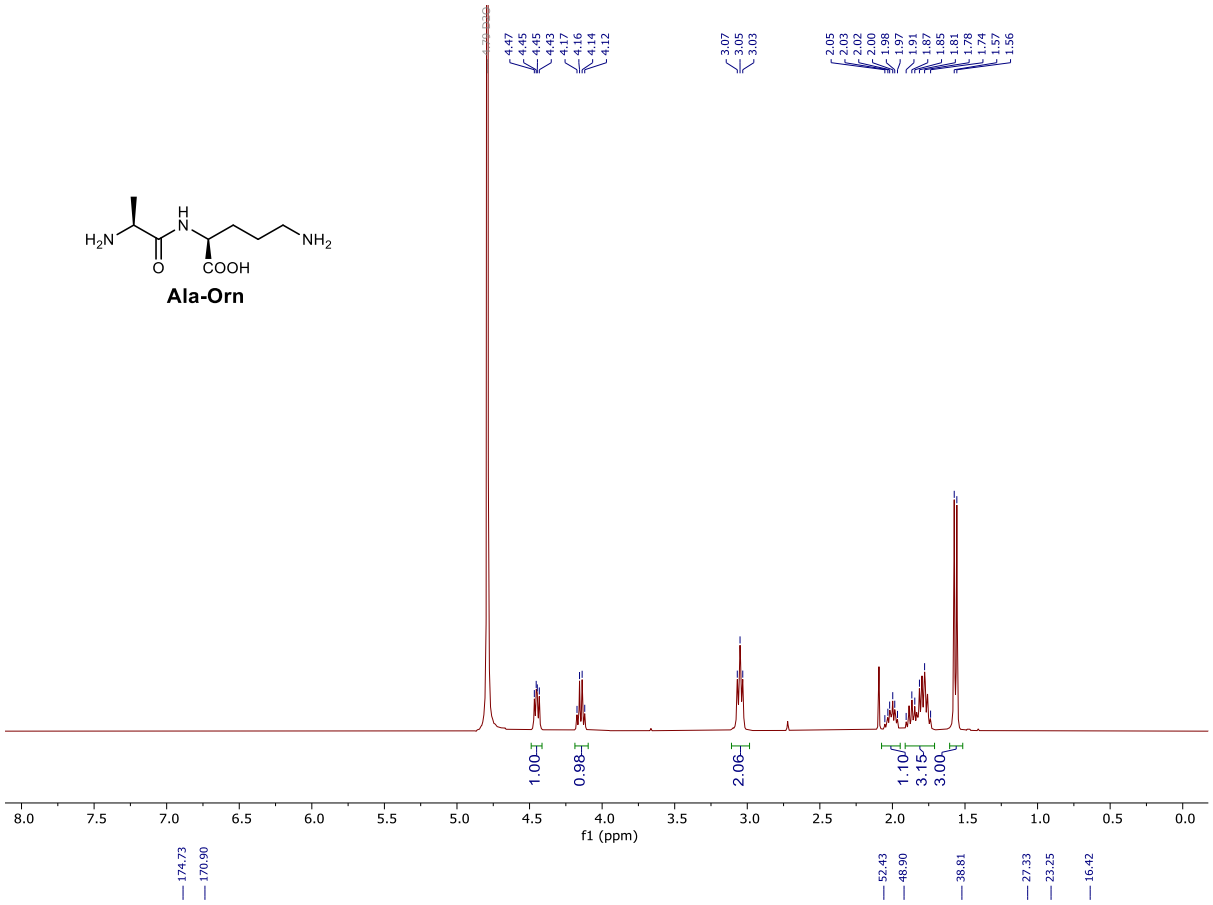

7
174.34
172.63
170.36
70.97
61.97
54.35
38.64
32.81
27.74
26.23
24.17
15.42
10.29
f1 (ppm)


FmocHN
NHBoc
COOBn
S17
f1 (ppm)

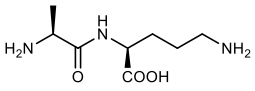

FmocHN
NHBoc
COOBn
S17
f1 (ppm)


7.38
7.36
7.35
7.34
7.30
7.29
7.26 CDCl3
6.90
5.50
5.20
5.17
5.14
5.10
5.09
5.06
4.62
4.61
4.60
4.59
4.57
4.30
4.28
4.26
4.24
3.07
3.05
3.04
3.00
1.92
1.87
1.83
1.82
1.71
1.66
1.62
1.54
1.50
1.48
1.45
1.41
1.38
1.36
CbzHN
H
N
O
COOBn
NHBoc
S19
10.11
0.89
0.93
4.23
1.98
1.00
2.07
1.75
1.10
1.73
9.27
3.25
f1 (ppm)

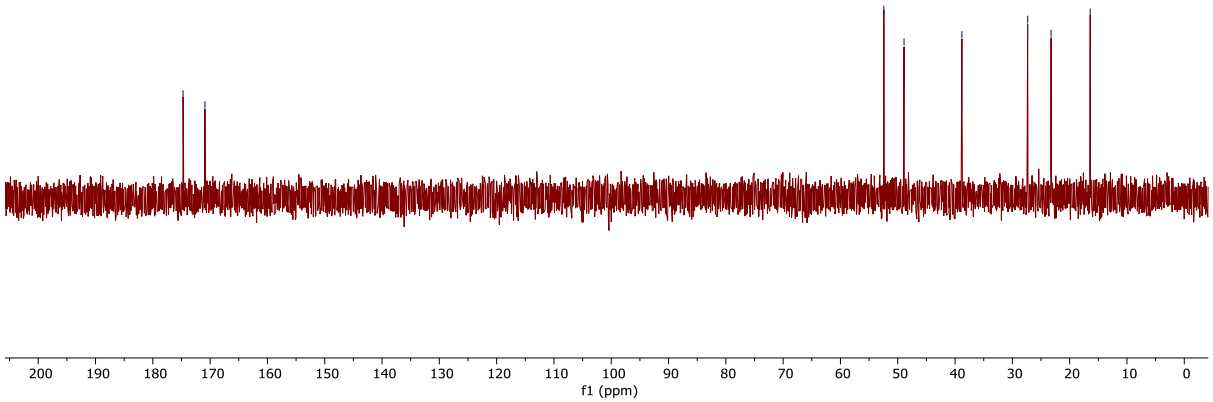

172.36
171.91
156.29
156.08
136.36
135.33
128.77
128.65
128.46
128.30
128.20
79.50
67.39
67.12
52.29
50.64
39.89
29.04
28.52
26.24
18.65
CbzHN
H
N
O
COOBn
NHBoc
S19
f1 (ppm)

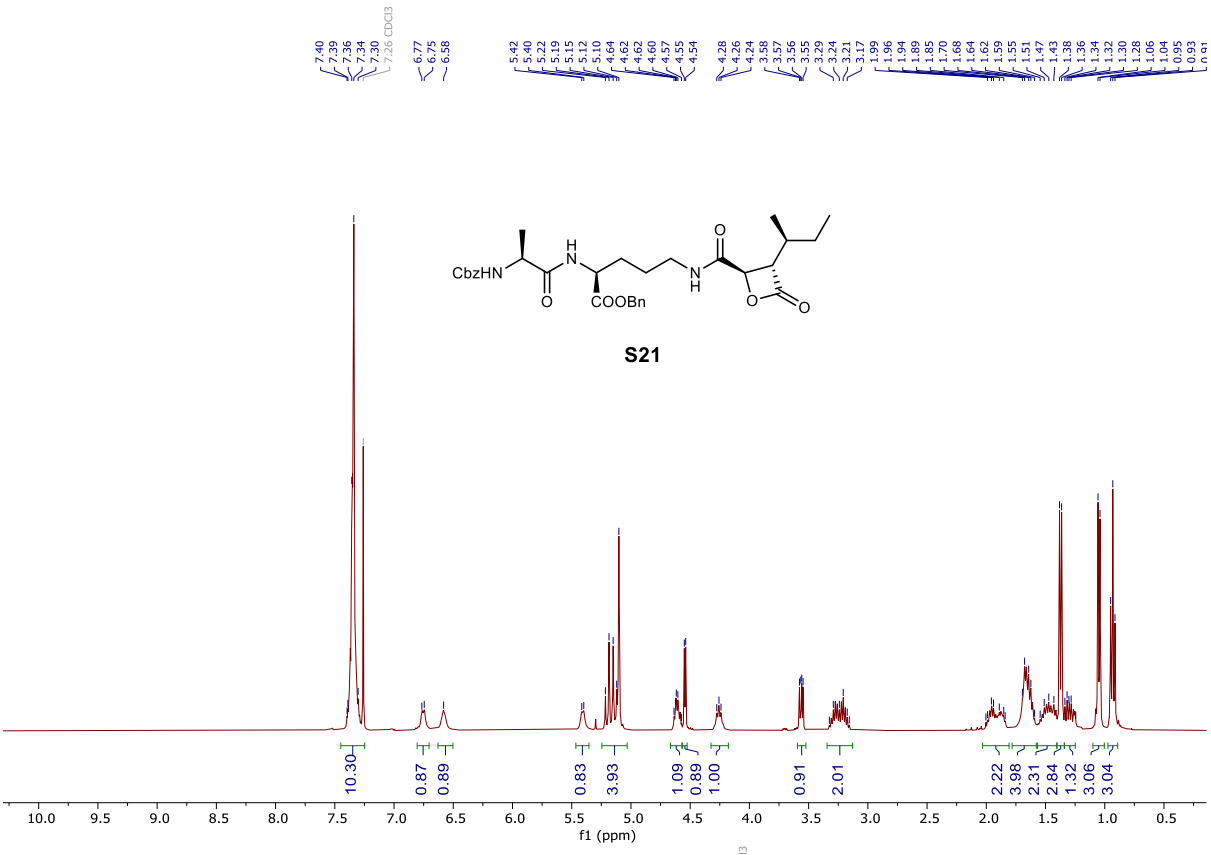


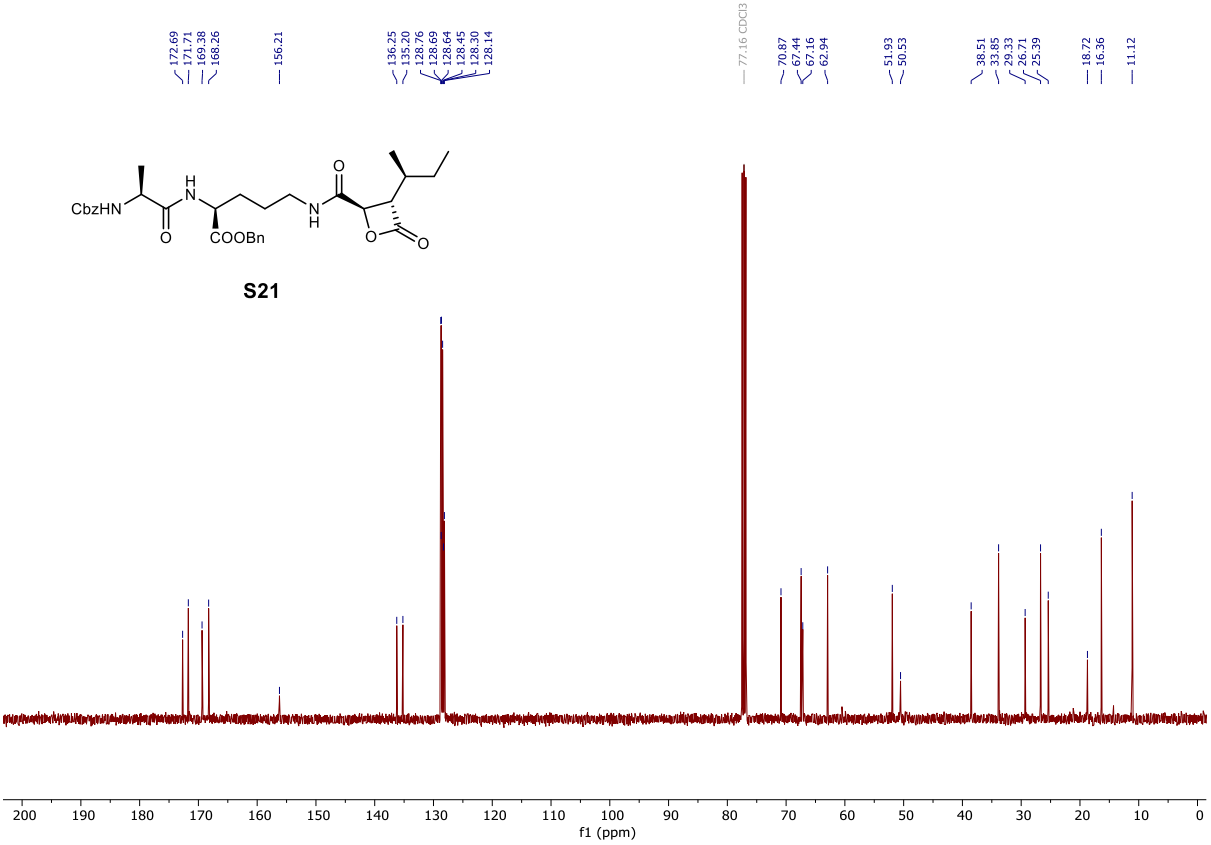
Ala-Orn

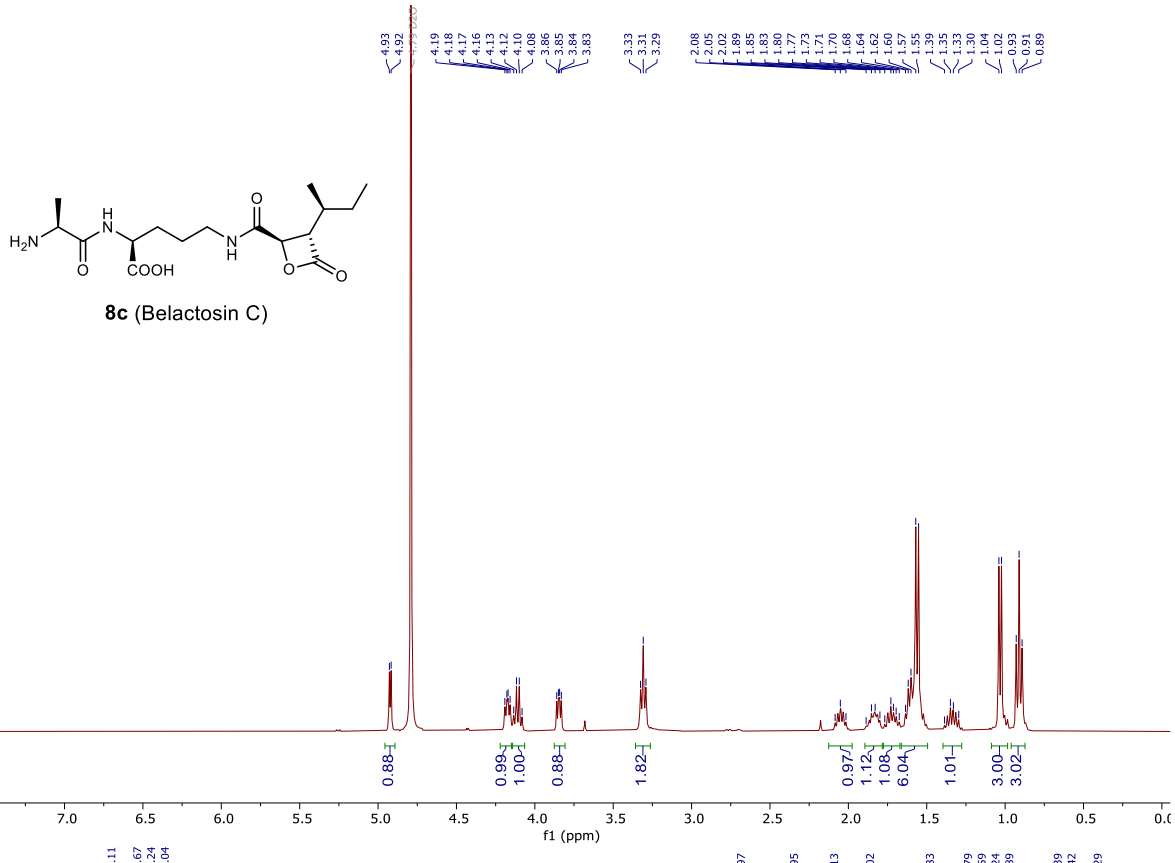


Ala-Orn

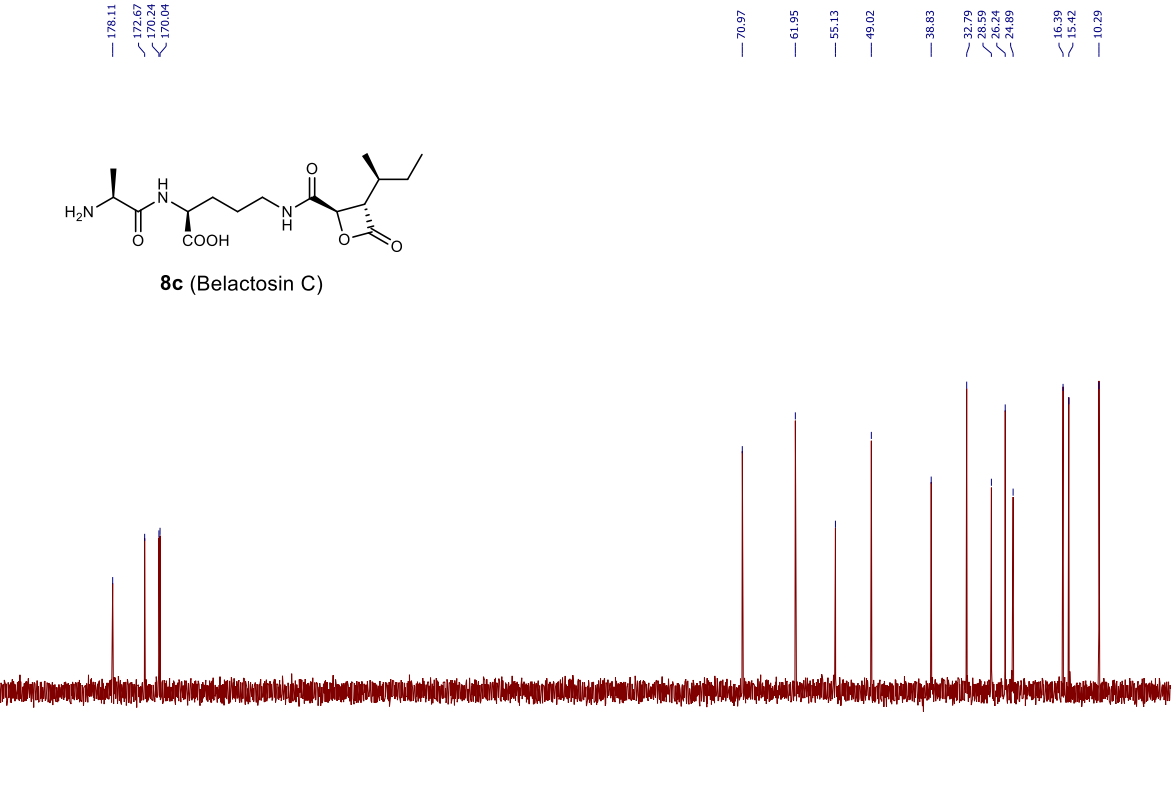



CbzHN
H
N
O
COOBn
N
H
O
O
O
S21
10.30
0.87
0.89
0.83
3.93
1.09
0.89
1.00
0.91
2.01
2.22
3.98
2.31
2.84
1.32
3.06
3.04
f1 (ppm)

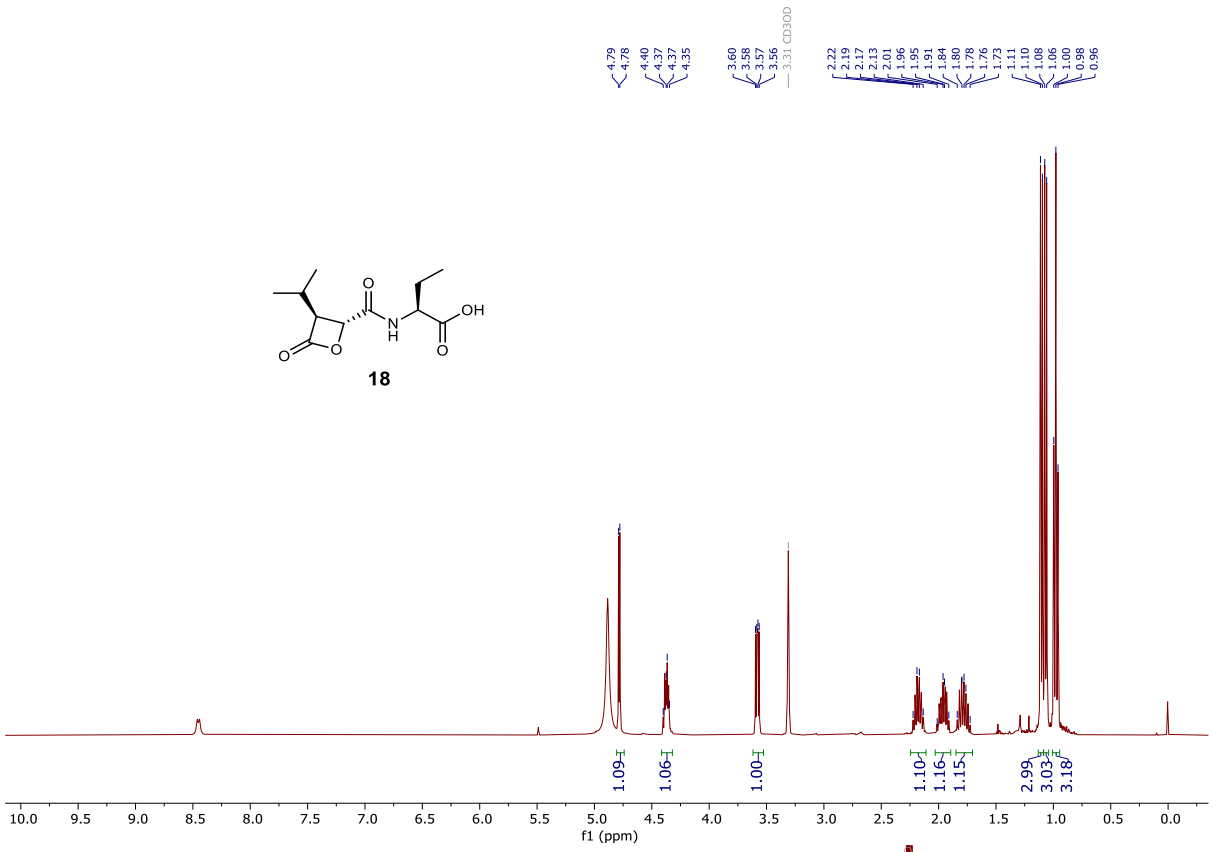

CbzHN
H
N
O
COOBn
N
H
O
O
O
S21
f1 (ppm)

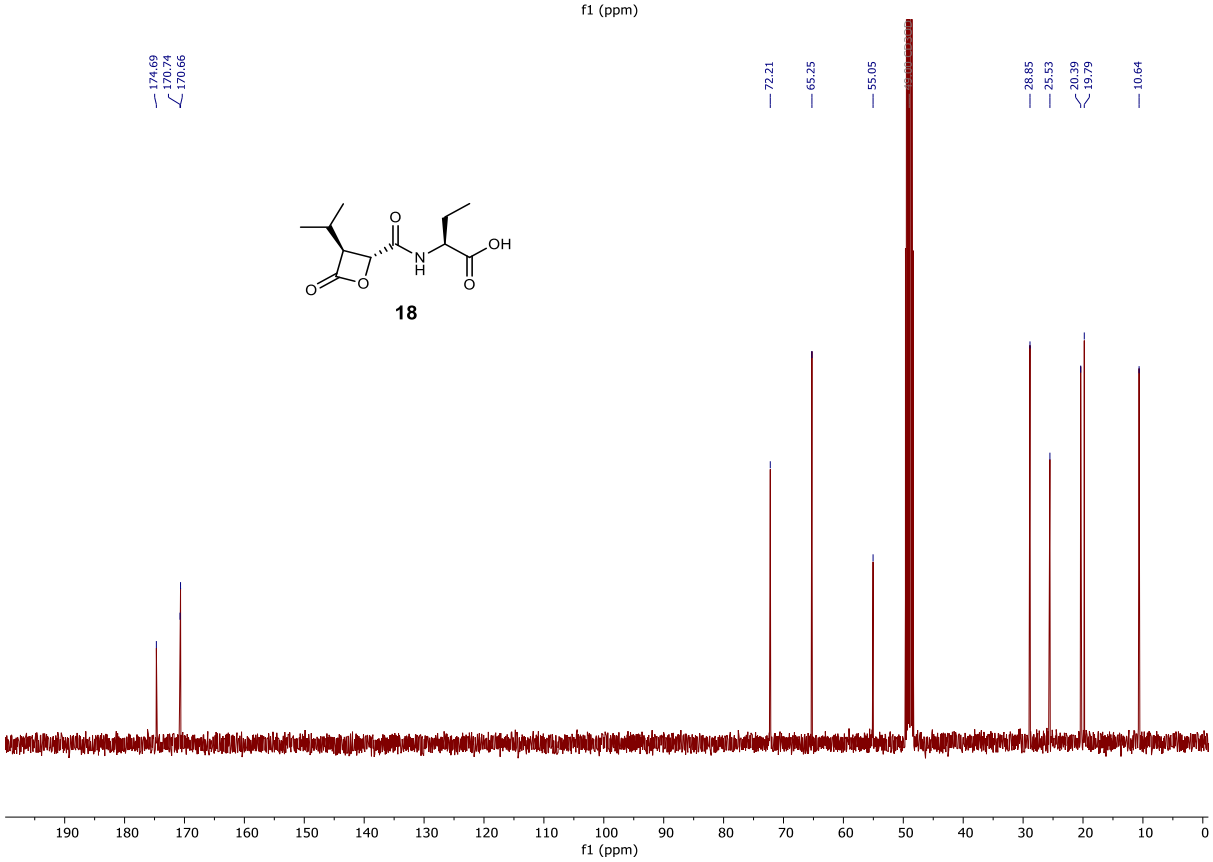


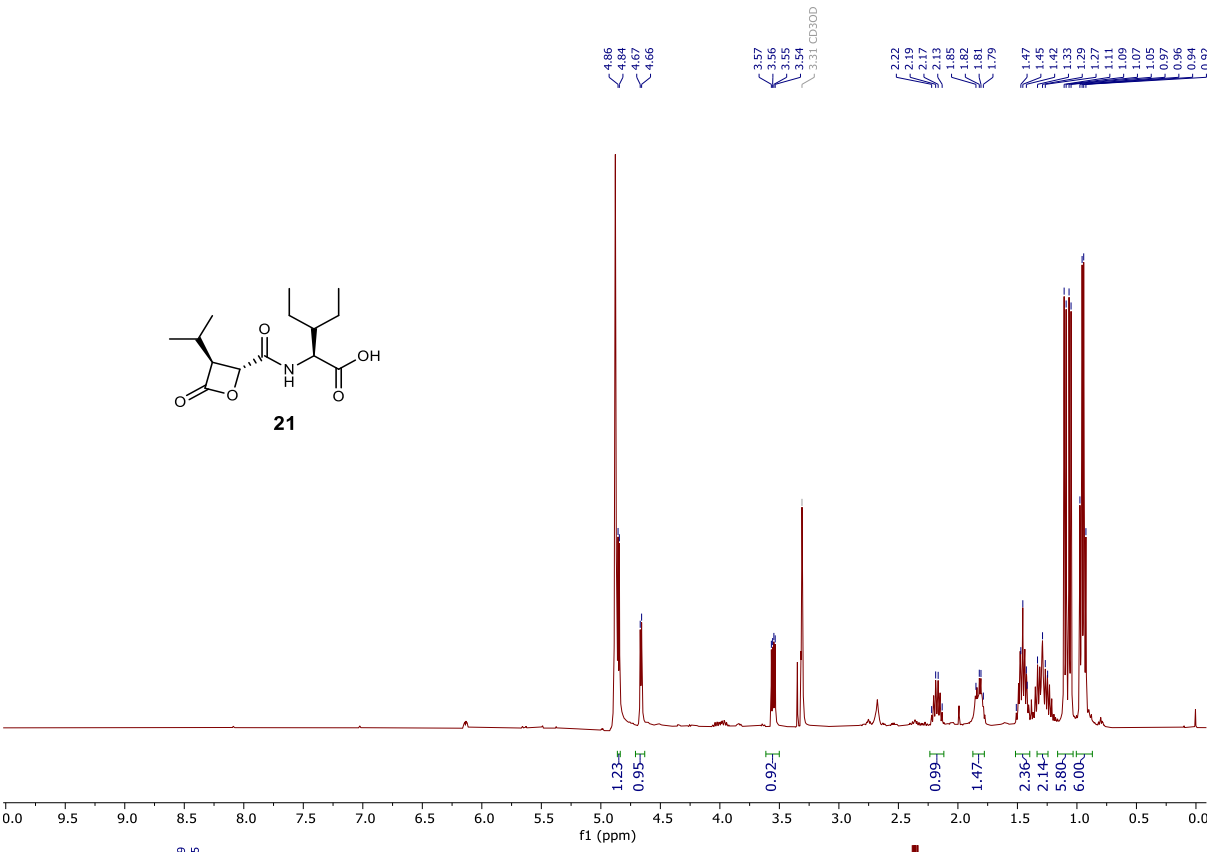
8c (Belactosin C)

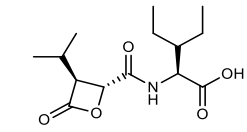


8c (Belactosin C)


4.79
4.78
4.40
4.37
4.37
4.35
3.60
3.58
3.57
3.56
3.31 CD3OD
2.22
2.19
2.17
2.13
2.01
1.96
1.95
1.91
1.84
1.80
1.78
1.76
1.73
1.11
1.10
1.08
1.06
1.00
0.98
0.96
O
N
H
OH
O
O
18
1.09
1.06
1.00
1.10
1.16
1.15
2.99
3.03
3.18
10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)


174.69
170.74
170.66
72.21
65.25
55.05
28.85
25.53
20.39
19.79
10.64
O
N
H
OH
O
O
18
190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

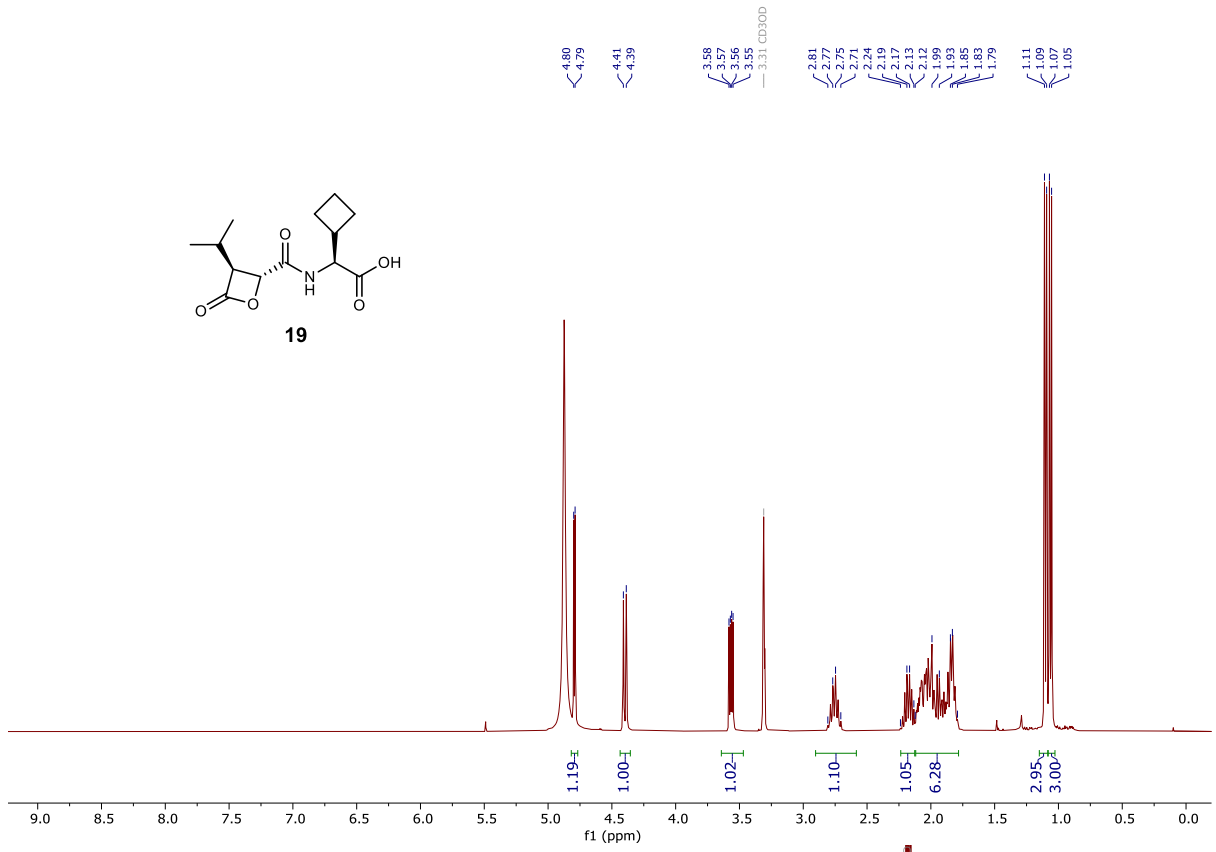

4.86
4.84
4.67
4.66
3.57
3.56
3.55
3.54
3.31 CD3OD
2.22
2.19
2.17
2.13
1.85
1.82
1.81
1.79
1.47
1.45
1.42
1.33
1.29
1.27
1.11
1.09
1.07
1.05
0.97
0.96
0.94
1.23
0.95
0.92
0.99
1.47
2.36
2.14
5.80
6.00
f1 (ppm)



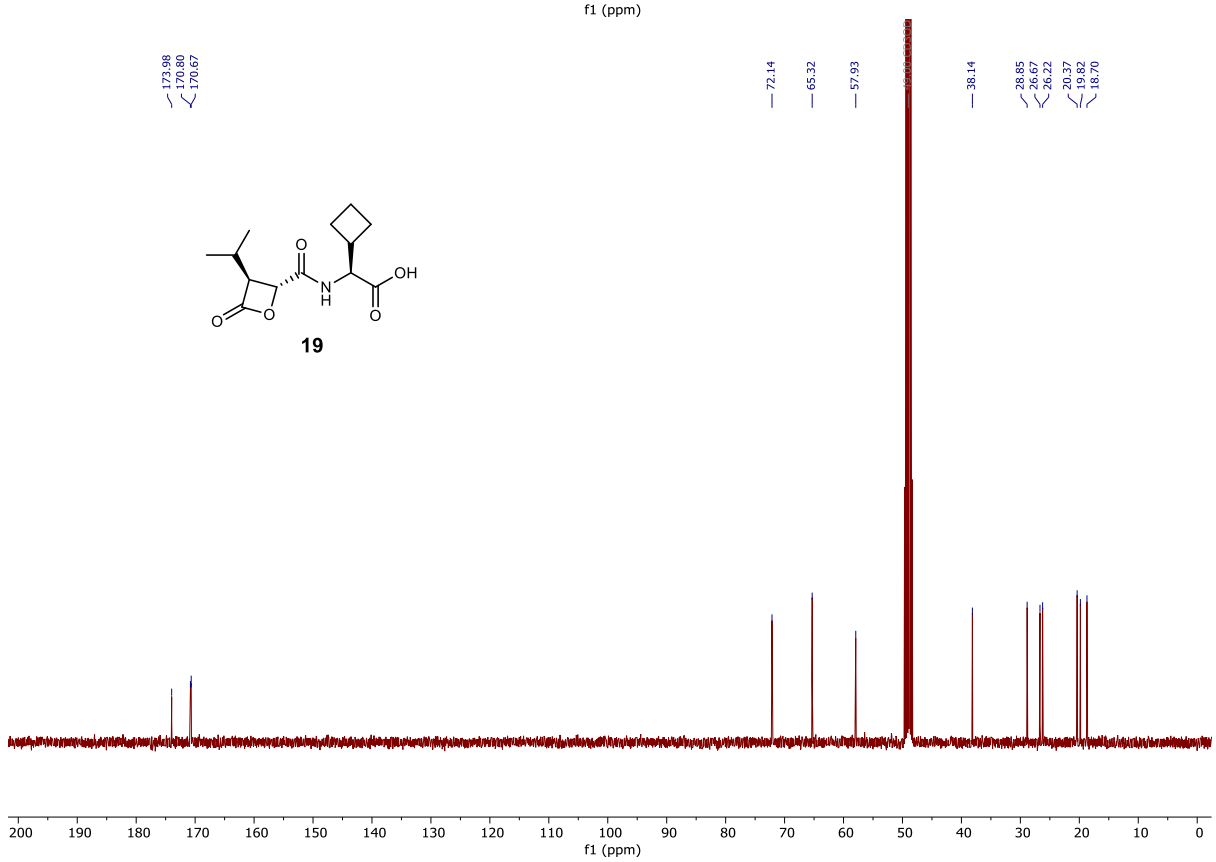
21

169.29
169.25
70.62
63.87
53.71
43.62
27.41
22.45
22.14
18.96
18.47
10.56
10.54
f1 (ppm)

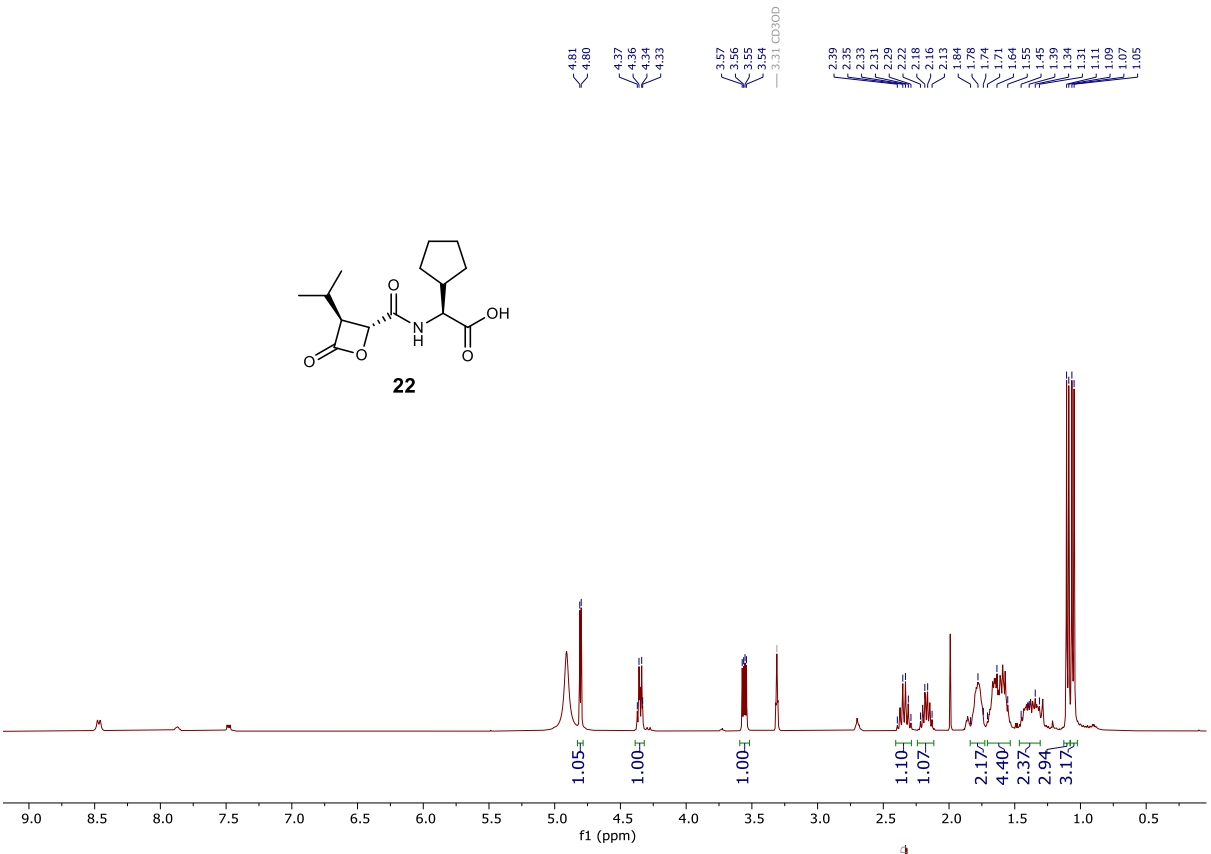
21

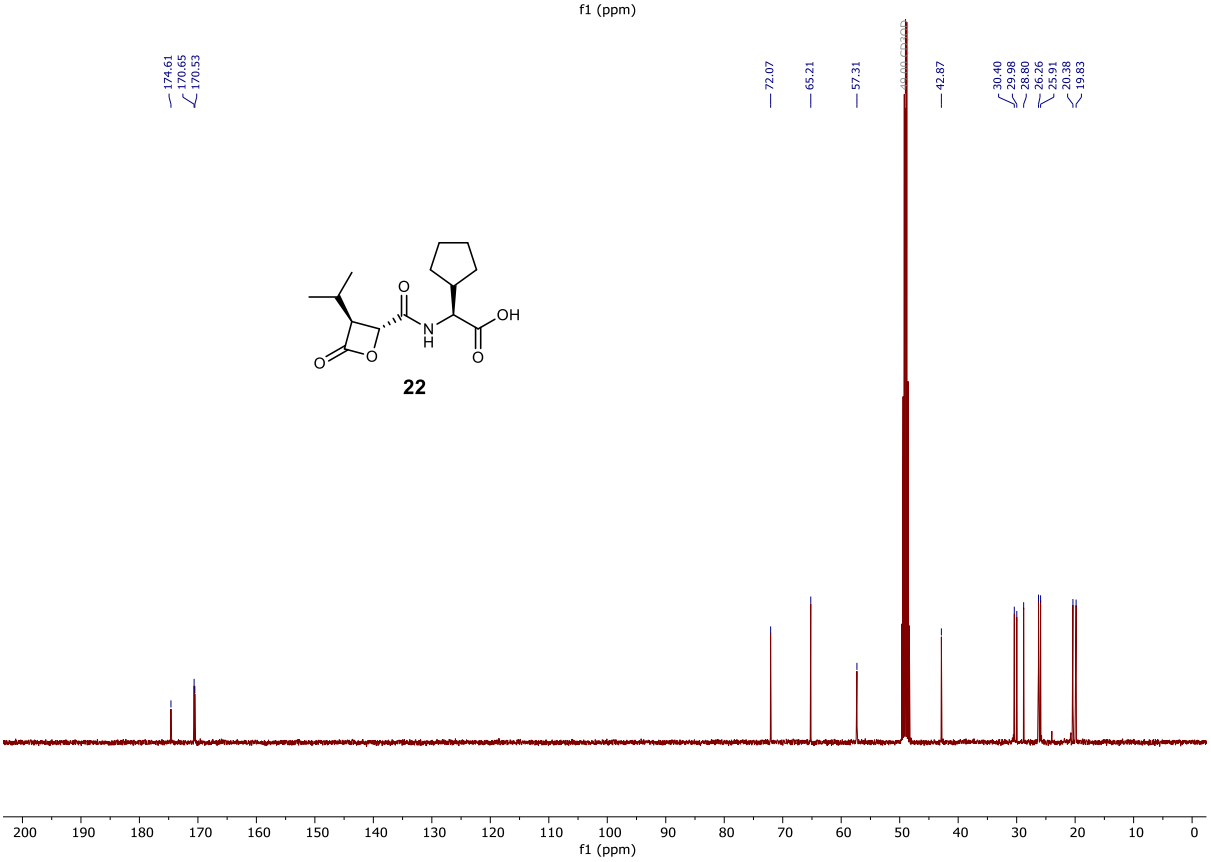

4.80
4.79
4.41
4.39
3.58
3.57
3.56
3.55
3.31 CD3OD
2.81
2.77
2.75
2.71
2.24
2.19
2.17
2.13
2.12
1.99
1.93
1.85
1.83
1.79
1.11
1.09
1.07
1.05
1.19
1.00
1.02
1.10
1.05
6.28
2.95
3.00
f1 (ppm)

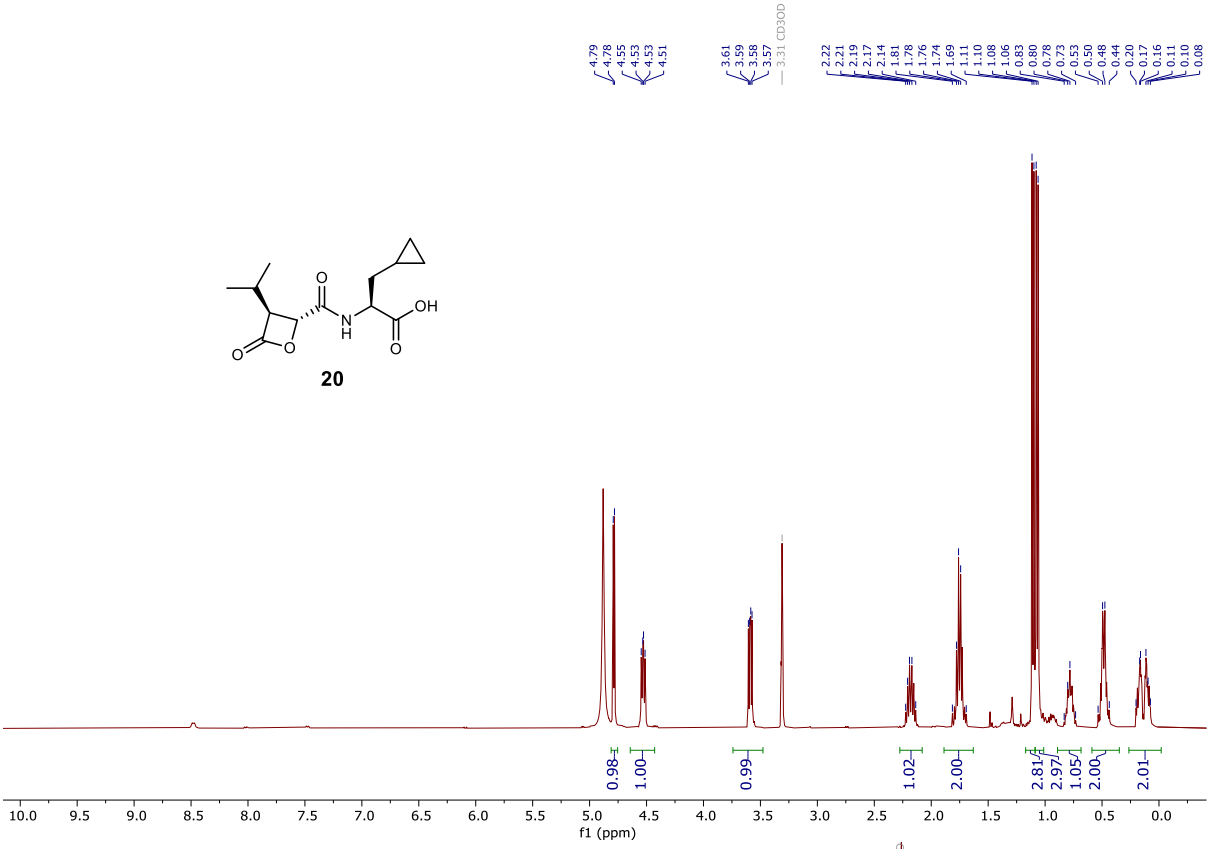
19

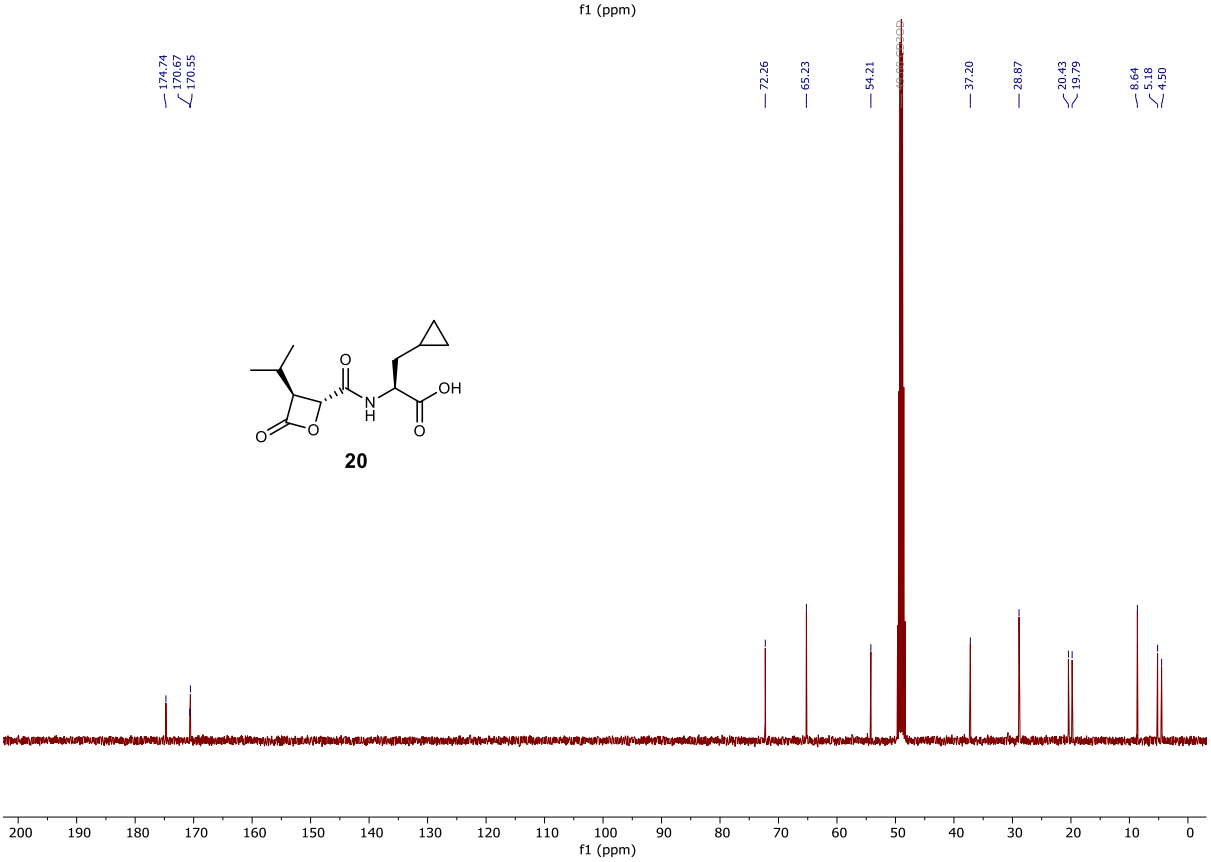

173.98
170.80
170.67
72.14
65.32
57.93
38.14
28.85
26.67
26.22
20.37
19.82
18.70
f1 (ppm)

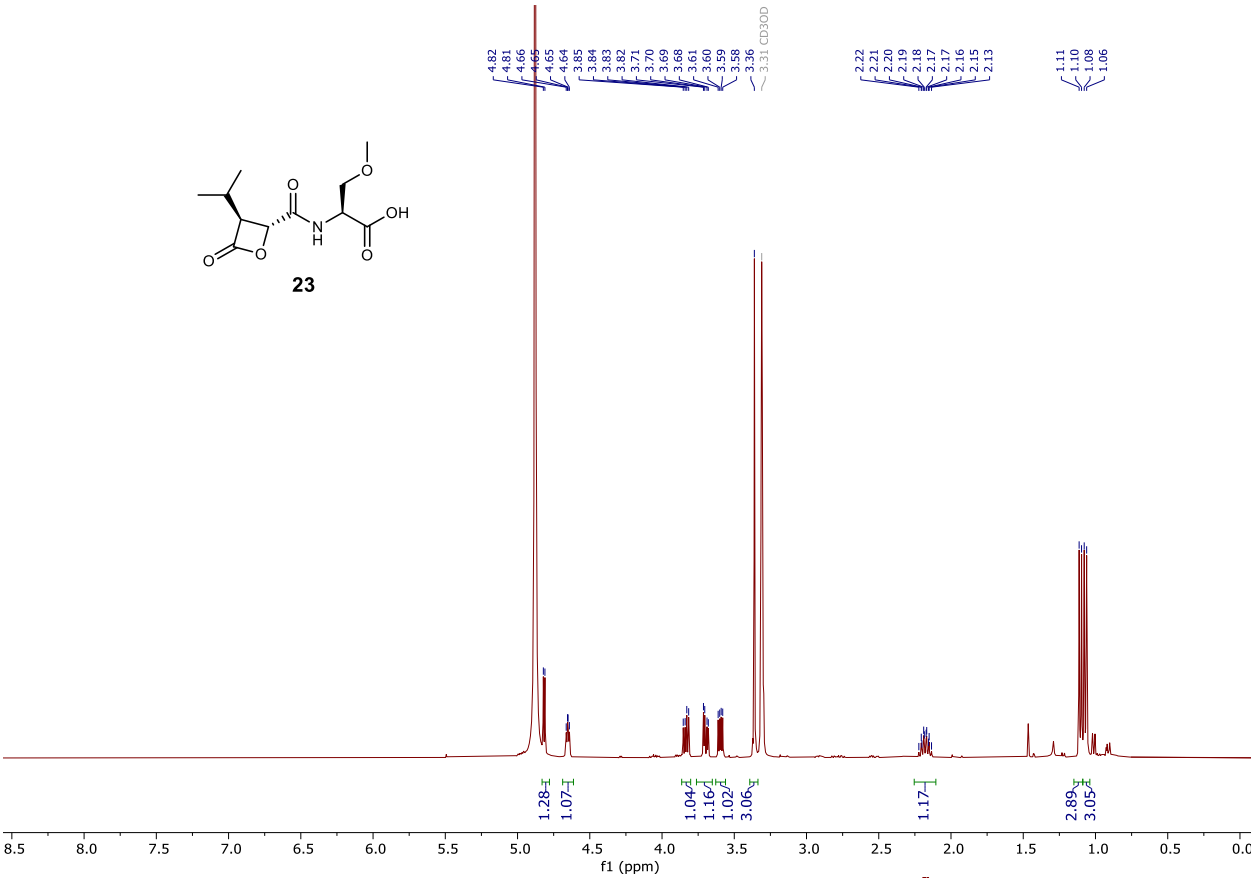
19

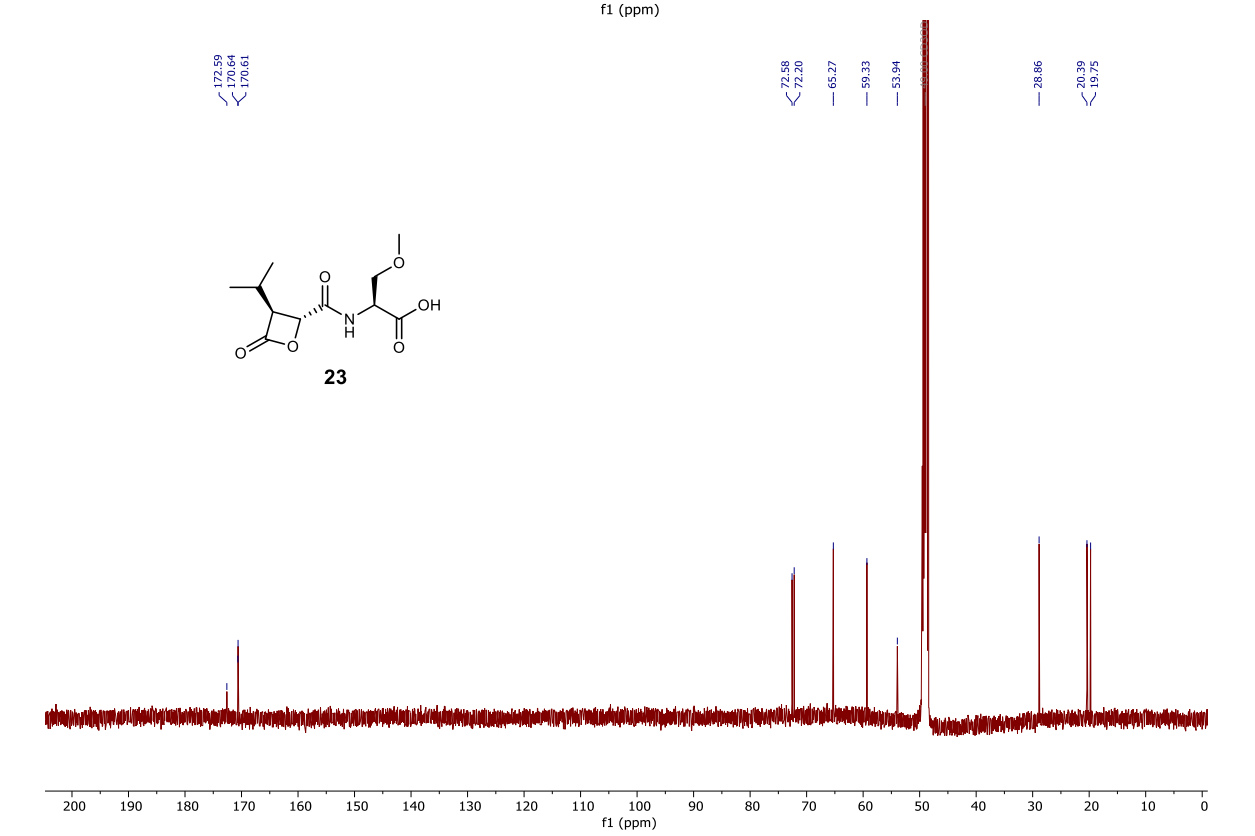

4.81
4.80
4.37
4.36
4.34
4.33
3.57
3.56
3.55
3.54
3.31 CD3OD
2.39
2.35
2.33
2.31
2.29
2.22
2.18
2.16
2.13
1.84
1.78
1.74
1.71
1.64
1.55
1.45
1.39
1.34
1.31
1.11
1.09
1.07
1.05
O
N
H
OH
O
O
O
22
1.05
1.00
1.00
1.10
1.07
2.17
4.40
2.37
2.94
3.17
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
f1 (ppm)

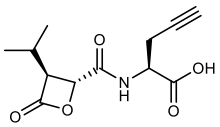

174.61
170.65
170.53
72.07
65.21
57.31
49.00 CD3OD
42.87
30.40
29.98
28.80
26.26
25.91
20.38
19.83
O
N
H
OH
O
O
O
22
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)


4.79
4.78
4.55
4.53
4.53
4.51
3.61
3.59
3.58
3.57
3.31 CD3OD
2.22
2.21
2.19
2.17
2.14
1.81
1.78
1.76
1.74
1.69
1.11
1.10
1.08
1.06
0.83
0.80
0.78
0.73
0.53
0.50
0.48
0.44
0.20
0.17
0.16
0.11
0.10
0.08
OH
O
N
H
O
O
O
20
0.98
1.00
0.99
1.02
2.00
2.81
2.97
1.05
2.00
2.01
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)


174.74
170.67
170.55
72.26
65.23
54.21
49.00 CD3OD
37.20
28.87
20.43
19.79
8.64
5.18
4.50
OH
O
N
H
O
O
O
20
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

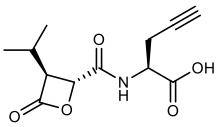

4.82
4.81
4.66
4.65
4.65
4.64
3.85
3.84
3.83
3.82
3.71
3.70
3.69
3.68
3.61
3.60
3.59
3.58
3.36
3.31 CD3OD
2.22
2.21
2.20
2.19
2.18
2.17
2.17
2.16
2.15
2.13
1.11
1.10
1.08
1.06
O
O
N
H
OH
O
O
O
23
1.28
1.07
1.04
1.16
1.02
3.06
1.17
2.89
3.05
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)


172.59
170.64
170.61
72.58
72.20
65.27
59.33
53.94
28.86
20.39
19.75
O
O
N
H
OH
O
O
O
23
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)



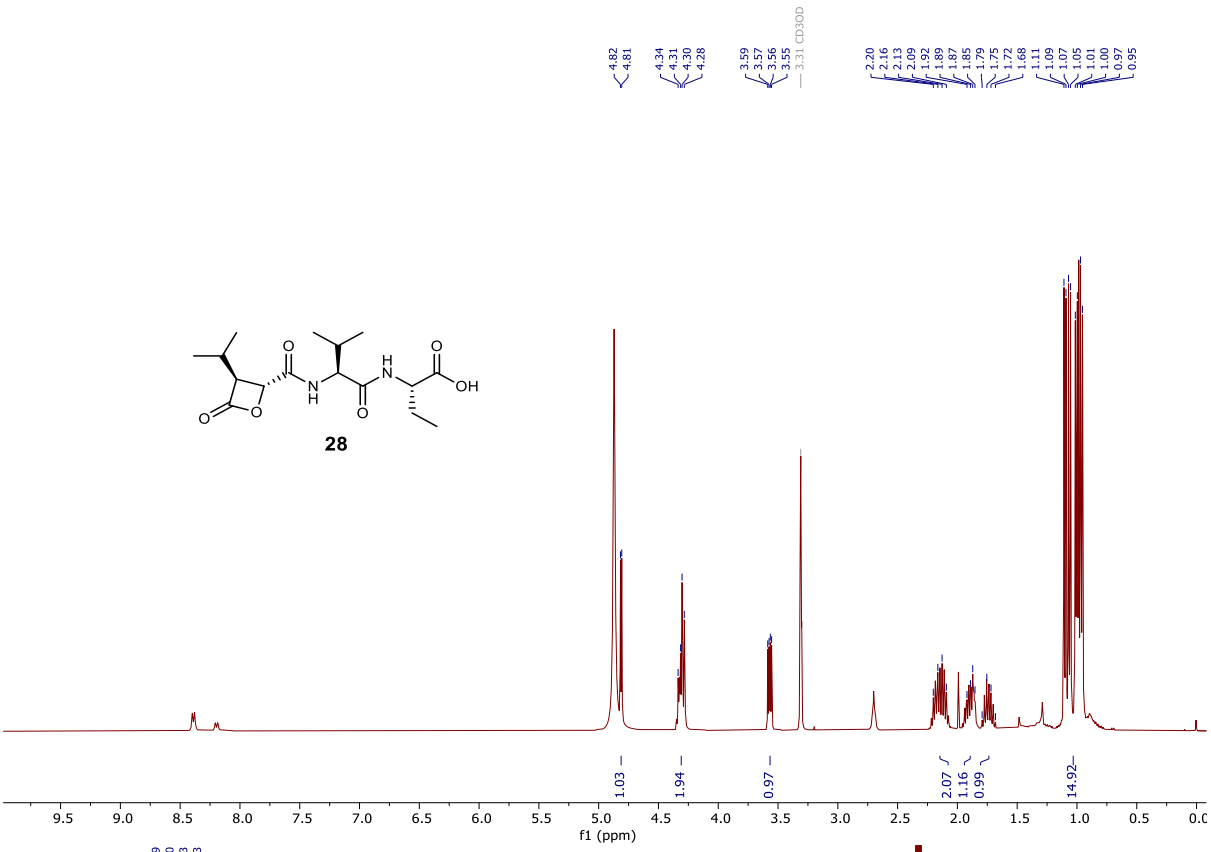
24

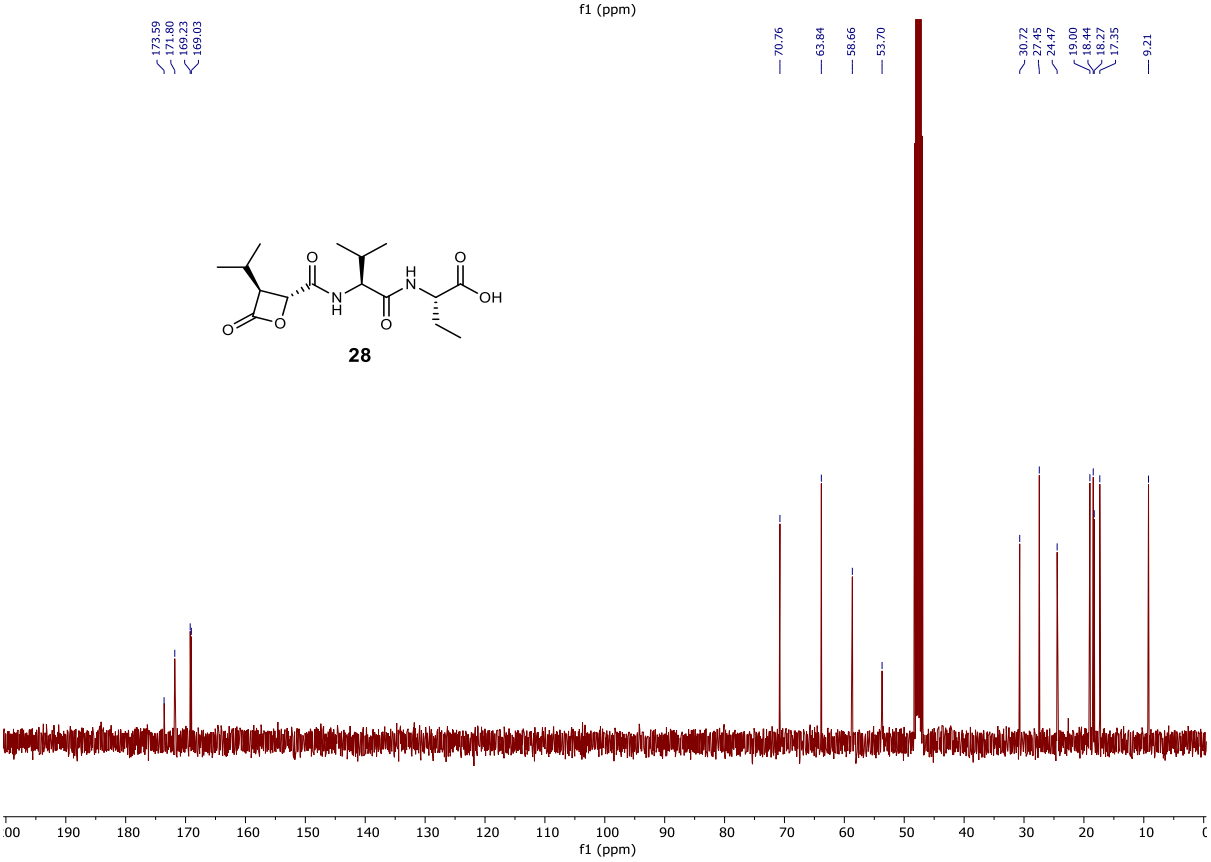
4.82
4.81
4.64
4.62
4.62
4.61
3.60
3.59
3.58
3.57
3.31 CD3OD
2.87
2.86
2.85
2.83
2.82
2.79
2.77
2.74
2.72
2.40
2.39
2.39
2.25
2.20
2.17
2.12
1.12
1.10
1.08
1.07
1.14
1.00
1.01
2.14
0.94
1.09
2.95
3.05
f1 (ppm)

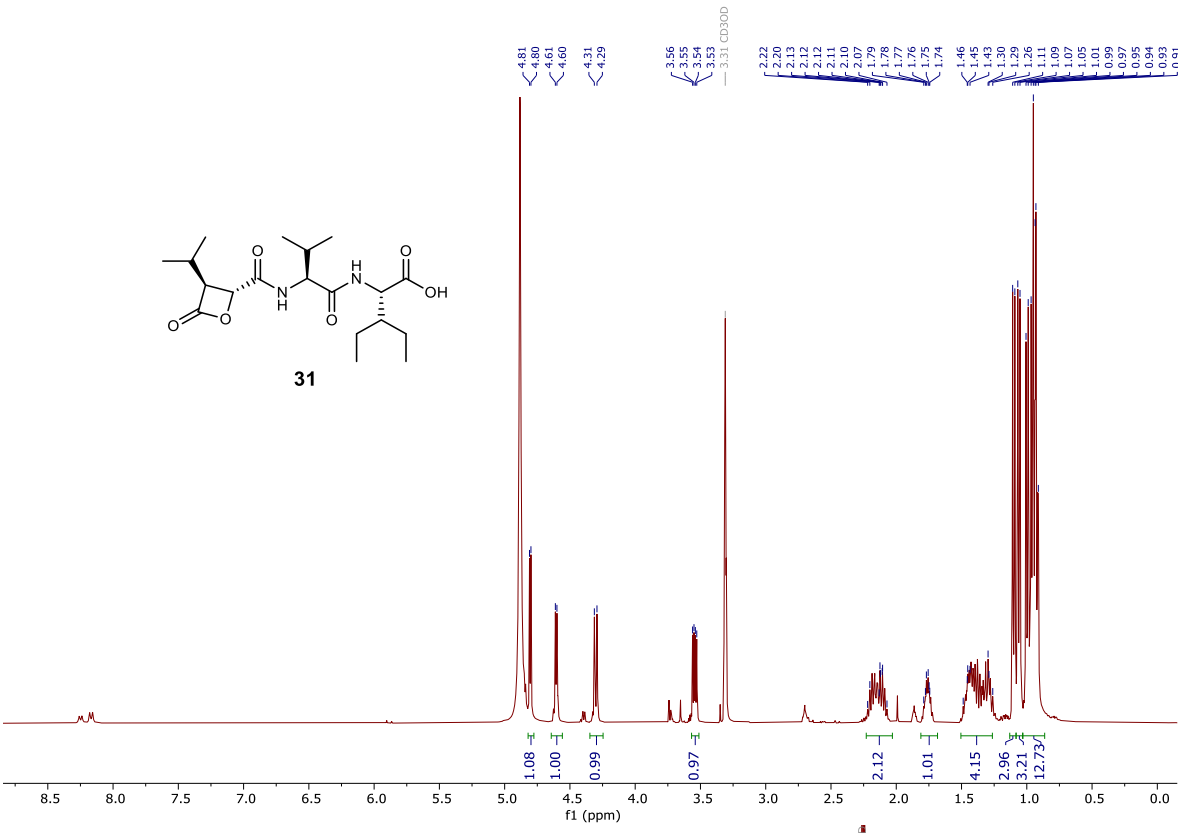


24

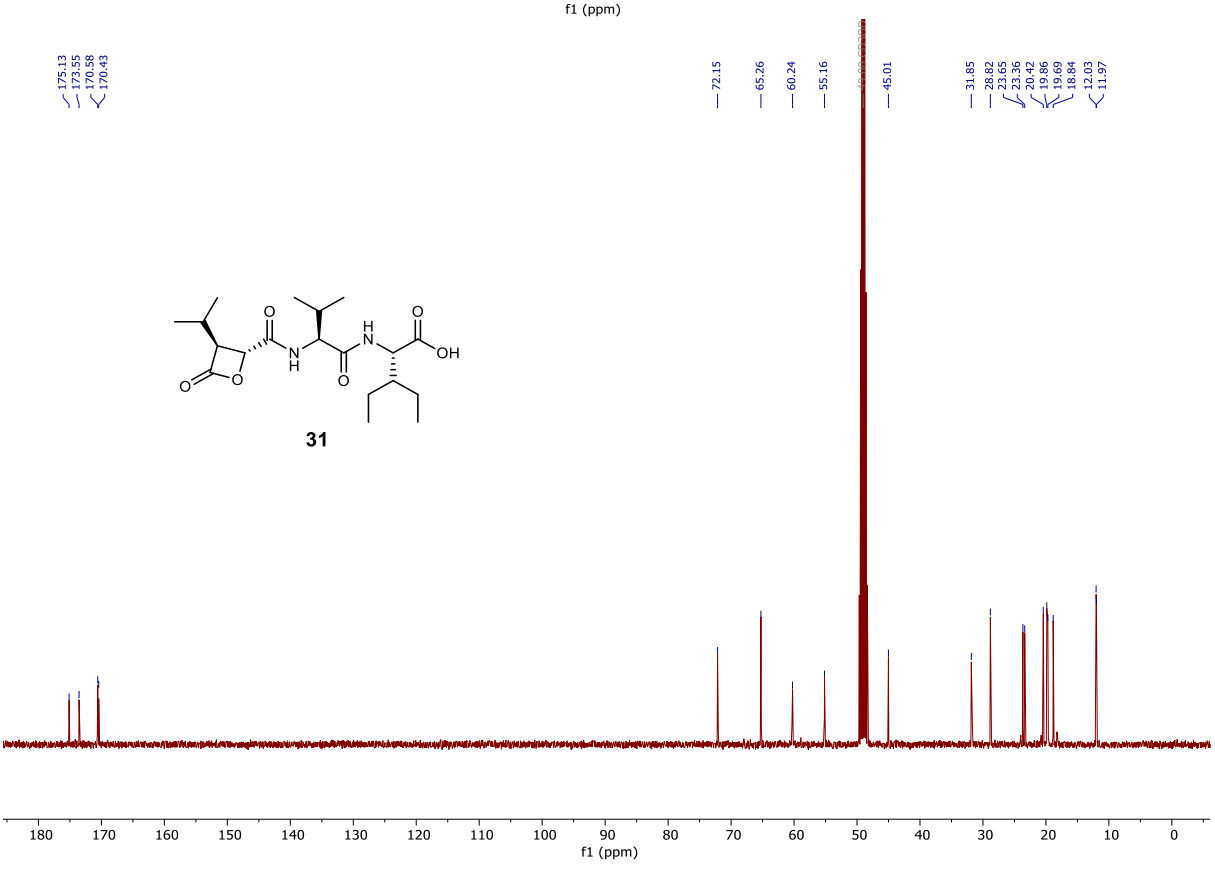
172.68
170.59
79.94
72.34
72.13
65.43
52.33
28.88
22.05
20.43
19.81
f1 (ppm)


4.82
4.81
4.34
4.31
4.30
4.28
3.59
3.57
3.56
3.55
3.31 CD3OD
2.20
2.16
2.13
2.09
1.92
1.89
1.87
1.85
1.79
1.75
1.72
1.68
1.11
1.09
1.07
1.05
1.01
1.00
0.97
0.95
28
1.03
1.94
0.97
2.07
1.16
0.99
14.92
f1 (ppm)

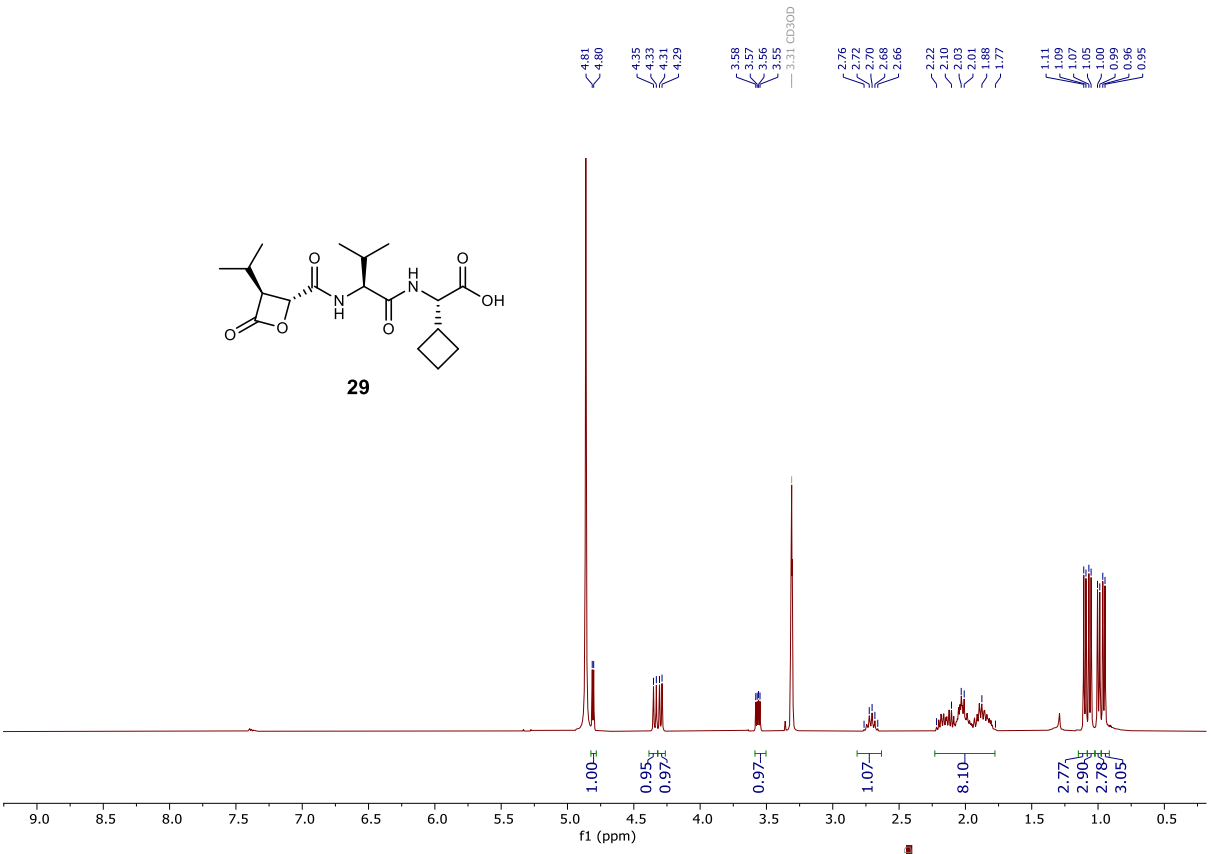

173.59
171.80
169.23
169.03
70.76
63.84
58.66
53.70
30.72
27.45
24.47
19.00
18.44
18.27
17.35
9.21
28
f1 (ppm)


4.81
4.80
4.61
4.60
4.31
4.29
3.56
3.55
3.54
3.53
3.31 CD3OD
2.22
2.20
2.13
2.12
2.11
2.10
2.07
1.79
1.78
1.77
1.76
1.75
1.74
1.46
1.45
1.43
1.30
1.29
1.26
1.11
1.09
1.07
1.05
1.01
0.99
0.97
0.95
0.94
0.93
1.08
1.00
0.99
0.97
2.12
1.01
4.15
2.96
3.21
12.73
f1 (ppm)

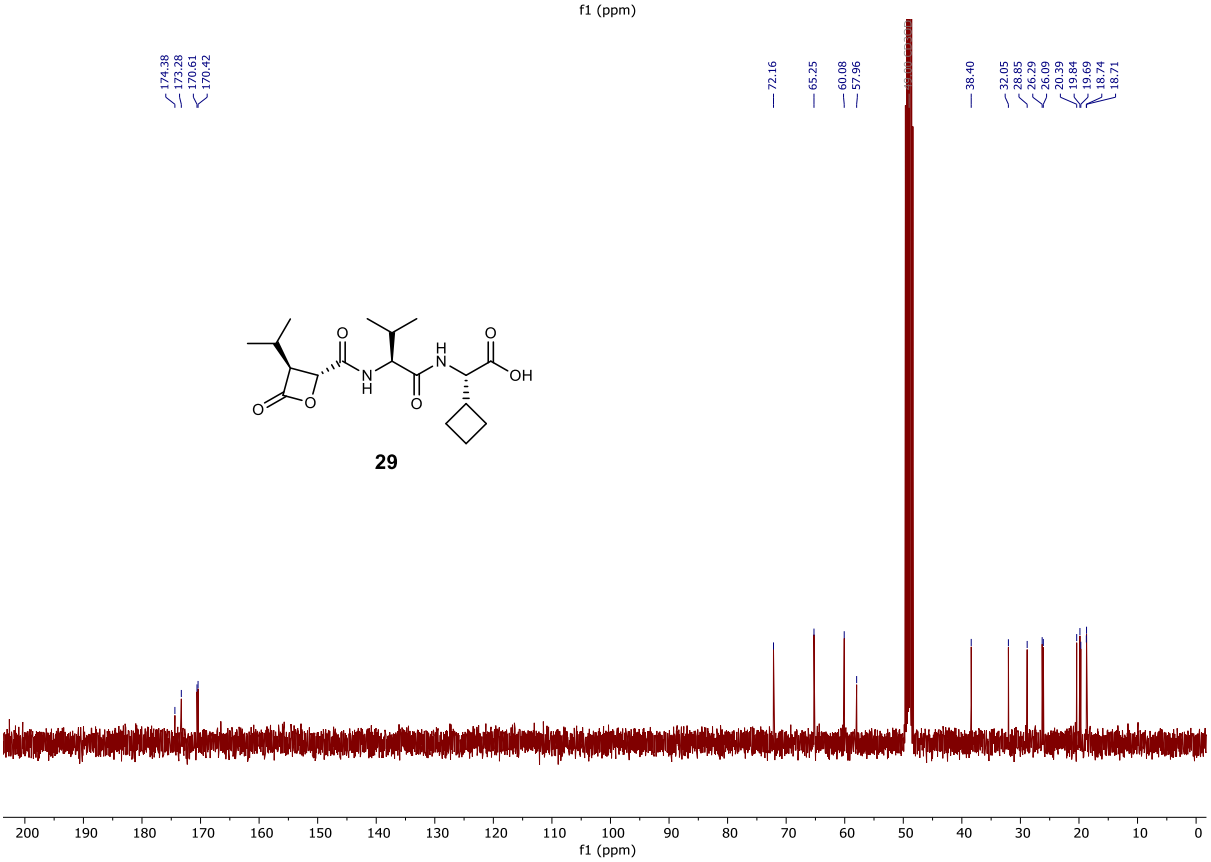
31


175.13
173.55
170.58
170.43
72.15
65.26
60.24
55.16
49.00 CD3OD
45.01
31.85
28.82
23.65
23.36
20.42
19.86
19.69
18.84
12.03
11.97
f1 (ppm)

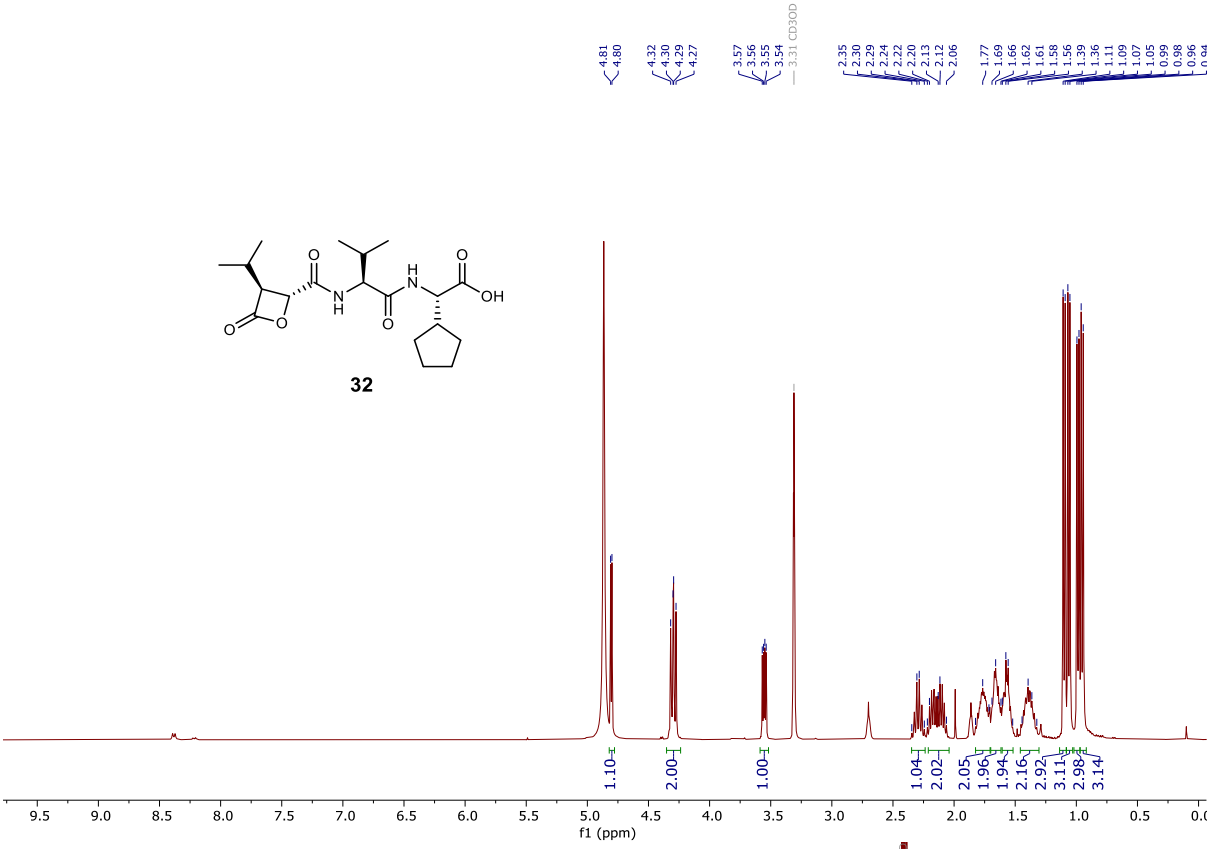
31


4.81
4.80
4.35
4.33
4.31
4.29
3.58
3.57
3.56
3.55
3.31 CD3OD
2.76
2.72
2.70
2.68
2.66
2.22
2.10
2.03
2.01
1.88
1.77
1.11
1.09
1.07
1.05
1.00
0.99
0.96
0.95
1.00
0.95
0.97
0.97
1.07
8.10
2.77
2.90
2.78
3.05
f1 (ppm)

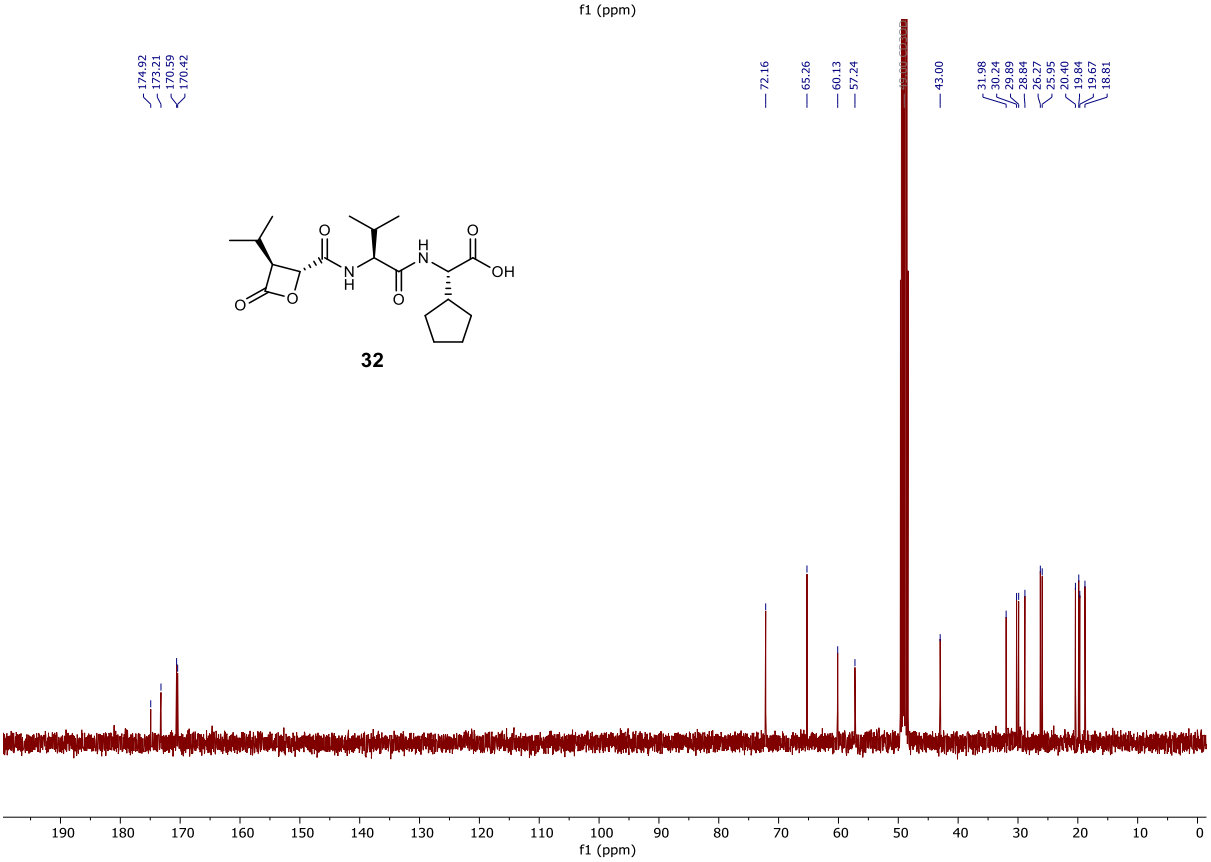
29


174.38
173.28
170.61
170.42
72.16
65.25
60.08
57.96
38.40
32.05
28.85
26.29
26.09
20.39
19.84
19.69
18.74
18.71
f1 (ppm)

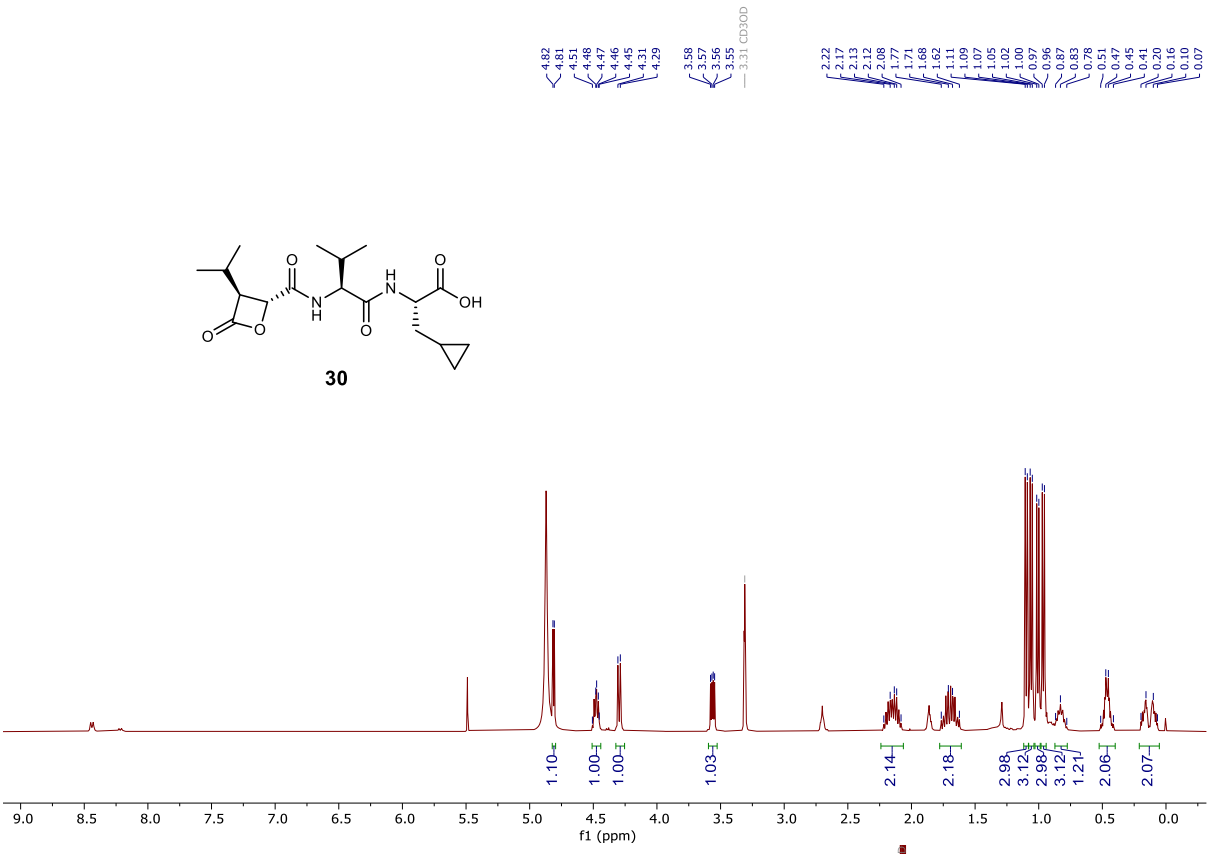
29

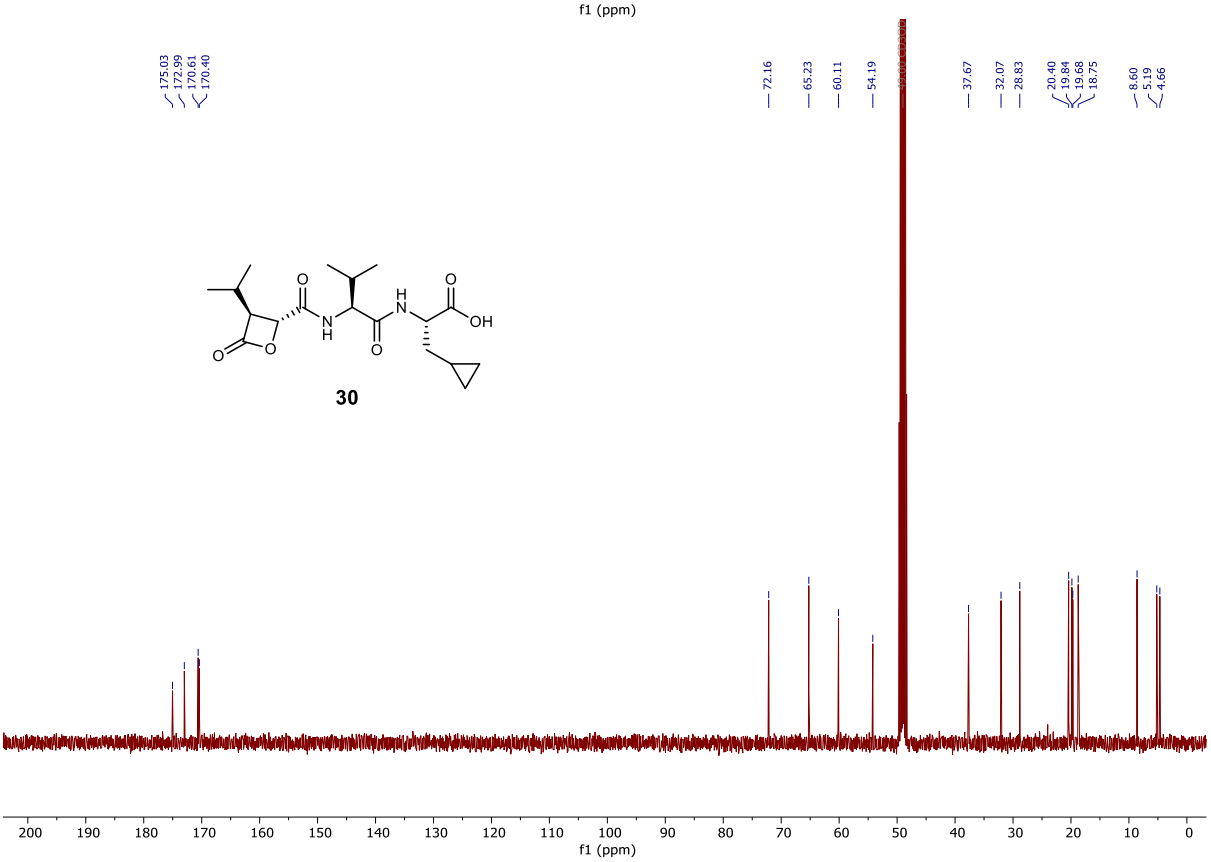

4.81
4.80
4.32
4.30
4.29
4.27
3.57
3.56
3.55
3.54
3.31 CD3OD
2.35
2.30
2.29
2.24
2.22
2.20
2.13
2.12
2.06
1.77
1.69
1.66
1.62
1.61
1.58
1.56
1.39
1.36
1.11
1.09
1.07
1.05
0.99
0.98
0.96
1.10
2.00
1.00
1.04
2.02
2.05
1.96
1.94
2.16
2.92
3.11
2.98
3.14
f1 (ppm)

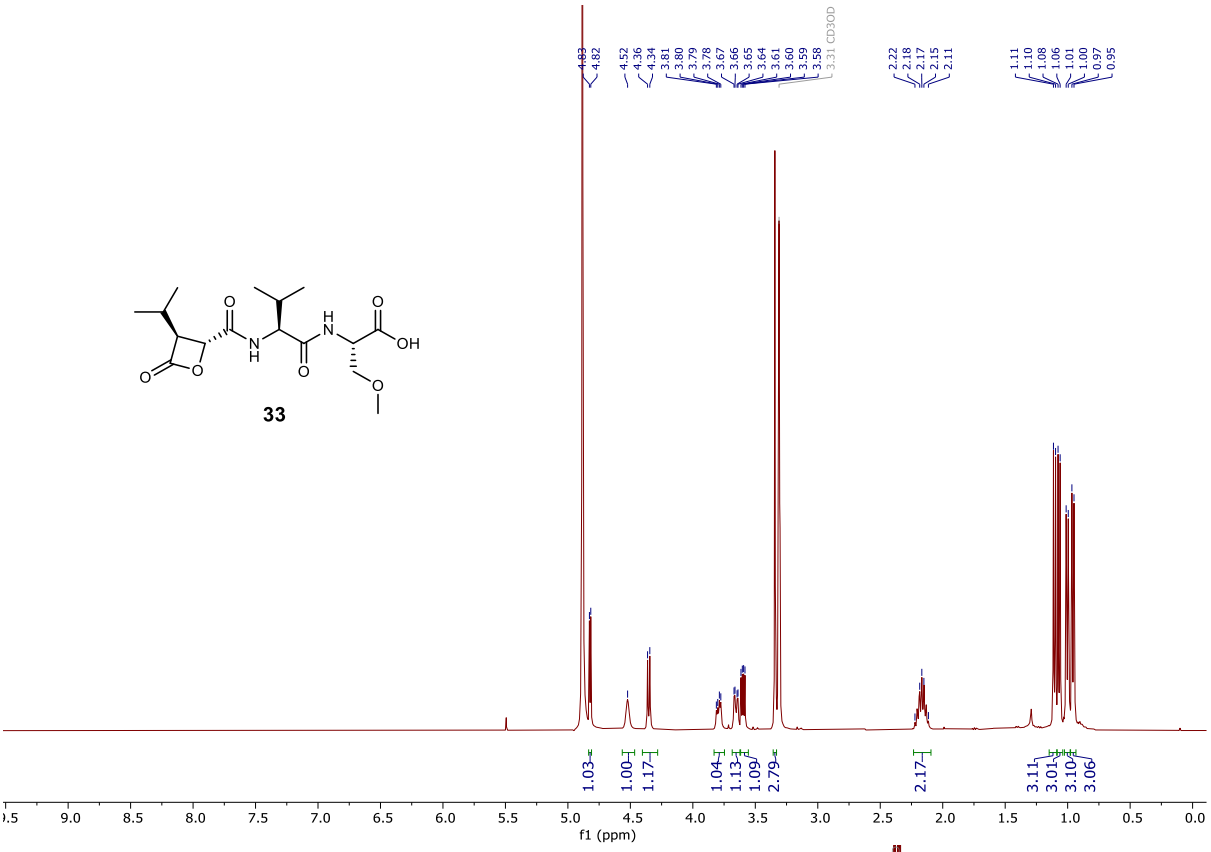
32

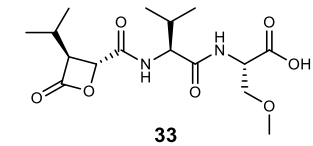

174.92
173.21
170.59
170.42
72.16
65.26
60.13
57.24
43.00
31.98
30.24
29.89
28.84
26.27
25.95
20.40
19.84
19.67
18.81
f1 (ppm)

32


4.82
4.81
4.51
4.48
4.47
4.46
4.45
4.31
4.29
3.58
3.57
3.56
3.55
3.31 CD3OD
2.22
2.17
2.13
2.12
2.08
1.77
1.71
1.68
1.62
1.11
1.09
1.07
1.05
1.02
1.00
0.97
0.96
0.87
0.83
0.78
0.51
0.47
0.45
0.41
0.20
0.16
0.10
0.07
30
1.10
1.00
1.00
1.03
2.14
2.18
2.98
3.12
2.98
3.12
1.21
2.06
2.07
f1 (ppm)


175.03
172.99
170.61
170.40
72.16
65.23
60.11
54.19
37.67
32.07
28.83
20.40
19.84
19.68
18.75
8.60
5.19
4.66
30
f1 (ppm)


4.83
4.82
4.52
4.36
4.34
3.81
3.80
3.79
3.78
3.67
3.66
3.65
3.64
3.61
3.60
3.59
3.58
3.31 CD3OD
2.22
2.18
2.17
2.15
2.11
1.11
1.10
1.08
1.06
1.01
1.00
0.97
0.95
1.03
1.00
1.17
1.04
1.13
1.09
2.79
2.17
3.11
3.01
3.10
3.06
f1 (ppm)

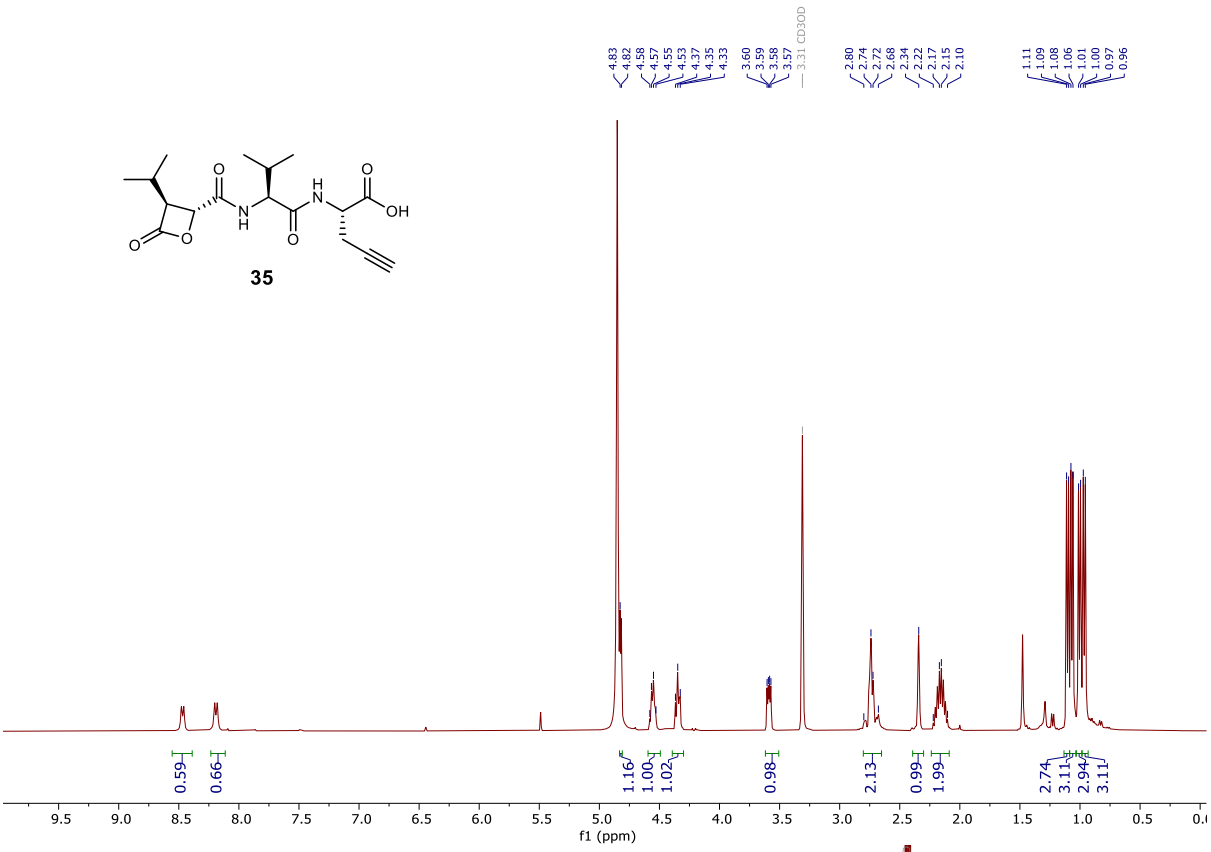

O
N
H
O
N
H
O
OH
O
O
O

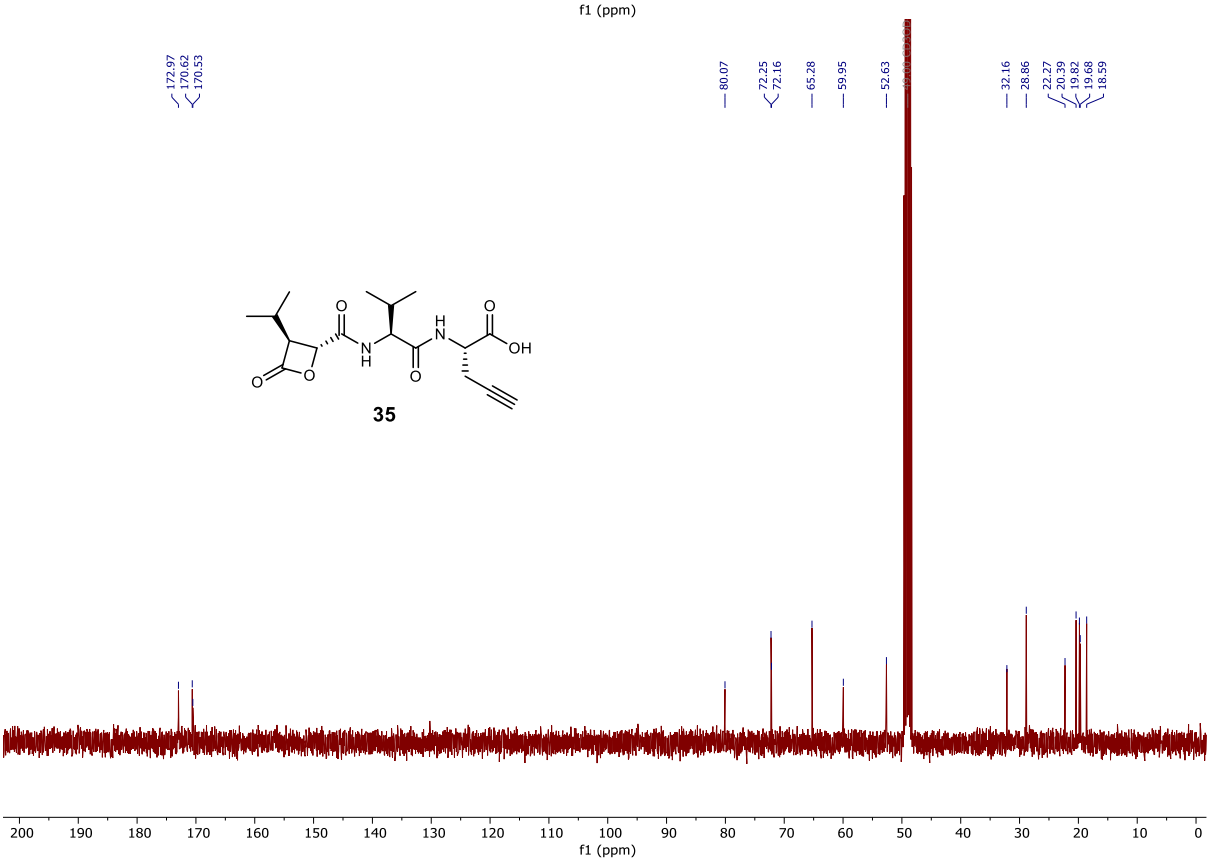
33

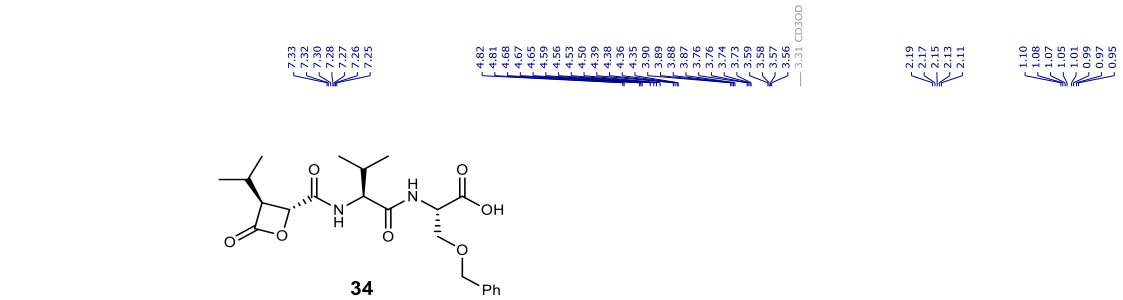
172.84
170.64
170.49
73.32
72.17
65.25
59.95
59.30
32.10
28.87
20.40
19.84
19.72
18.55
f1 (ppm)

O
N
H
O
N
H
O
OH
O
O
O

33

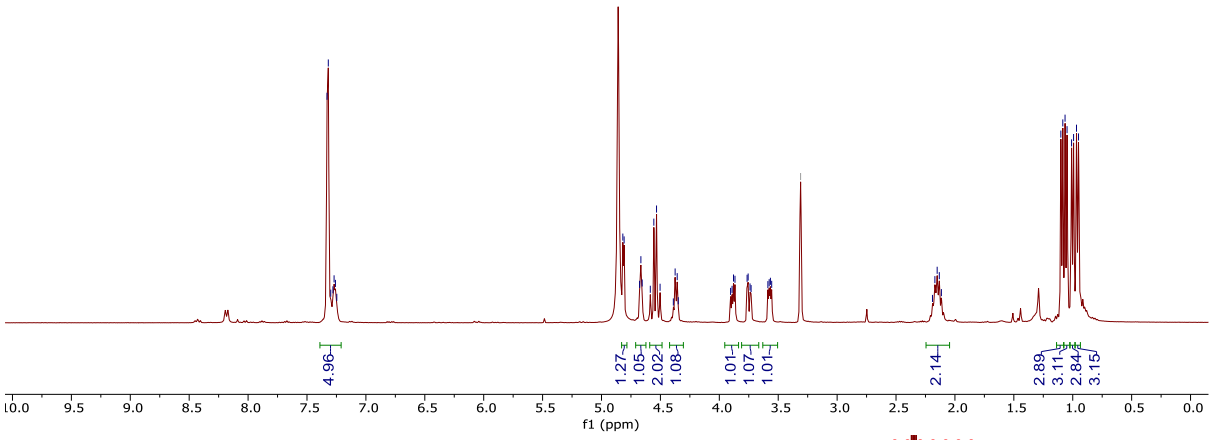

4.83
4.82
4.58
4.57
4.55
4.53
4.37
4.35
4.33
3.60
3.59
3.58
3.57
3.31 CD3OD
2.80
2.74
2.72
2.68
2.34
2.22
2.17
2.15
2.10
1.11
1.09
1.08
1.06
1.01
1.00
0.97
0.96
O
N
H
O
N
H
O
OH
O
O
35
0.59
0.66
1.16
1.00
1.02
0.98
2.13
0.99
1.99
2.74
3.11
2.94
3.11
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
f1 (ppm)

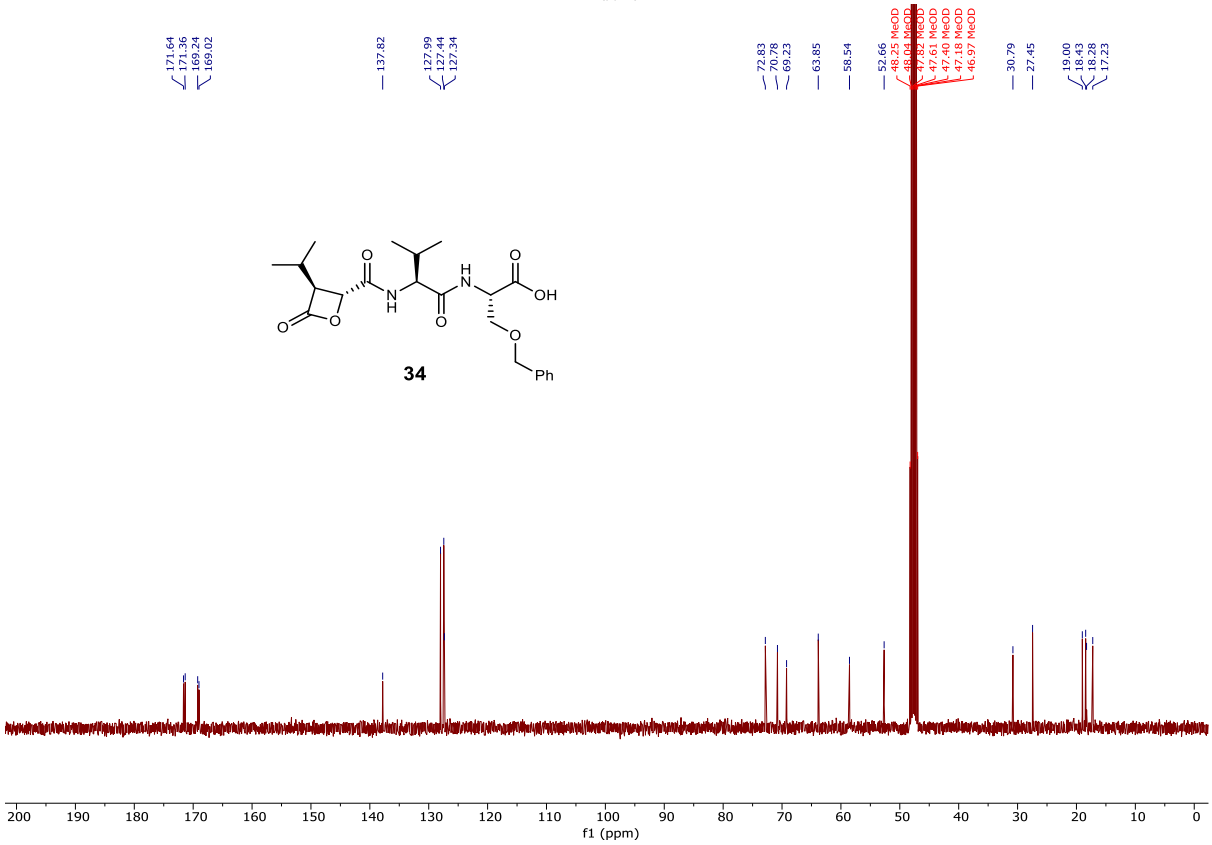

172.97
170.62
170.53
80.07
72.25
72.16
65.28
59.95
52.63
32.16
28.86
22.27
20.39
19.82
19.68
18.59
O
N
H
O
N
H
O
OH
O
O
35
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)


7.33
7.32
7.30
7.28
7.27
7.26
7.25
4.82
4.81
4.68
4.67
4.65
4.59
4.56
4.53
4.50
4.39
4.38
4.36
4.35
3.90
3.89
3.88
3.87
3.76
3.74
3.73
3.59
3.58
3.57
3.56
3.31 CD3OD
2.19
2.17
2.15
2.13
2.11
1.10
1.08
1.07
1.05
1.01
0.99
0.97
0.95

O
N
H
O
H
N
O
OH
O
O
O
Ph

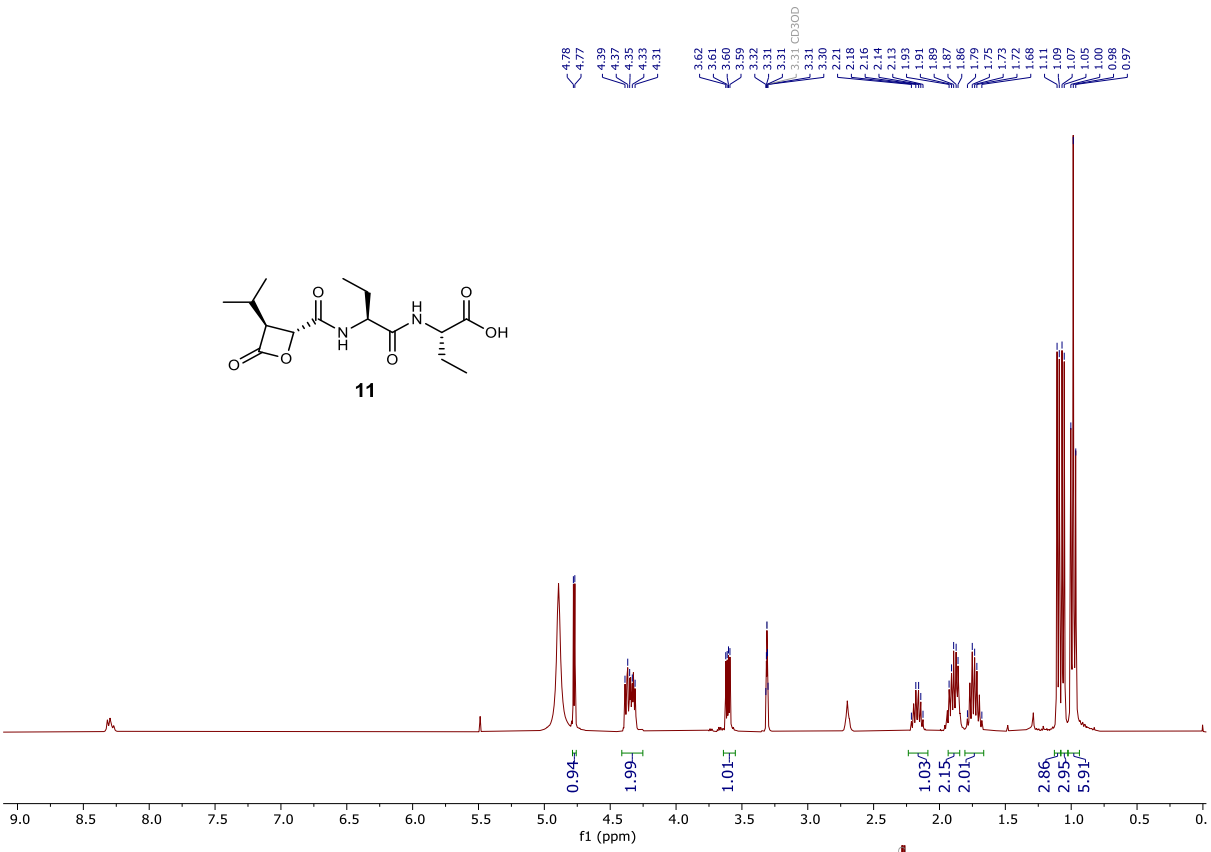
34

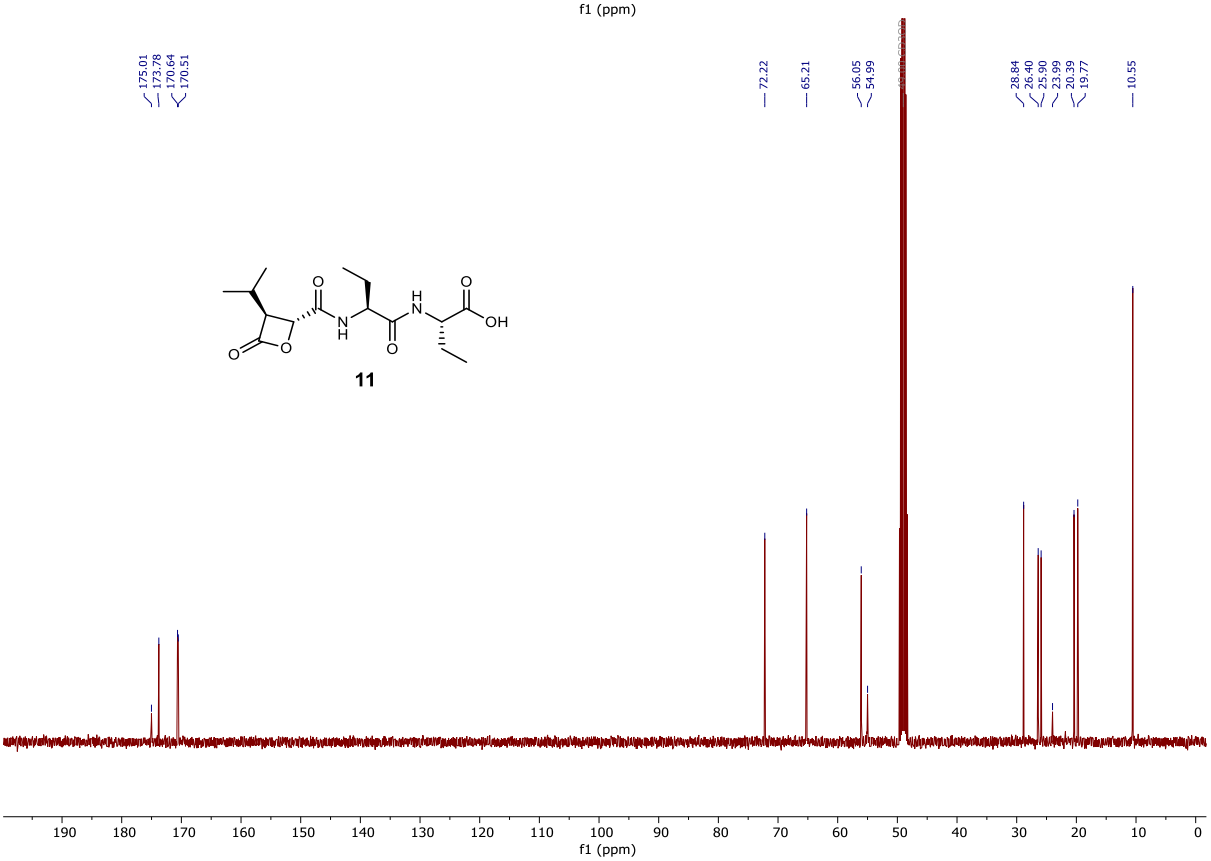

4.96
1.27
1.05
2.02
1.08
1.01
1.07
1.01
2.14
2.89
3.11
2.84
3.15
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

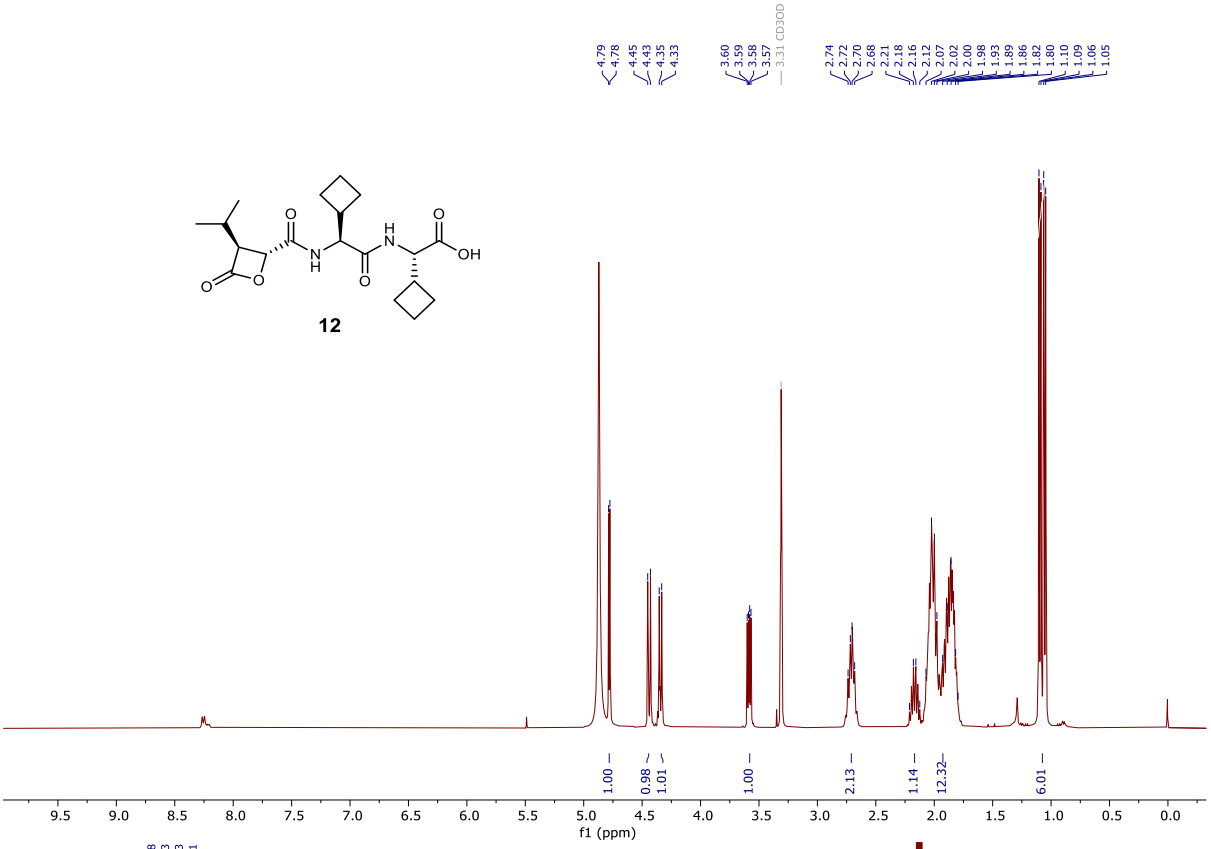

171.64
171.36
169.24
169.02
137.82
127.99
127.44
127.34
72.83
70.78
69.23
63.85
58.54
52.66
48.25 MeOD
48.04 MeOD
47.82 MeOD
47.61 MeOD
47.40 MeOD
47.18 MeOD
46.97 MeOD
30.79
27.45
19.00
18.43
18.28
17.23
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

O
N
H
O
H
N
O
OH
O
O
O
Ph

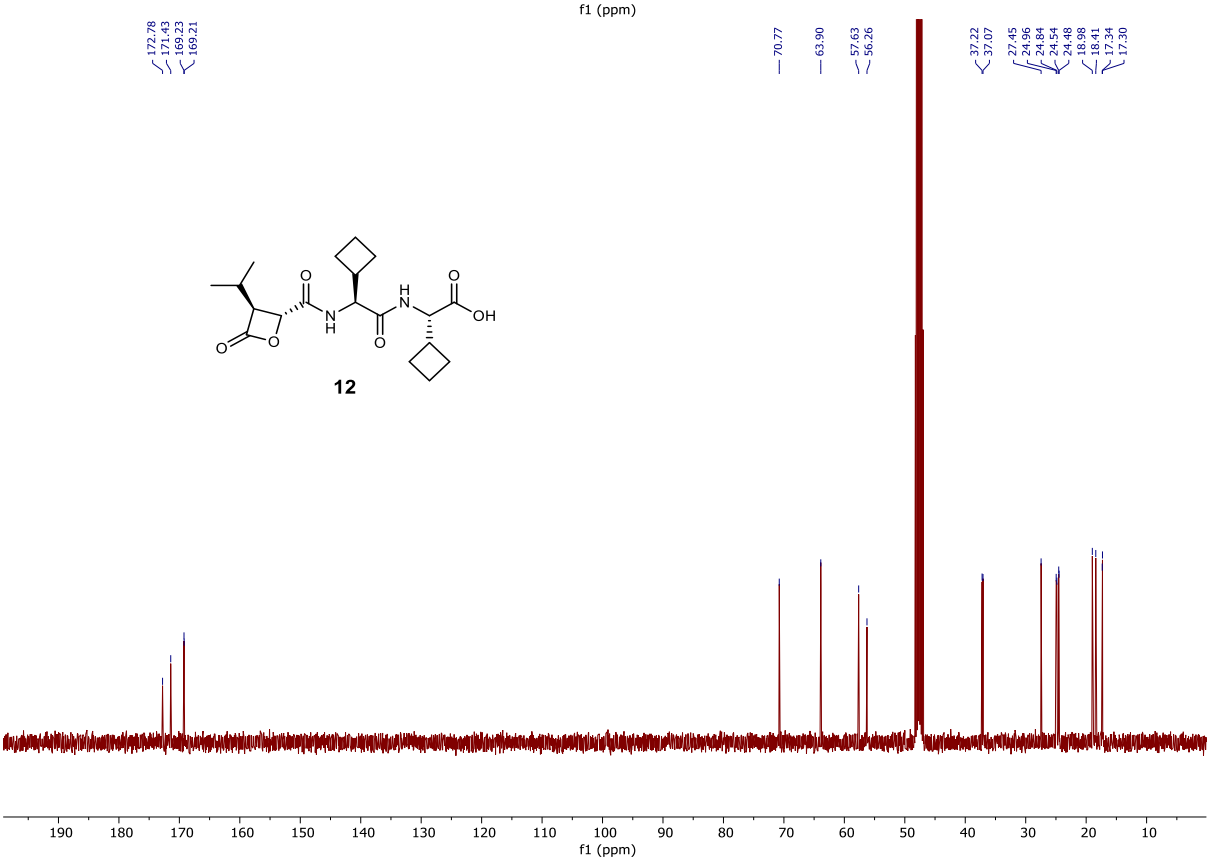
34


4.78
4.77
4.39
4.37
4.35
4.33
4.31
3.62
3.61
3.60
3.59
3.32
3.31
3.31
CD3OD
3.31
3.30
2.21
2.18
2.16
2.14
2.13
1.93
1.91
1.89
1.87
1.86
1.79
1.75
1.73
1.72
1.68
1.11
1.09
1.07
1.05
1.00
0.98
0.97
OH
11
0.94
1.99
1.01
1.03
2.15
2.01
2.86
2.95
5.91
f1 (ppm)

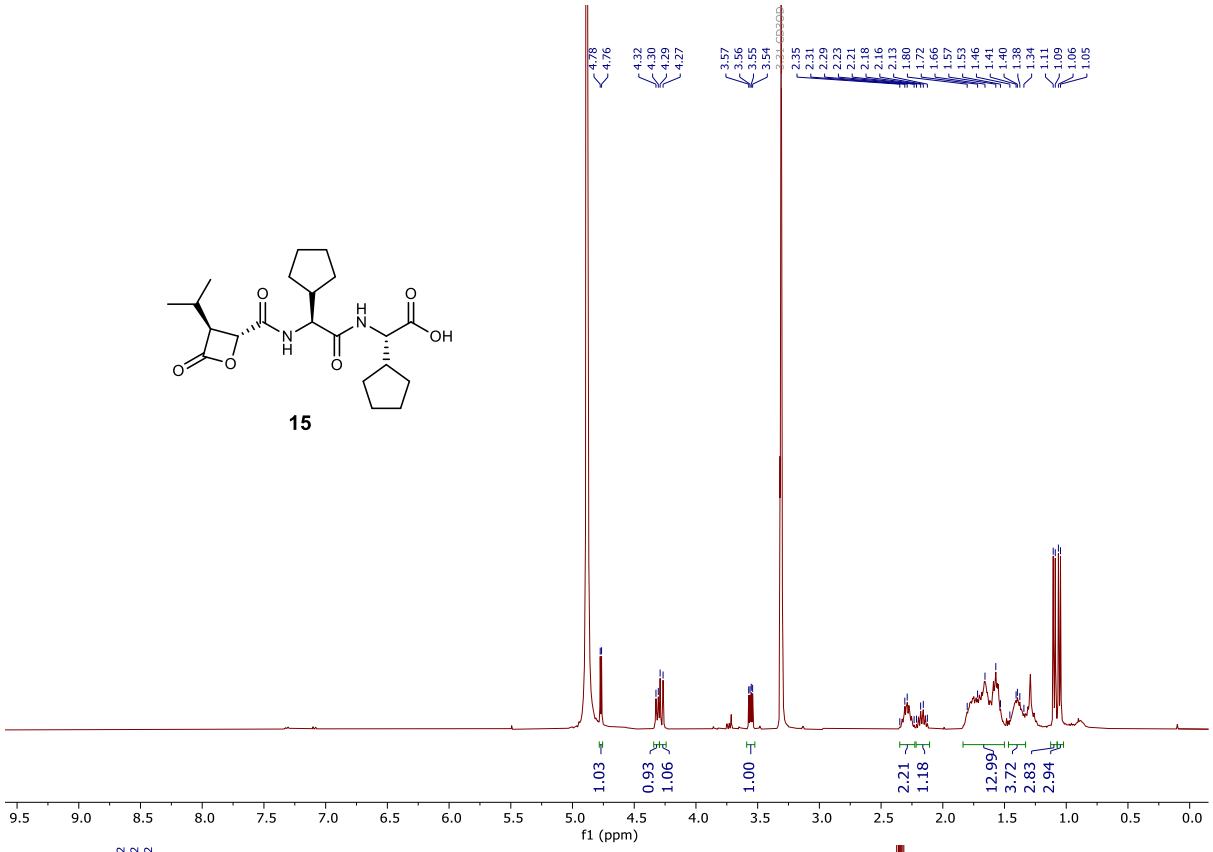

175.01
173.78
170.64
170.51
72.22
65.21
56.05
54.99
28.84
26.40
25.90
23.99
20.39
19.77
10.55
OH
11
f1 (ppm)

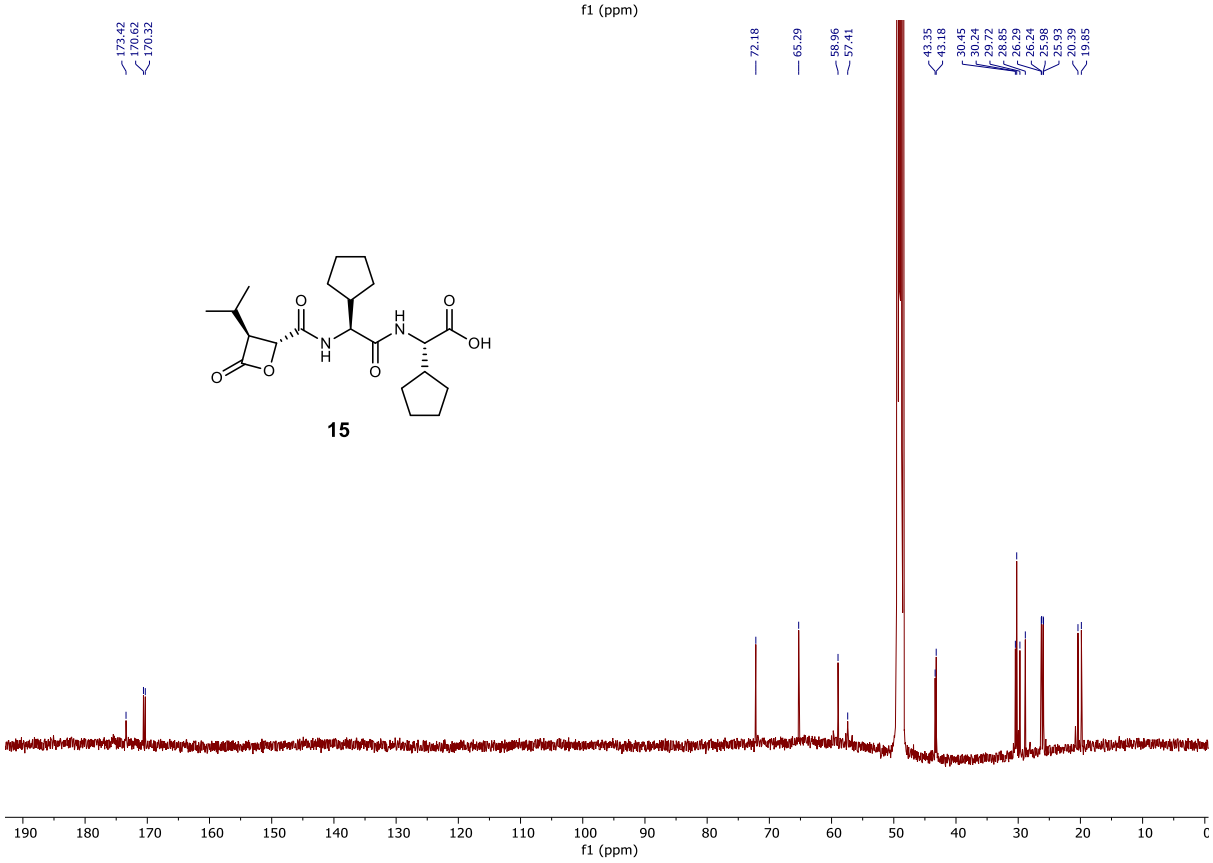

4.79
4.78
4.45
4.43
4.35
4.33
3.60
3.59
3.58
3.57
3.31 CD3OD
2.74
2.72
2.70
2.68
2.21
2.18
2.16
2.12
2.07
2.02
2.00
1.98
1.93
1.89
1.86
1.82
1.80
1.10
1.09
1.06
1.05
1.00
0.98
1.01
1.00
2.13
1.14
12.32
6.01
f1 (ppm)

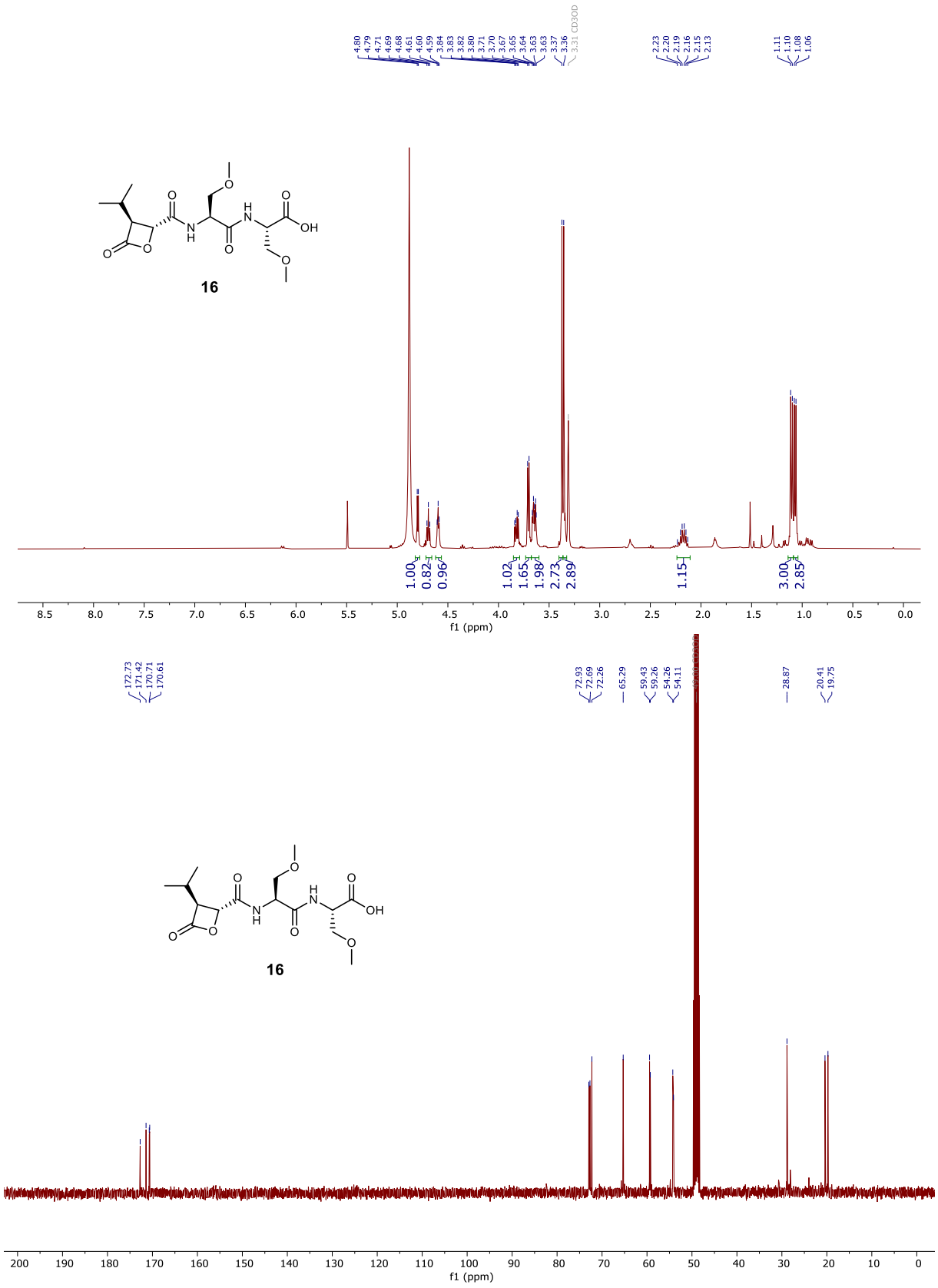
12

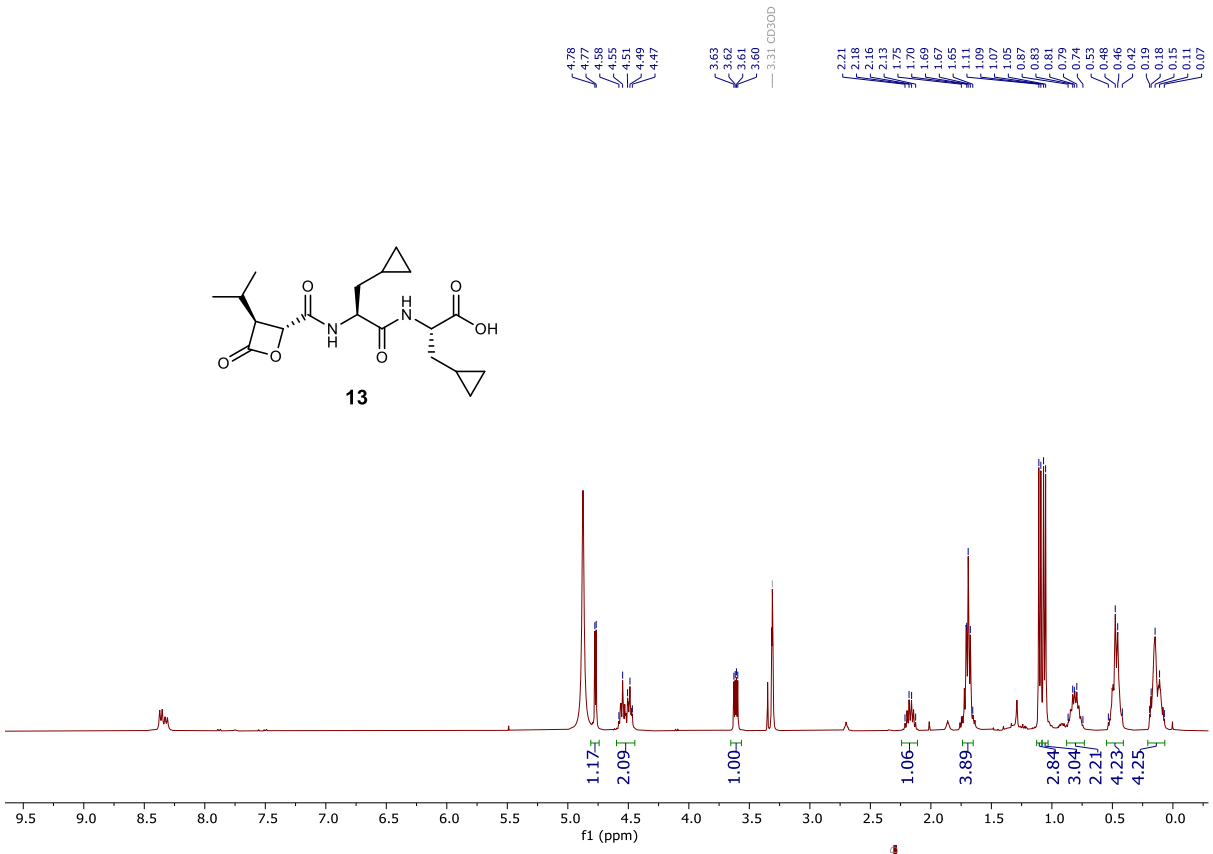

172.78
171.43
169.23
169.21
70.77
63.90
57.63
56.26
37.22
37.07
27.45
24.96
24.84
24.54
24.48
18.98
18.41
17.34
17.30
f1 (ppm)

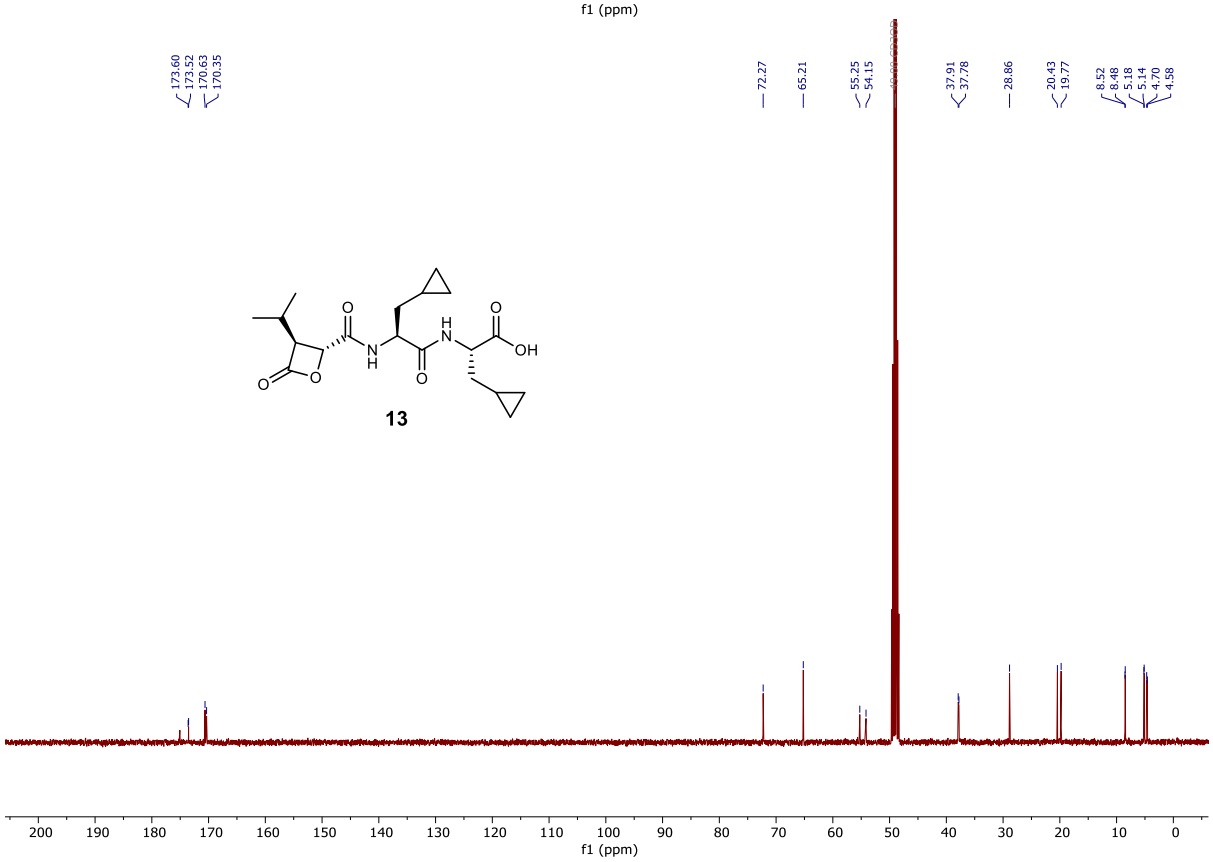
12

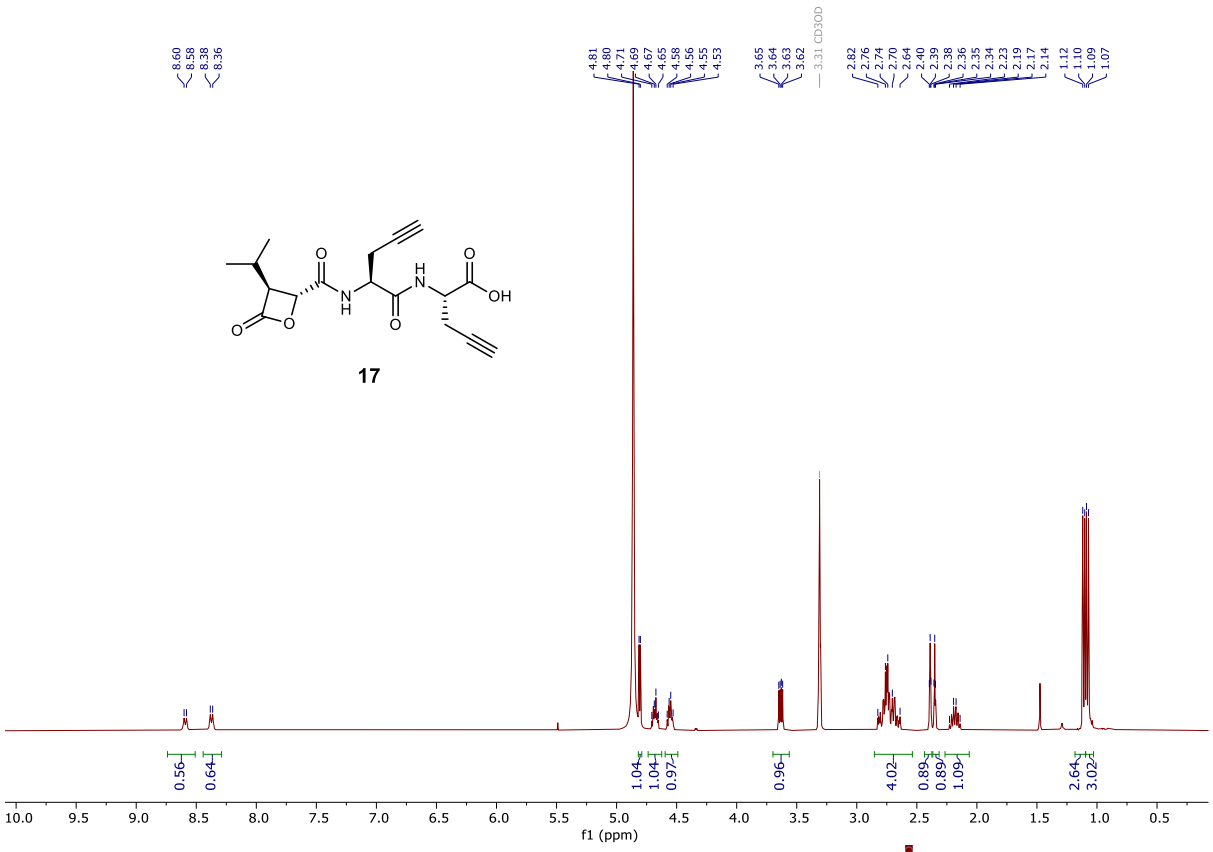

15
4.78
4.76
4.32
4.30
4.29
4.27
3.57
3.56
3.55
3.54
2.35
2.31
2.29
2.23
2.21
2.18
2.16
2.13
1.80
1.72
1.66
1.57
1.53
1.46
1.41
1.40
1.38
1.34
1.11
1.09
1.06
1.05
1.03
0.93
1.06
1.00
2.21
1.18
12.99
3.72
2.83
2.94
f1 (ppm)


15
173.42
170.62
170.32
72.18
65.29
58.96
57.41
43.35
43.18
30.45
30.24
29.72
28.85
26.29
26.24
25.98
25.93
20.39
19.85
f1 (ppm)

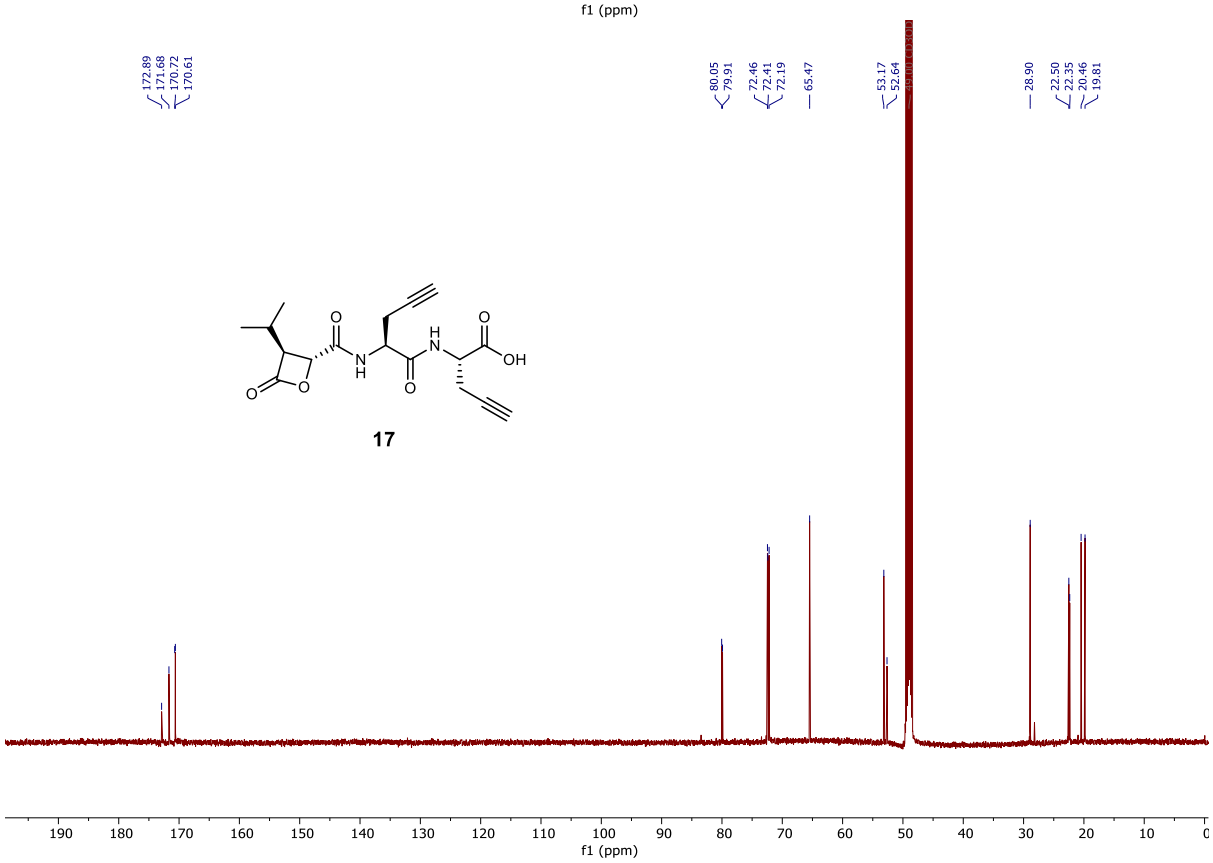

4.80
4.79
4.71
4.69
4.68
4.61
4.60
4.59
3.84
3.83
3.82
3.80
3.71
3.70
3.67
3.65
3.64
3.63
3.63
3.37
3.36
3.31 CD3OD
2.23
2.20
2.19
2.16
2.15
2.13
1.11
1.10
1.08
1.06
16
1.00
0.82
0.96
1.02
1.65
1.98
2.73
2.89
1.15
3.00
2.85
f1 (ppm)
172.73
171.42
170.71
170.61
72.93
72.69
72.26
65.29
59.43
59.26
54.26
54.11
28.87
20.41
19.75
16
f1 (ppm)


4.78
4.77
4.58
4.55
4.51
4.49
4.47
3.63
3.62
3.61
3.60
3.31 CD3OD
2.21
2.18
2.16
2.13
1.75
1.70
1.69
1.67
1.65
1.11
1.09
1.07
1.05
0.87
0.83
0.81
0.79
0.74
0.53
0.48
0.46
0.42
0.19
0.18
0.15
0.11
0.07
13
1.17
2.09
1.00
1.06
3.89
2.84
3.04
2.21
4.23
4.25
f1 (ppm)


173.60
173.52
170.63
170.35
72.27
65.21
55.25
54.15
37.91
37.78
28.86
20.43
19.77
8.52
8.48
5.18
5.14
4.70
4.58
13
f1 (ppm)


8.60
8.58
8.38
8.36
4.81
4.80
4.71
4.69
4.67
4.65
4.58
4.56
4.55
4.53
3.65
3.64
3.63
3.62
3.31 CD3OD
2.82
2.76
2.74
2.70
2.64
2.40
2.39
2.38
2.36
2.35
2.34
2.23
2.19
2.17
2.14
1.12
1.10
1.09
1.07
0.56
0.64
1.04
1.04
0.97
0.96
4.02
0.89
0.89
1.09
2.64
3.02
f1 (ppm)

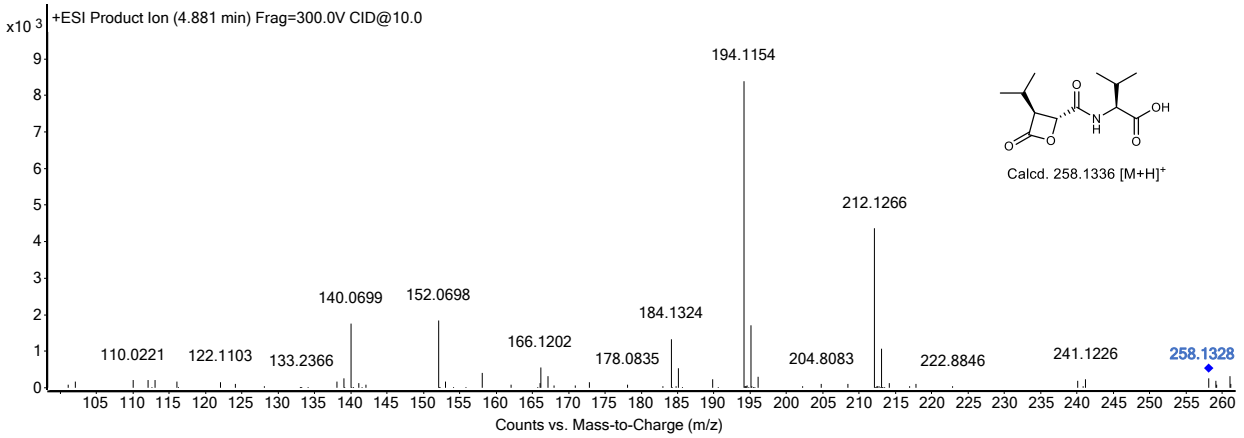
17

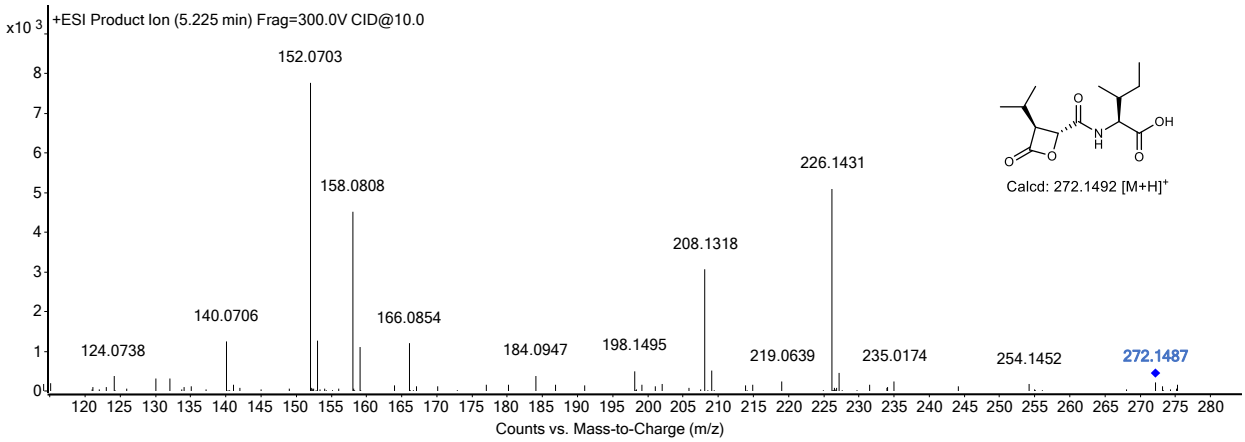

172.89
171.68
170.72
170.61
80.05
79.91
72.46
72.41
72.19
65.47
53.17
52.64
28.90
22.50
22.35
20.46
19.81
f1 (ppm)

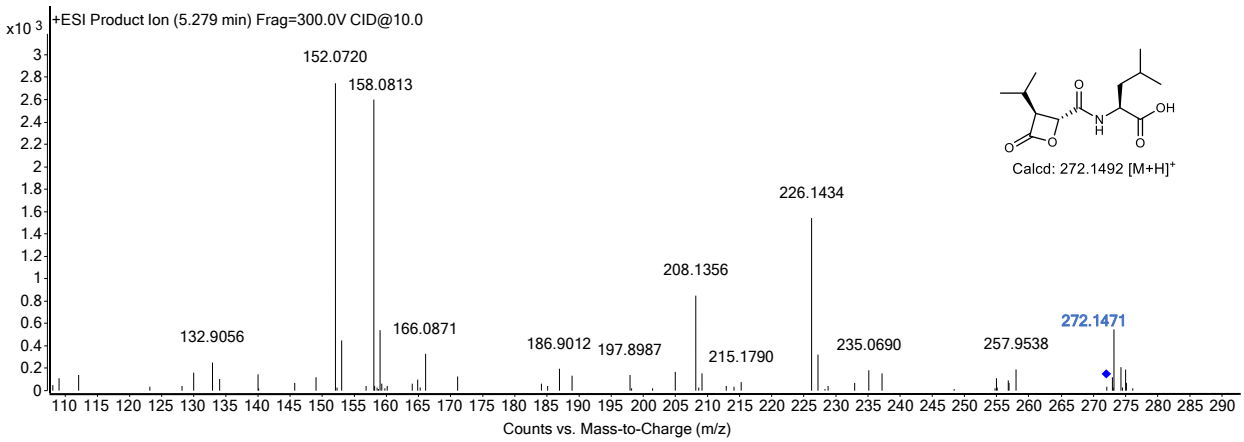
17

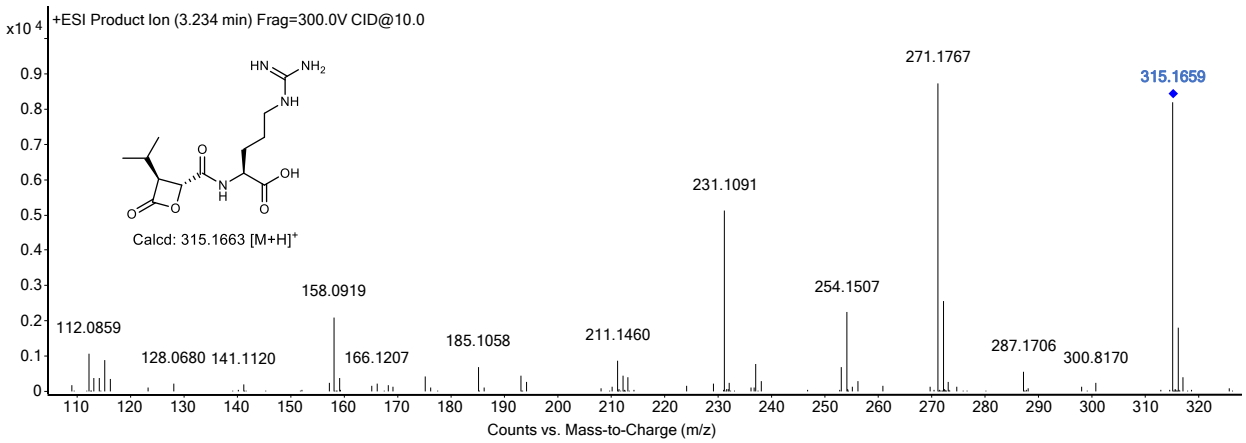
9. LC-MS/MS data for assays described in supplementary figure 18-20.

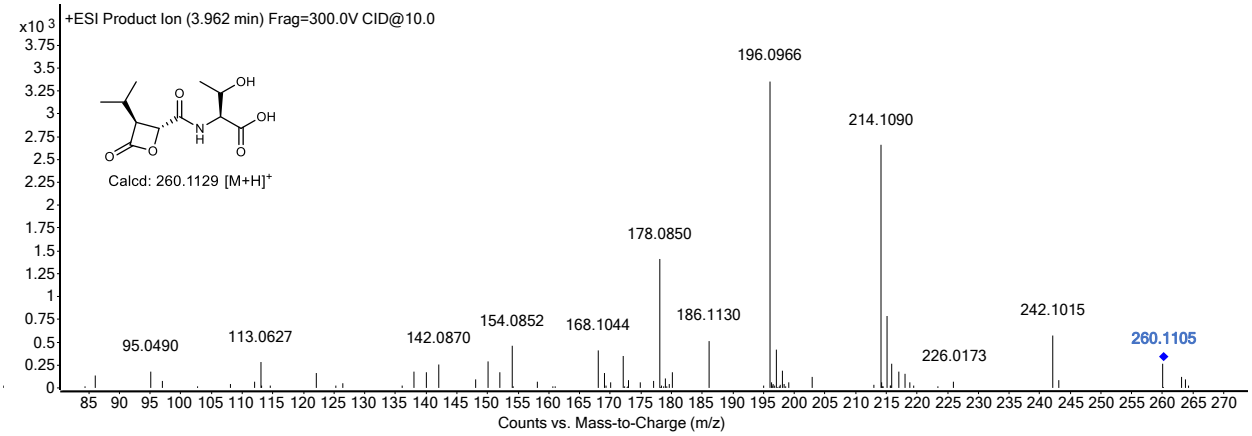
Observed mass (m/z) of precursor ion $[M+H]^+$ is indicated in blue.

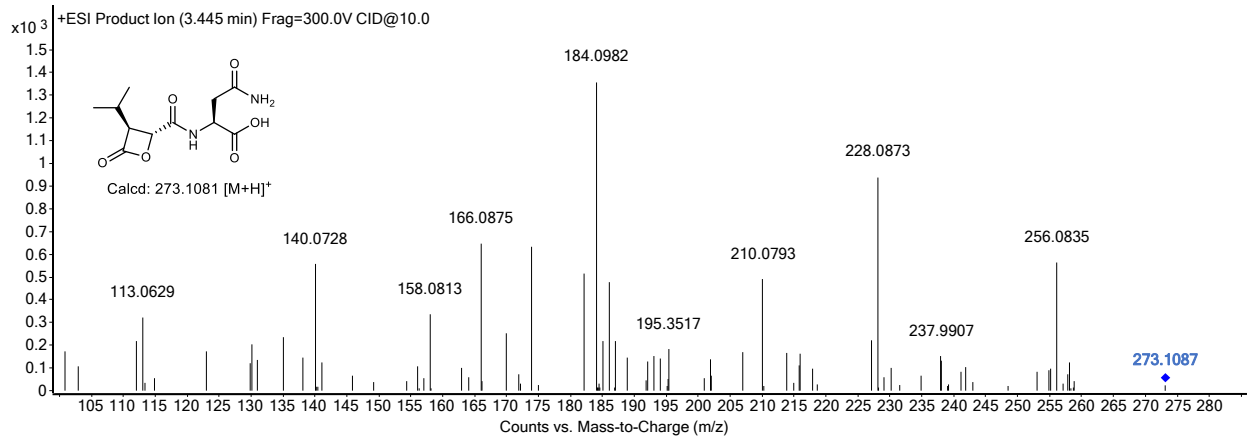


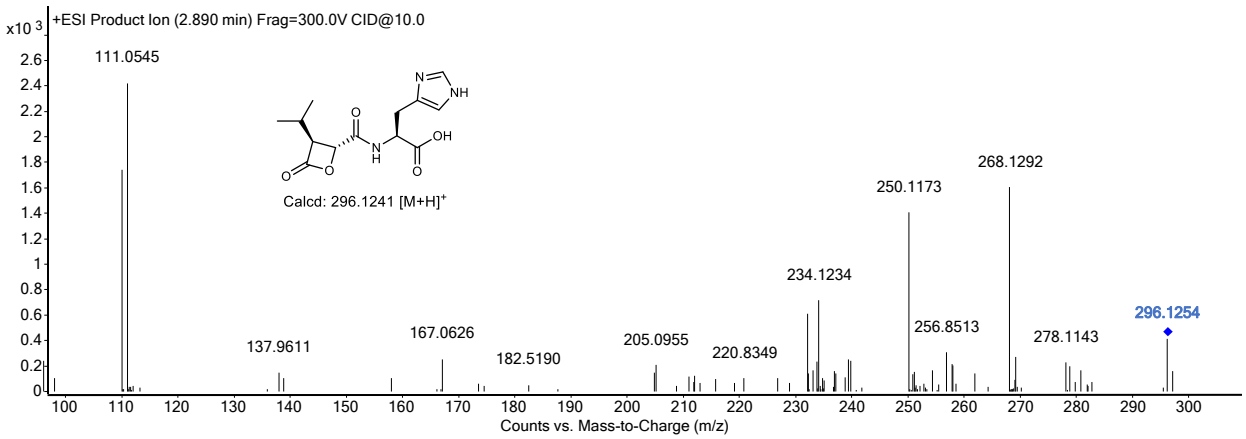


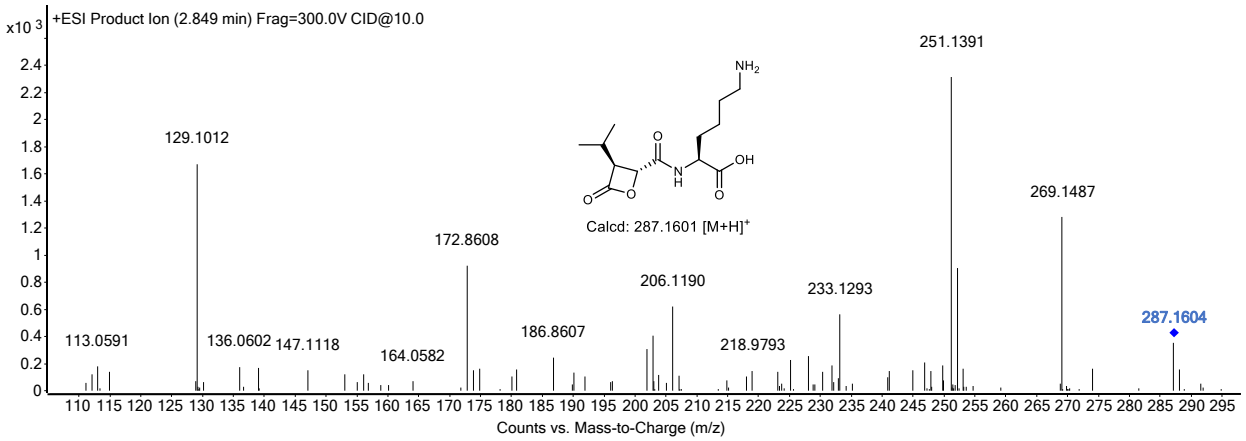


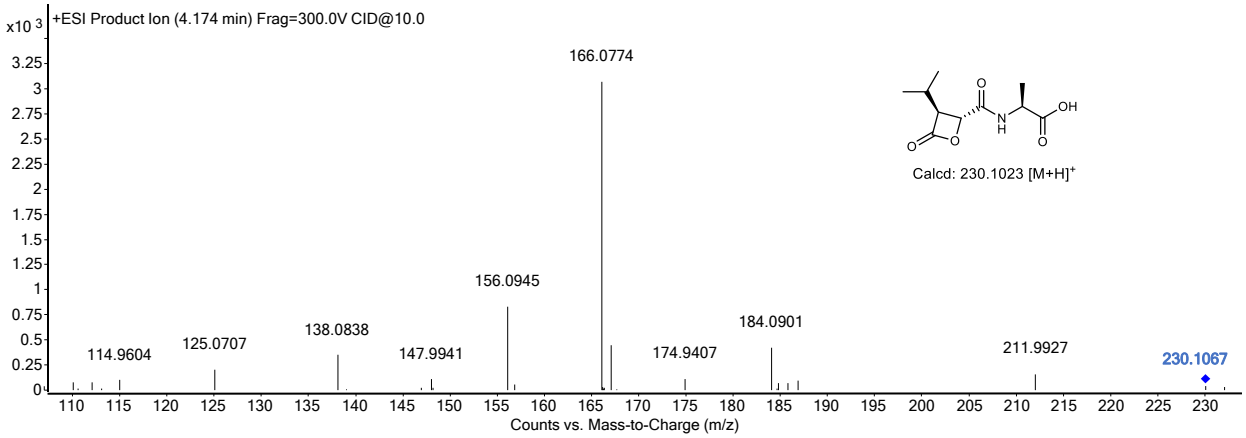

+ESI Product Ion (3.234 min) Frag=300.0V CID@10.0
x10 4
Calcd: 315.1663 [M+H]+
315.1659
271.1767
231.1091
254.1507
158.0919
112.0859
128.0680
141.1120
166.1207
185.1058
211.1460
287.1706
300.8170
Counts vs. Mass-to-Charge (m/z)

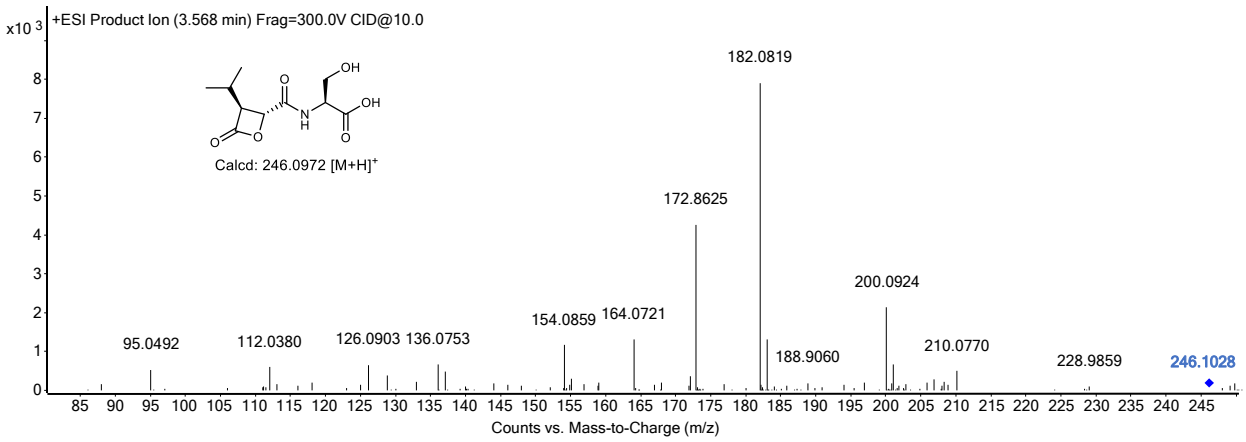

+ESI Product Ion (3.962 min) Frag=300.0V CID@10.0
x10 3
Calcd: 260.1129 [M+H]+
196.0966
214.1090
178.0850
260.1105
95.0490
113.0627
142.0870
154.0852
168.1044
186.1130
226.0173
242.1015
Counts vs. Mass-to-Charge (m/z)

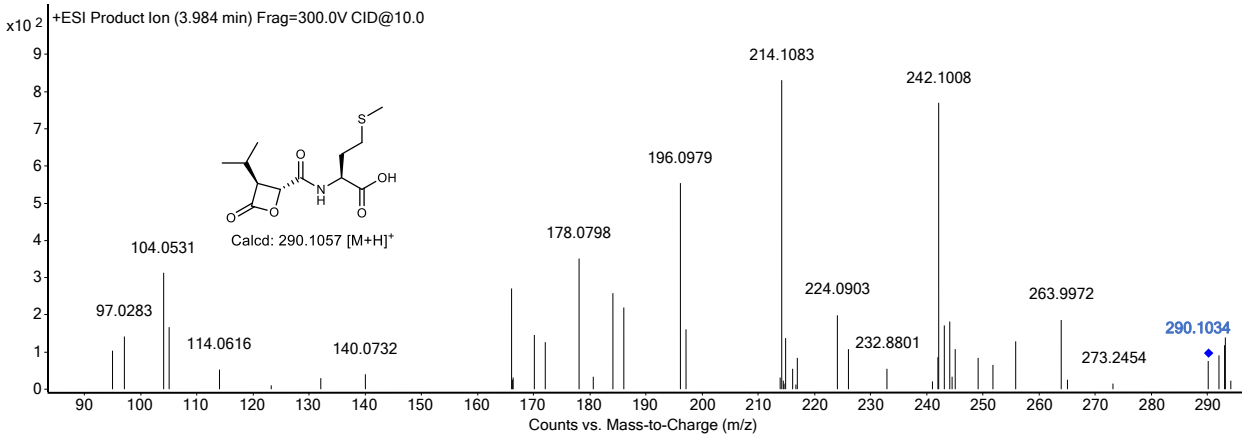

+ESI Product Ion (3.445 min) Frag=300.0V CID@10.0
x10 3
Calcd: 273.1081 [M+H]+
184.0982
228.0873
166.0875
140.0728
210.0793
256.0835
113.0629
158.0813
195.3517
237.9907
273.1087
Counts vs. Mass-to-Charge (m/z)

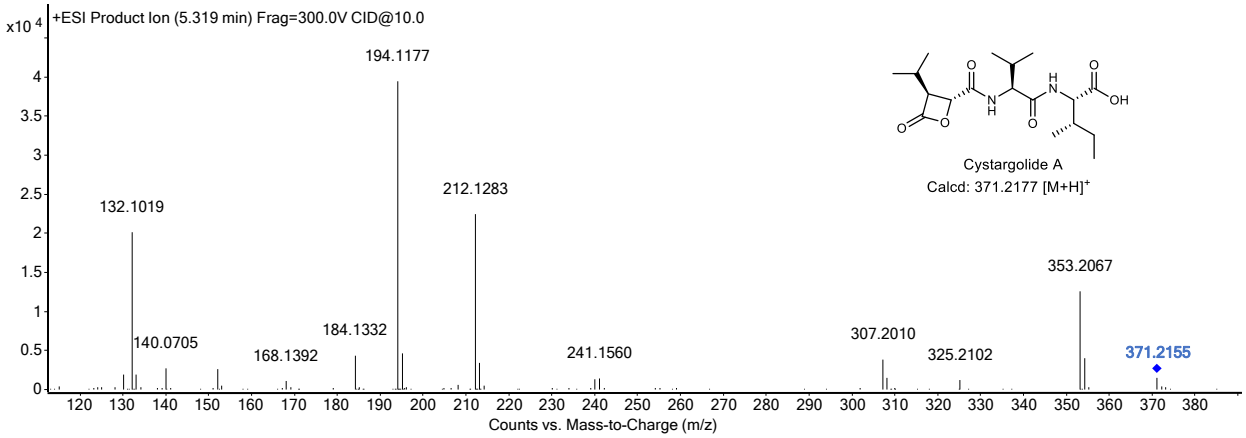

+ESI Product Ion (2.890 min) Frag=300.0V CID@10.0
x10 3
111.0545
137.9611
167.0626
182.5190
205.0955
220.8349
234.1234
250.1173
256.8513
268.1292
278.1143
296.1254
Calcd: 296.1241 [M+H]+
Counts vs. Mass-to-Charge (m/z)

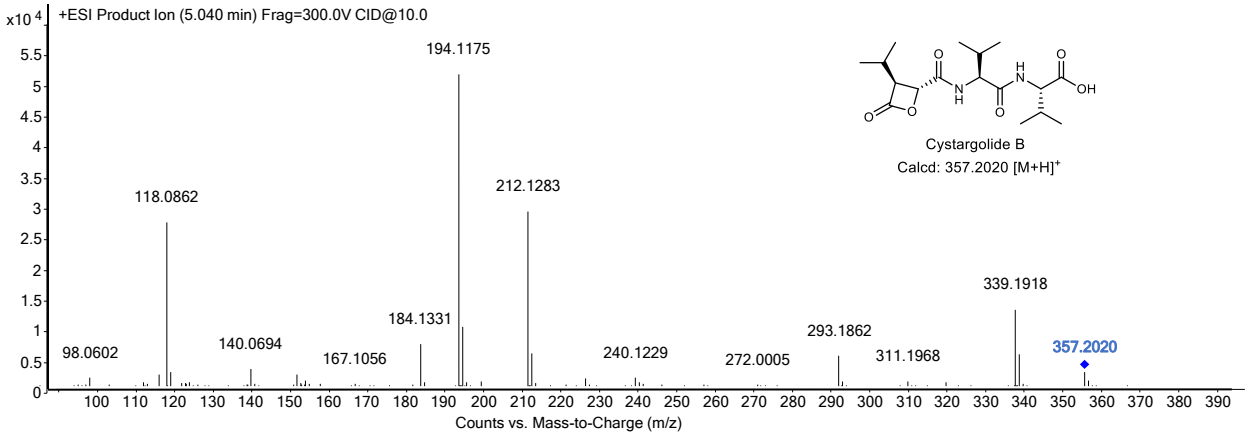

+ESI Product Ion (2.849 min) Frag=300.0V CID@10.0
x10 3
113.0591
129.1012
136.0602
147.1118
164.0582
172.8608
186.8607
206.1190
218.9793
233.1293
251.1391
269.1487
287.1604
Calcd: 287.1601 [M+H]+
Counts vs. Mass-to-Charge (m/z)

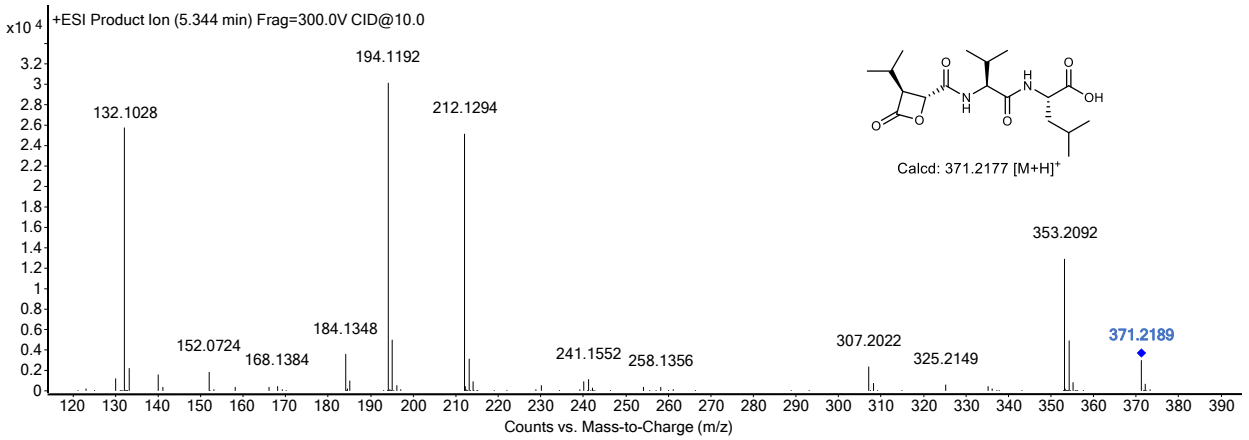

+ESI Product Ion (4.174 min) Frag=300.0V CID@10.0
x10 3
114.9604
125.0707
138.0838
147.9941
156.0945
166.0774
174.9407
184.0901
211.9927
230.1067
Calcd: 230.1023 [M+H]+
Counts vs. Mass-to-Charge (m/z)

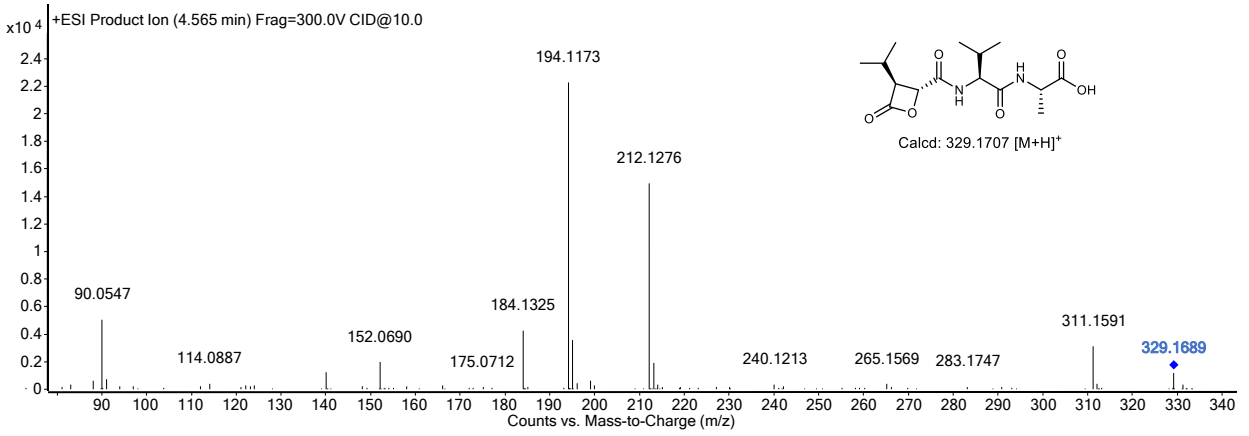

+ESI Product Ion (3.568 min) Frag=300.0V CID@10.0
x10 3
Calcd: 246.0972 [M+H]+
182.0819
172.8625
200.0924
154.0859
164.0721
95.0492
112.0380
126.0903
136.0753
188.9060
210.0770
228.9859
246.1028
Counts vs. Mass-to-Charge (m/z)

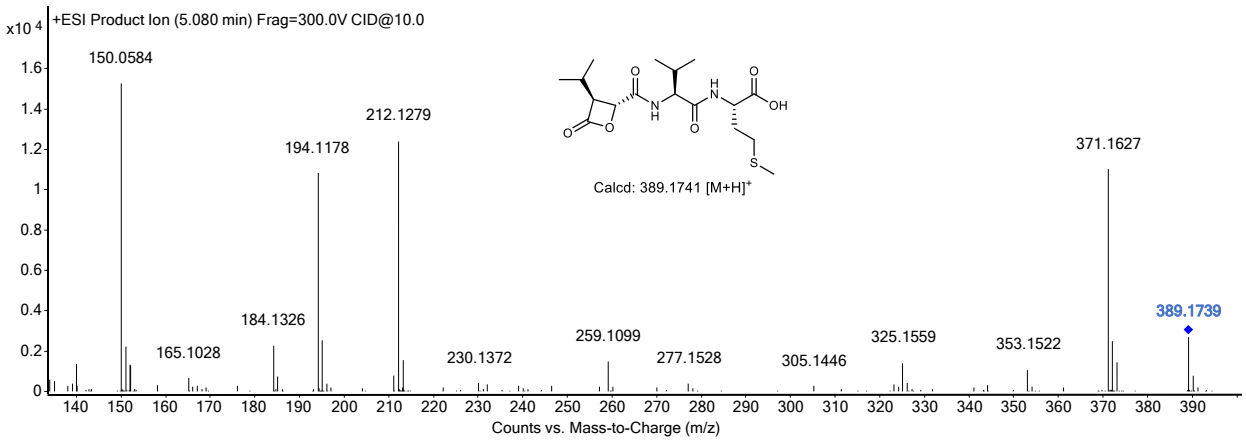

+ESI Product Ion (3.984 min) Frag=300.0V CID@10.0
x10 2
Calcd: 290.1057 [M+H]+
214.1083
242.1008
196.0979
178.0798
104.0531
224.0903
263.9972
97.0283
114.0616
140.0732
232.8801
273.2454
290.1034
Counts vs. Mass-to-Charge (m/z)

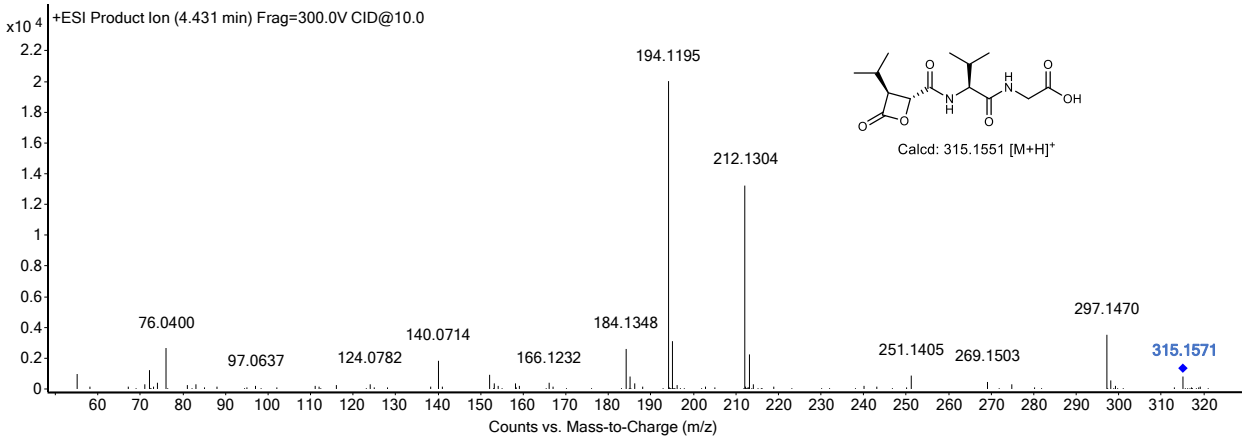

+ESI Product Ion (5.319 min) Frag=300.0V CID@10.0
x10 4
Cystargolide A
Calcd: 371.2177 [M+H]+
194.1177
212.1283
132.1019
353.2067
140.0705
184.1332
168.1392
241.1560
307.2010
325.2102
371.2155
Counts vs. Mass-to-Charge (m/z)

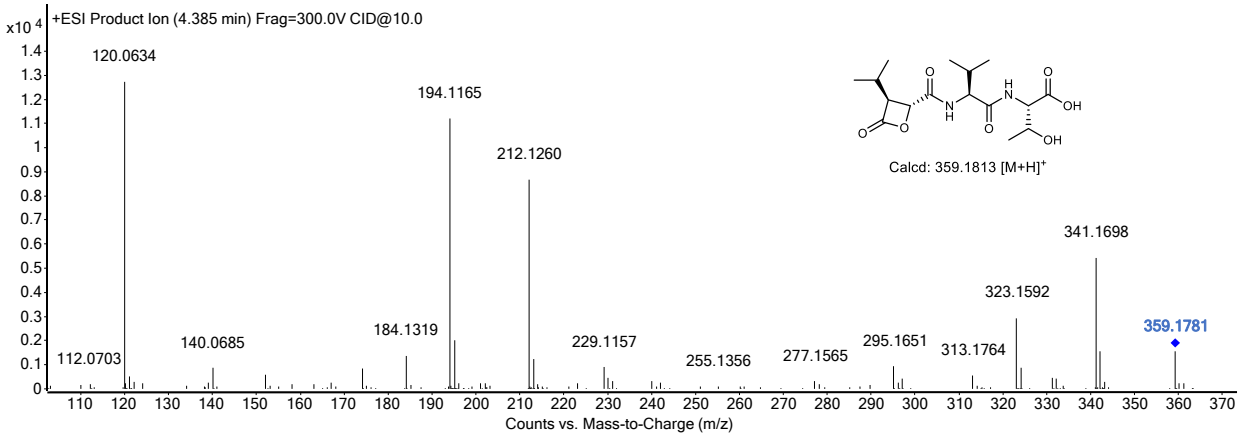

x10 4
+ESI Product Ion (5.040 min) Frag=300.0V CID@10.0
194.1175
212.1283
118.0862
339.1918
184.1331
293.1862
140.0694
98.0602
167.1056
240.1229
272.0005
311.1968
357.2020
Cystargolide B
Calcd: 357.2020 [M+H]+
Counts vs. Mass-to-Charge (m/z)

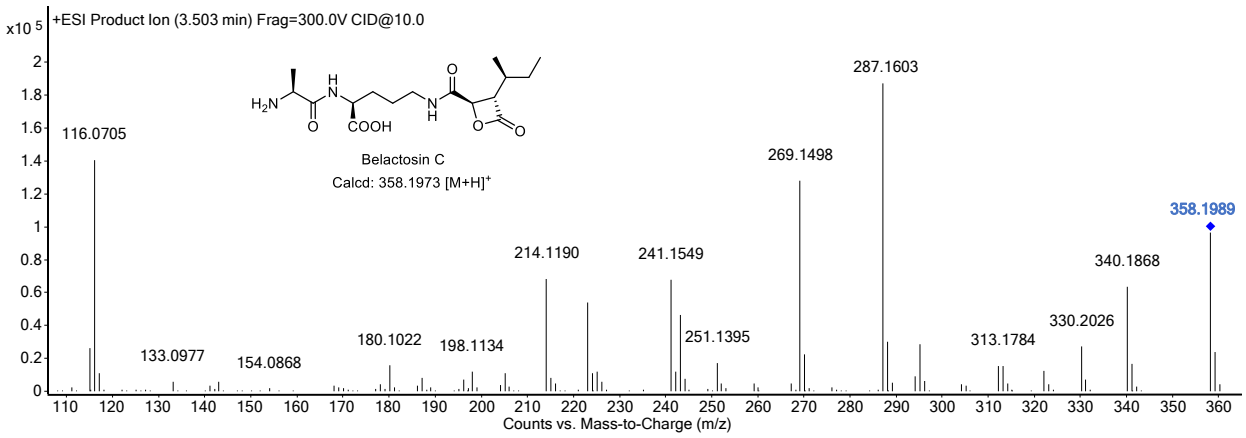

x10 4
+ESI Product Ion (5.344 min) Frag=300.0V CID@10.0
194.1192
212.1294
132.1028
353.2092
184.1348
152.0724
168.1384
241.1552
258.1356
307.2022
325.2149
371.2189
Calcd: 371.2177 [M+H]+
Counts vs. Mass-to-Charge (m/z)

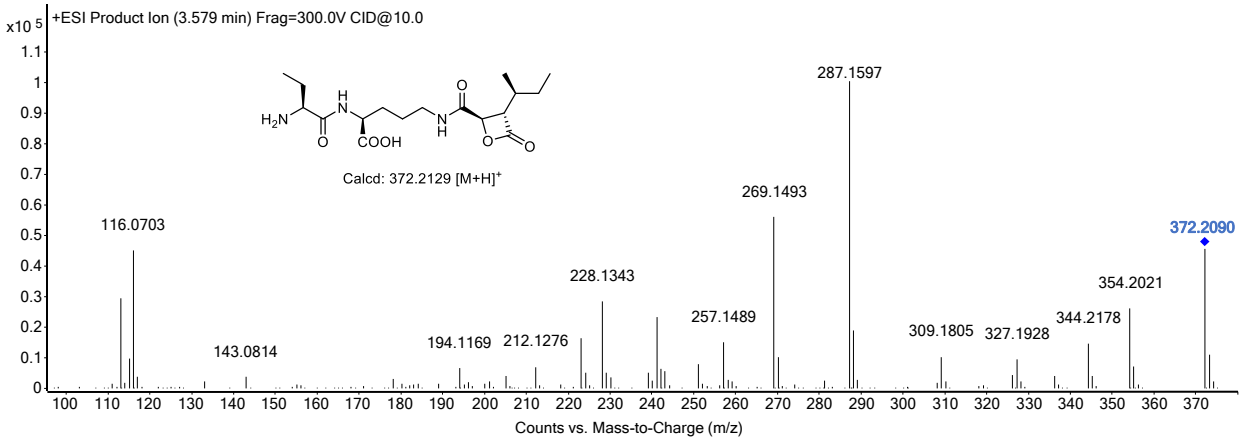

x10 4
+ESI Product Ion (4.565 min) Frag=300.0V CID@10.0
194.1173
212.1276
90.0547
184.1325
311.1591
152.0690
329.1689
114.0887
175.0712
240.1213
265.1569
283.1747
Calcd: 329.1707 [M+H]+
Counts vs. Mass-to-Charge (m/z)

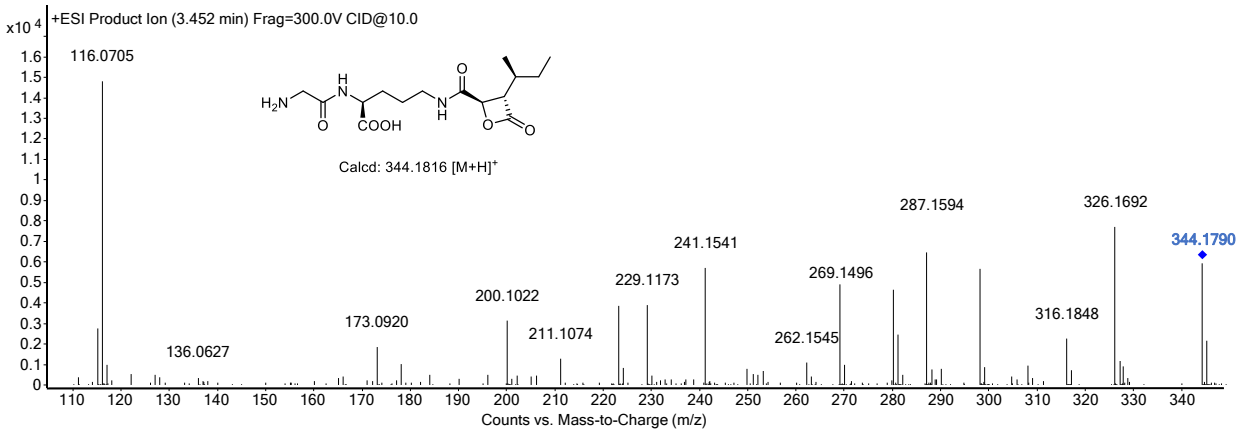

+ESI Product Ion (5.080 min) Frag=300.0V CID@10.0
x10 4
150.0584
212.1279
194.1178
371.1627
184.1326
165.1028
259.1099
230.1372
277.1528
305.1446
325.1559
353.1522
389.1739
Calcd: 389.1741 [M+H]+
Counts vs. Mass-to-Charge (m/z)

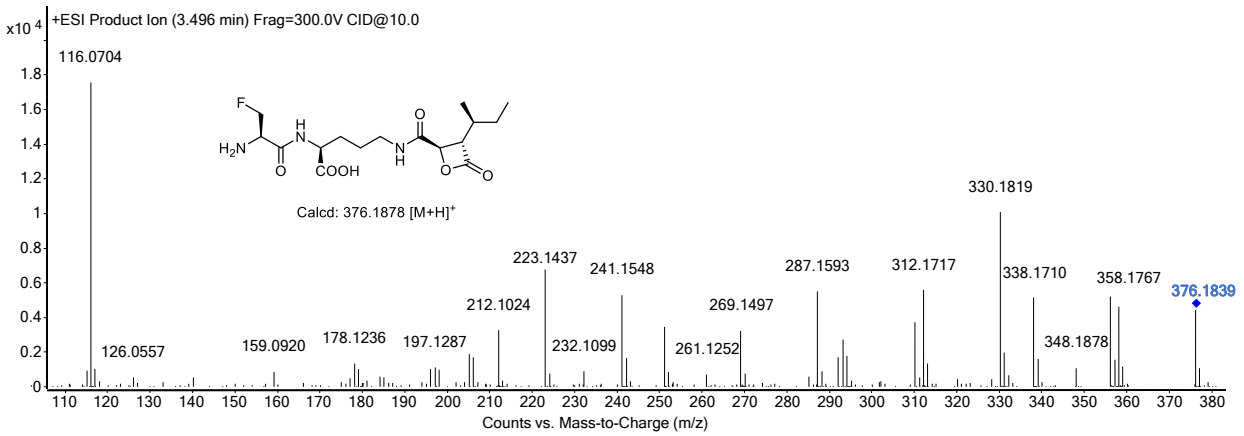

+ESI Product Ion (4.431 min) Frag=300.0V CID@10.0
x10 4
194.1195
212.1304
76.0400
140.0714
184.1348
97.0637
124.0782
166.1232
251.1405
269.1503
297.1470
315.1571
Calcd: 315.1551 [M+H]+
Counts vs. Mass-to-Charge (m/z)

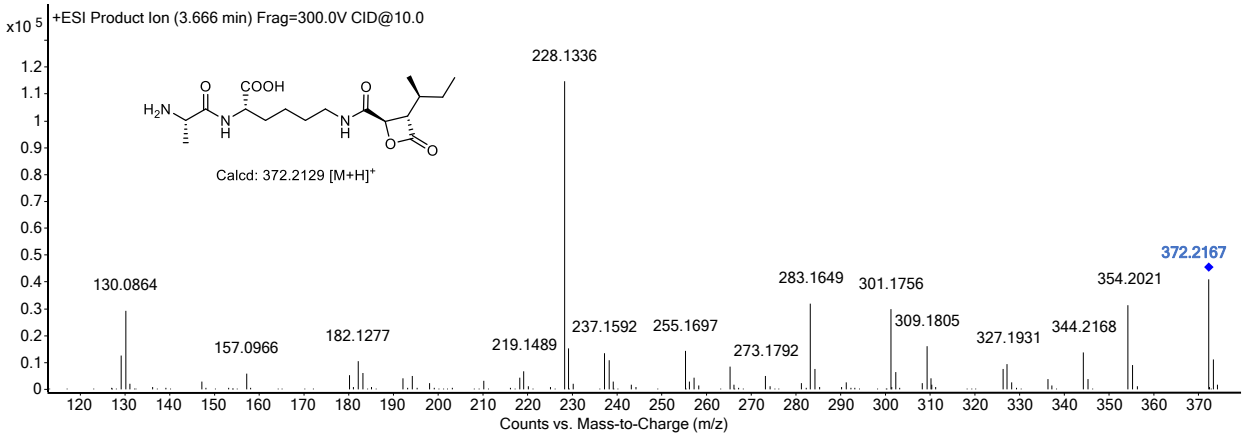

+ESI Product Ion (4.385 min) Frag=300.0V CID@10.0
x10 4
120.0634
194.1165
212.1260
341.1698
323.1592
184.1319
140.0685
229.1157
112.0703
255.1356
277.1565
295.1651
313.1764
359.1781
Calcd: 359.1813 [M+H]+
Counts vs. Mass-to-Charge (m/z)

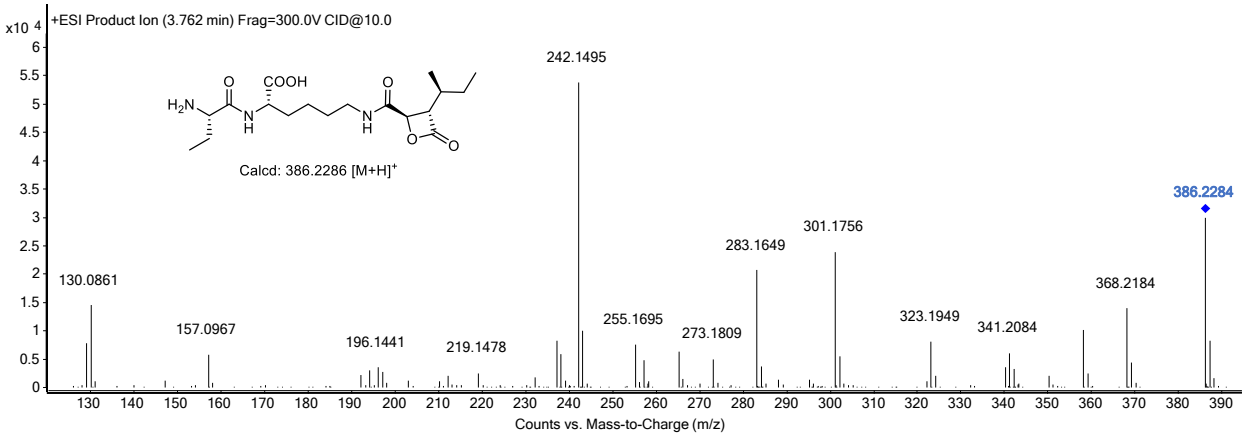

+ESI Product Ion (3.503 min) Frag=300.0V CID@10.0
x10 5
Belactosin C
Calcd: 358.1973 [M+H]+
116.0705
133.0977
154.0868
180.1022
198.1134
214.1190
241.1549
251.1395
269.1498
287.1603
313.1784
330.2026
340.1868
358.1989
Counts vs. Mass-to-Charge (m/z)

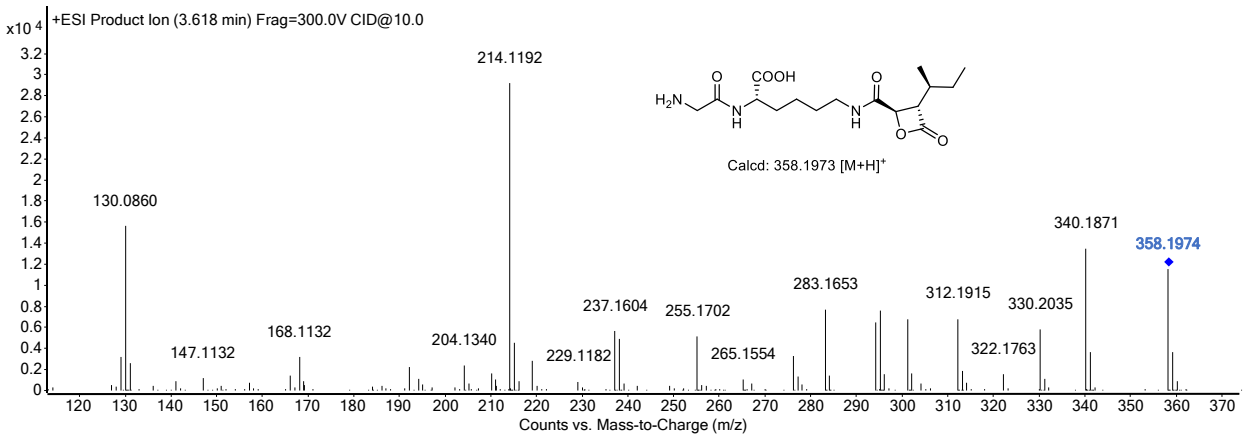

+ESI Product Ion (3.579 min) Frag=300.0V CID@10.0
x10 5
Calcd: 372.2129 [M+H]+
116.0703
143.0814
194.1169
212.1276
228.1343
257.1489
269.1493
287.1597
309.1805
327.1928
344.2178
354.2021
372.2090
Counts vs. Mass-to-Charge (m/z)

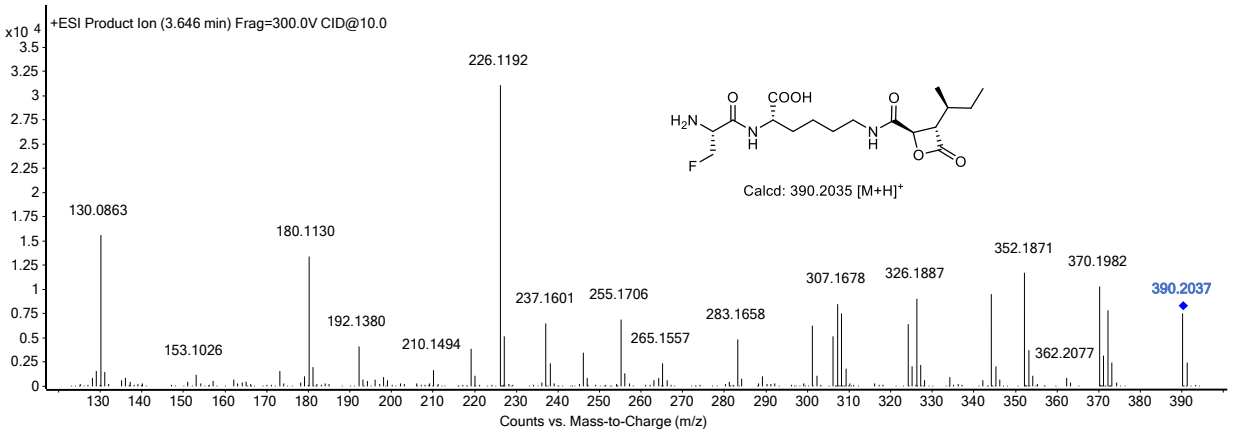

+ESI Product Ion (3.452 min) Frag=300.0V CID@10.0
x10 4
Calcd: 344.1816 [M+H]+
116.0705
136.0627
173.0920
200.1022
211.1074
229.1173
241.1541
262.1545
269.1496
287.1594
316.1848
326.1692
344.1790
Counts vs. Mass-to-Charge (m/z)

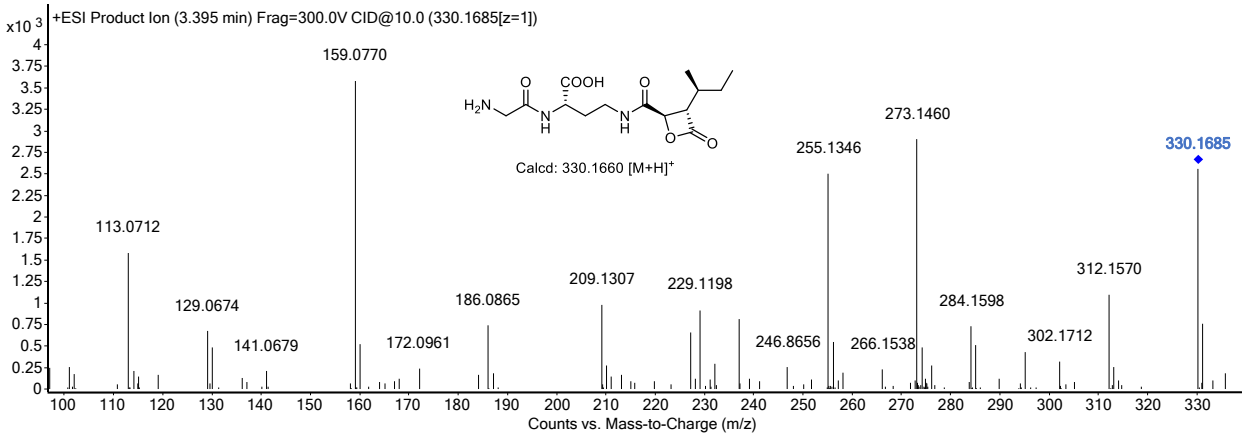

+ESI Product Ion (3.496 min) Frag=300.0V CID@10.0
x10 4
Calcd: 376.1878 [M+H]+
116.0704
330.1819
223.1437
241.1548
287.1593
312.1717
338.1710
358.1767
376.1839
212.1024
269.1497
126.0557
159.0920
178.1236
197.1287
232.1099
261.1252
348.1878
Counts vs. Mass-to-Charge (m/z)

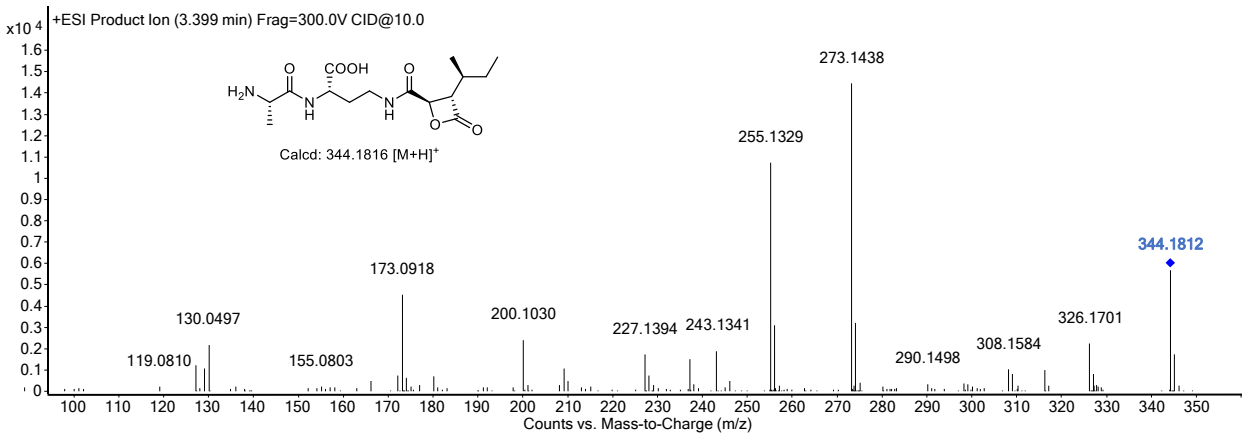

+ESI Product Ion (3.666 min) Frag=300.0V CID@10.0
x10 5
Calcd: 372.2129 [M+H]+
228.1336
372.2167
130.0864
283.1649
301.1756
354.2021
309.1805
157.0966
182.1277
219.1489
237.1592
255.1697
273.1792
327.1931
344.2168
Counts vs. Mass-to-Charge (m/z)

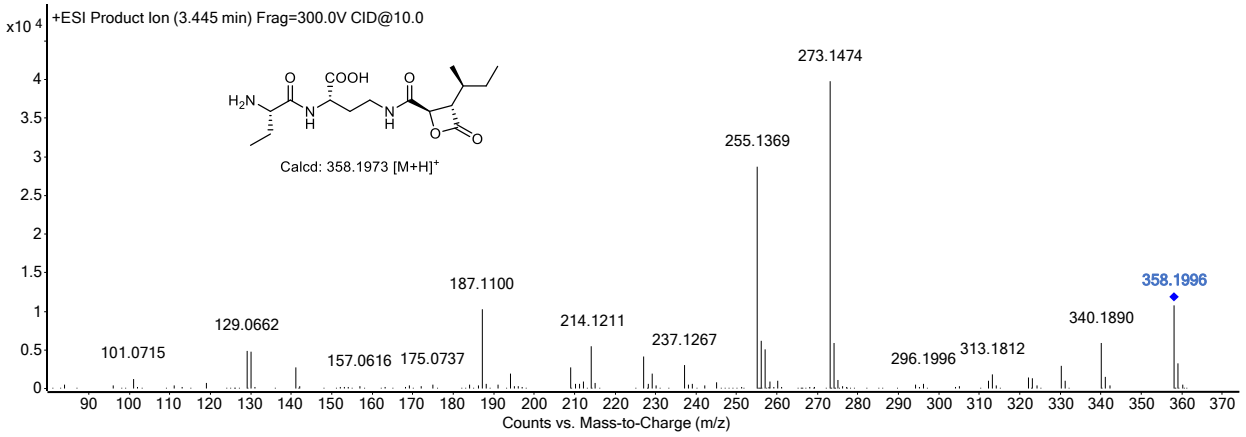

+ESI Product Ion (3.762 min) Frag=300.0V CID@10.0
x10 4
Calcd: 386.2286 [M+H]+
242.1495
386.2284
301.1756
283.1649
130.0861
368.2184
157.0967
255.1695
323.1949
273.1809
341.2084
196.1441
219.1478
Counts vs. Mass-to-Charge (m/z)

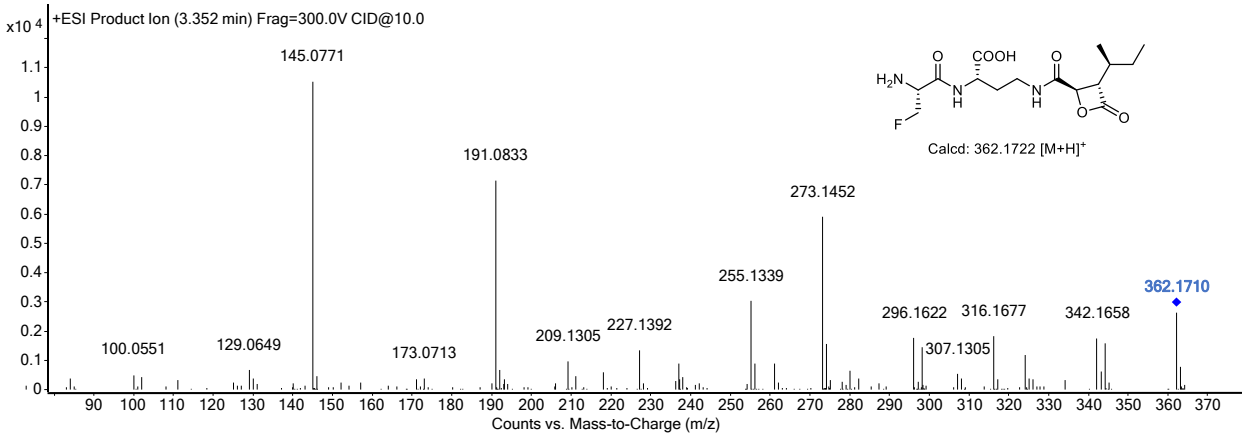

+ESI Product Ion (3.618 min) Frag=300.0V CID@10.0
x10 4
214.1192
130.0860
340.1871
358.1974
283.1653
312.1915
330.2035
237.1604
255.1702
168.1132
147.1132
204.1340
229.1182
265.1554
322.1763
Calcd: 358.1973 [M+H]+
Counts vs. Mass-to-Charge (m/z)

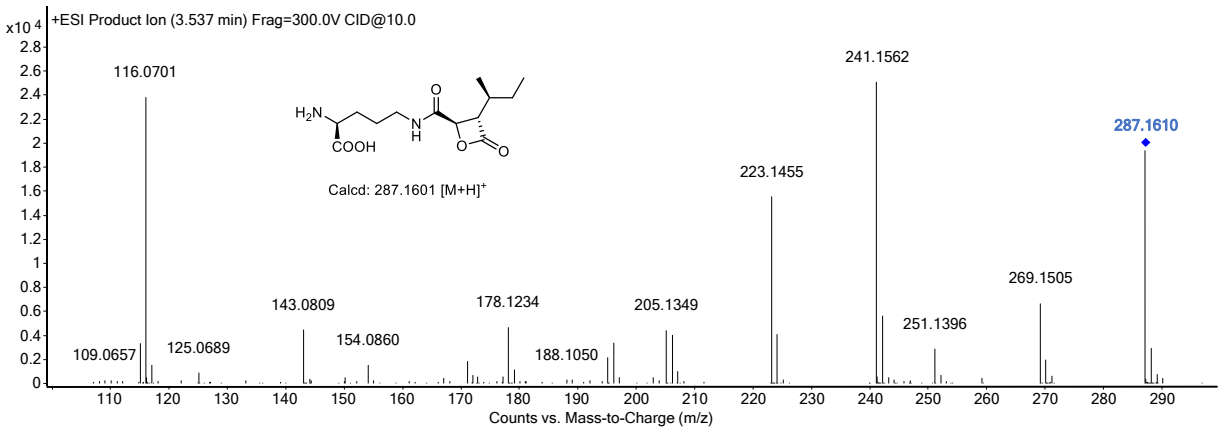

+ESI Product Ion (3.646 min) Frag=300.0V CID@10.0
x10 4
226.1192
130.0863
180.1130
352.1871
370.1982
307.1678
326.1887
390.2037
237.1601
255.1706
283.1658
192.1380
265.1557
153.1026
210.1494
362.2077
Calcd: 390.2035 [M+H]+
Counts vs. Mass-to-Charge (m/z)

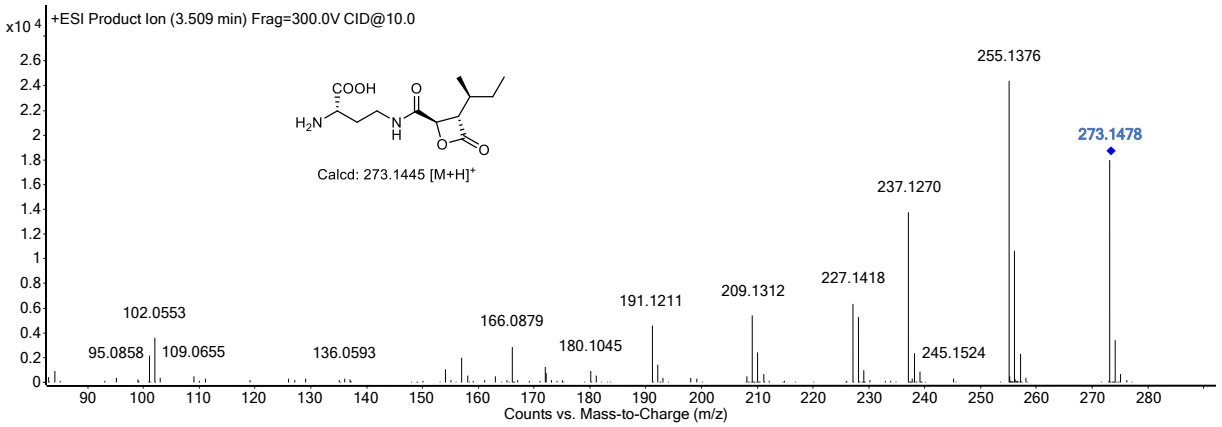

+ESI Product Ion (3.395 min) Frag=300.0V CID@10.0 (330.1685[z=1])
x10 3
159.0770
273.1460
330.1685
255.1346
113.0712
312.1570
209.1307
229.1198
186.0865
284.1598
129.0674
246.8656
266.1538
302.1712
141.0679
172.0961
Calcd: 330.1660 [M+H]+
Counts vs. Mass-to-Charge (m/z)

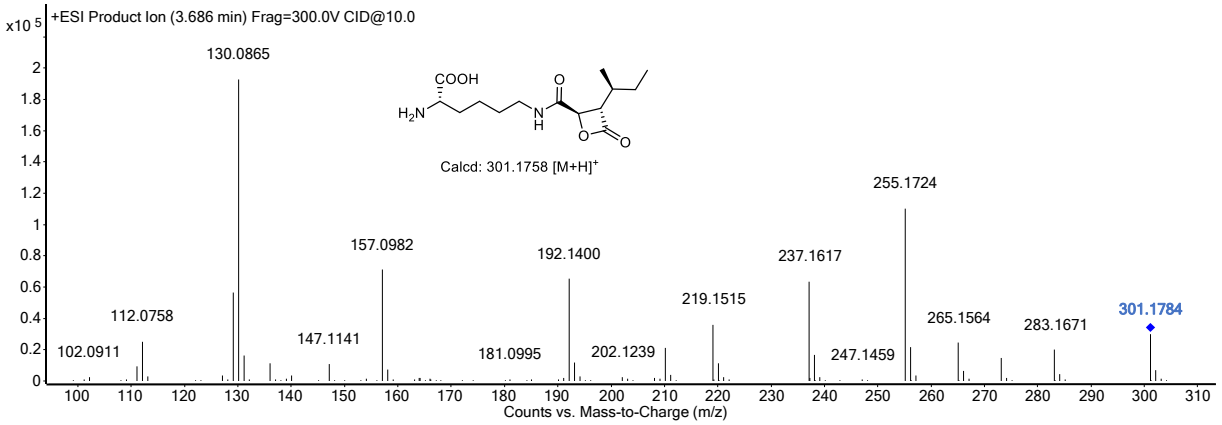

+ESI Product Ion (3.399 min) Frag=300.0V CID@10.0
x10 4
Calcd: 344.1816 [M+H]+
273.1438
255.1329
344.1812
173.0918
130.0497
200.1030
227.1394
243.1341
326.1701
308.1584
119.0810
155.0803
290.1498
Counts vs. Mass-to-Charge (m/z)

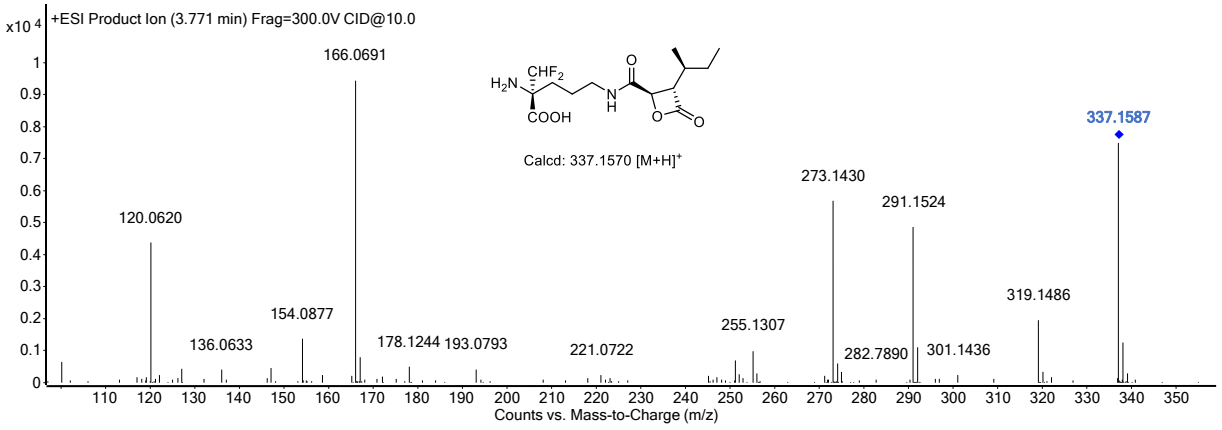

+ESI Product Ion (3.445 min) Frag=300.0V CID@10.0
x10 4
Calcd: 358.1973 [M+H]+
273.1474
255.1369
358.1996
187.1100
129.0662
214.1211
340.1890
237.1267
101.0715
157.0616
175.0737
296.1996
313.1812
Counts vs. Mass-to-Charge (m/z)


+ESI Product Ion (3.352 min) Frag=300.0V CID@10.0
x10 4
Calcd: 362.1722 [M+H]+
145.0771
191.0833
273.1452
255.1339
362.1710
296.1622
316.1677
342.1658
209.1305
227.1392
307.1305
100.0551
129.0649
173.0713
Counts vs. Mass-to-Charge (m/z)


+ESI Product Ion (3.537 min) Frag=300.0V CID@10.0
x10 4
Calcd: 287.1601 [M+H]+
116.0701
241.1562
287.1610
223.1455
269.1505
143.0809
178.1234
205.1349
251.1396
154.0860
109.0657
125.0689
188.1050
Counts vs. Mass-to-Charge (m/z)


+ESI Product Ion (3.509 min) Frag=300.0V CID@10.0
x10 4
Calcd: 273.1445 [M+H]+
255.1376
273.1478
237.1270
227.1418
209.1312
191.1211
102.0553
166.0879
95.0858
109.0655
136.0593
180.1045
245.1524
Counts vs. Mass-to-Charge (m/z)


+ESI Product Ion (3.686 min) Frag=300.0V CID@10.0
x10 5
Calcd: 301.1758 [M+H]+
130.0865
255.1724
157.0982
192.1400
237.1617
219.1515
112.0758
265.1564
283.1671
301.1784
147.1141
102.0911
181.0995
202.1239
247.1459
Counts vs. Mass-to-Charge (m/z)


+ESI Product Ion (3.771 min) Frag=300.0V CID@10.0
x10 4
Calcd: 337.1570 [M+H]+
166.0691
337.1587
273.1430
291.1524
120.0620
319.1486
154.0877
255.1307
136.0633
178.1244
193.0793
221.0722
282.7890
301.1436
Counts vs. Mass-to-Charge (m/z)

10. Supplementary references

S1. Blake, L. I. & Cann, M. J. Carbon Dioxide and the Carbamate Post-Translational Modification. *Front. Mol. Biosci.* **9**, 825706 (2022).

S2. Zallot, R., Oberg, N. & Gerlt, J. A. The EFI Web Resource for Genomic Enzymology Tools: Leveraging Protein, Genome, and Metagenome Databases to Discover Novel Enzymes and Metabolic Pathways. *Biochemistry* **58**, 4169–4182 (2019).

S3. An, J., Totrov, M. & Abagyan, R. Pocketome via Comprehensive Identification and Classification of Ligand Binding Envelopes. *Molecular & Cellular Proteomics*, **4**, 752-761 (2005).

S4. Wen, X., Leisinger, F., Leopold, V. & Seebeck, F. P. Synthetic Reagents for Enzyme-Catalyzed Methylation. *Angew. Chem. Int. Ed.* **61**, e202208746 (2022).

S5. Armstrong, A. & Scutt, J. N. Total synthesis of (+)-belactosin A. *Chem. Commun.* **4**, 510–511 (2004).

S6. Roeder, M., Spiegelstein, O., Schurig, V., Bialer, M. & Yagen, B. Absolute configuration of the four stereoisomers of valnoctamide (2- ethyl-3-methyl valeramide), a potentially new stereospecific antiepileptic and CNS drug. *Tetrahedron Asymmetry* **10**, 841–853 (1999).

S7. Jiang, B., Shi, H. ping, Xu, M., Wang, W. jun & Zhou, W. shan. Stereoselective synthesis of Certonardolsterol D3. *Tetrahedron* **64**, 9738–9744 (2008).

S8. Kawamura, S., Unno, Y., Asai, A., Arisawa, M. & Shuto, S. Design and synthesis of the stabilized analogs of belactosin A with the unnatural cis-cyclopropane structure. *Org. Biomol. Chem.* **11**, 6615–6622 (2013).

S9. Kumaraswamy, G. *et al.* Oppolzer sultam directed aldol as a key step for the stereoselective syntheses of antitumor antibiotic belactosin C and its synthetic congeners. *J. Org. Chem.* **71**, 337–340 (2006).

S10. Pirrung, M. C., Han, H. & Nunn, D. S. Kinetic Mechanism and Reaction Pathway of Thermus thermophilus Isopropylmalate Dehydrogenase. *J. Org. Chem.* **59**, 2423–2429 (1994).

S11. Kawamura, S. *et al.* Investigation of the noncovalent binding mode of covalent proteasome inhibitors around the transition state by combined use of cyclopropylic strain-based conformational restriction and computational modeling. *J. Med. Chem.* **56**, 5829–5842 (2013).

S12. Martin, N. I., Beeson, W. T., Woodward, J. J. & Marletta, M. A. NG-aminoguanidines from primary amines and the preparation of nitric oxide synthase inhibitors. *J. Med. Chem.* **51**, 924–931 (2008).

S13. Larionov, O. V. & De Meijere, A. Enantioselective total syntheses of belactosin A, belactosin C, and its homoanalogue. *Org. Lett.* **6**, 2153–2156 (2004).