Supporting Information

Detection of active mammalian GH31 α-glucosidases in health and disease using in-class, broad spectrum activity-based probes

Jianbing Jiang,[†,#] Chi-Lin Kuo,[‡,#] Liang Wu,[§] Christian Franke,[†] Wouter W. Kallemeijn,[‡] Bogdan I. Florea,[†] Eline van Meel,[‡] Gijsbert A. van der Marel,[†] Jeroen D. C. Codée,[†] Rolf G. Boot,[‡] Gideon J. Davies,[§] Herman S. Overkleeft[†,*] and Johannes M. F. G. Aerts, [‡,*]

[†]Department of Bio-organic Synthesis, Leiden Institute of Chemistry, Leiden University, Einsteinweg 55, 2333 CC Leiden, The Netherlands.

[‡]Department of Medical Biochemistry, Leiden Institute of Chemistry, Leiden University, Einsteinweg 55, 2333 CC Leiden, The Netherlands.

[§]Department of Chemistry, University of York, Heslington, York, YO10 5DD, UK.

[#]These authors contributed equally to this work.

[*]e-mail: h.s.overkleeft@chem.leidenuniv.nl or j.m.f.g.aerts@lic.leidenuniv.nl

Table of contents

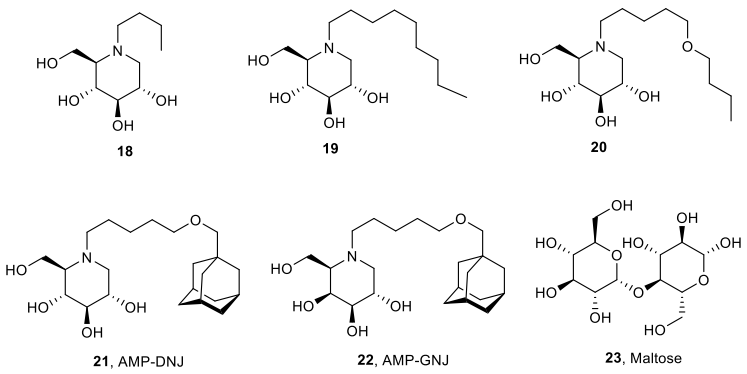


Figure S1. Chemical structures of iminosugar compounds **18**-**22** and maltose **23**.

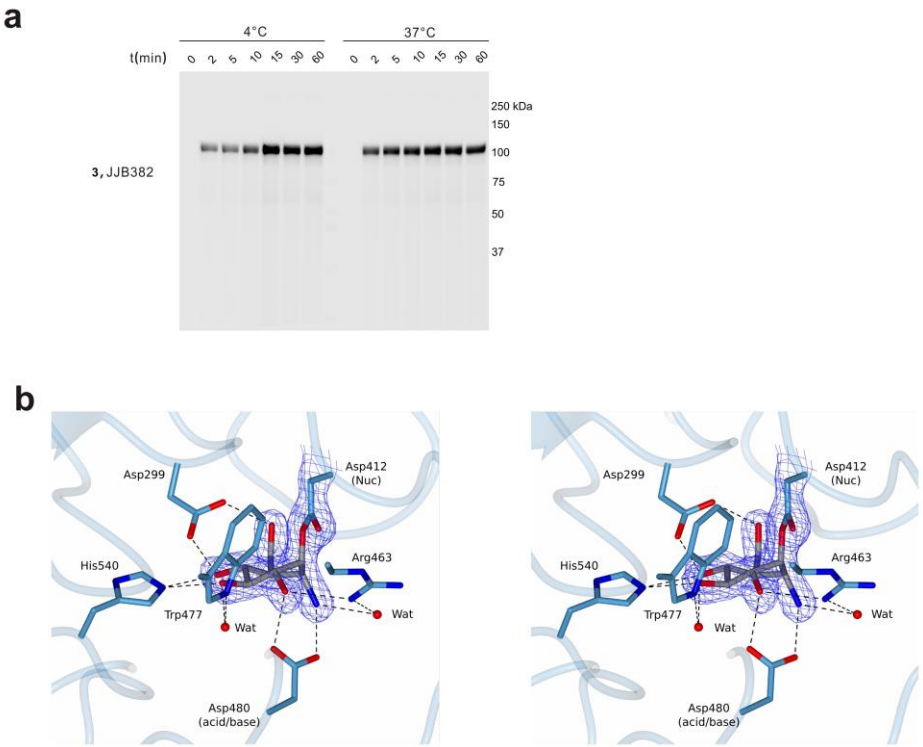
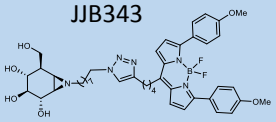


Figure S2. (**a**) Time dependent labeling of recombinant α-glucosidase with ABP **3** at 4°C and 37°C. (**b**) Stereoscopic view of the *Cj*Agd31B active site in complex with compound **1** CF021, showing covalent link to C*j*Agd31B enzymatic nucleophile Asp412, and H-bonding interactions to neighbouring residues. Electron density is REFMAC maximum-likelihood/σ_A-weighted 2 *Fo*−*Fc* synthesis contoured at 0.48 electrons per Å^3.

Table S1. Apparent IC_{50} values of α-configured cyclophellitol azridrines **1-6** and epoxide **7** JJB307 on rGBA, determined *in vitro*. JJB343 (cyclophellitol-β-gluco-aziridine-BODIPY) was included for comparison. Data were average values of two separate experiments measured in duplicate, error ranges depict standard deviation.

Compounds	*In vitro* IC_{50} (nM)
1 CF021	41,943.5 ± 490.0
2 CF 022	603.2 ± 28.0
3 JJB382	592.8 ± 214.3
4 JJB347	1,076.0 ± 63.6
5 JJB383	815.8 ± 526.4
6 JJB384	2,060.0 ± 219.2
7 JJB307	155,333.0 ± 1,773.4
JJB343 	2.05 ± 0.93

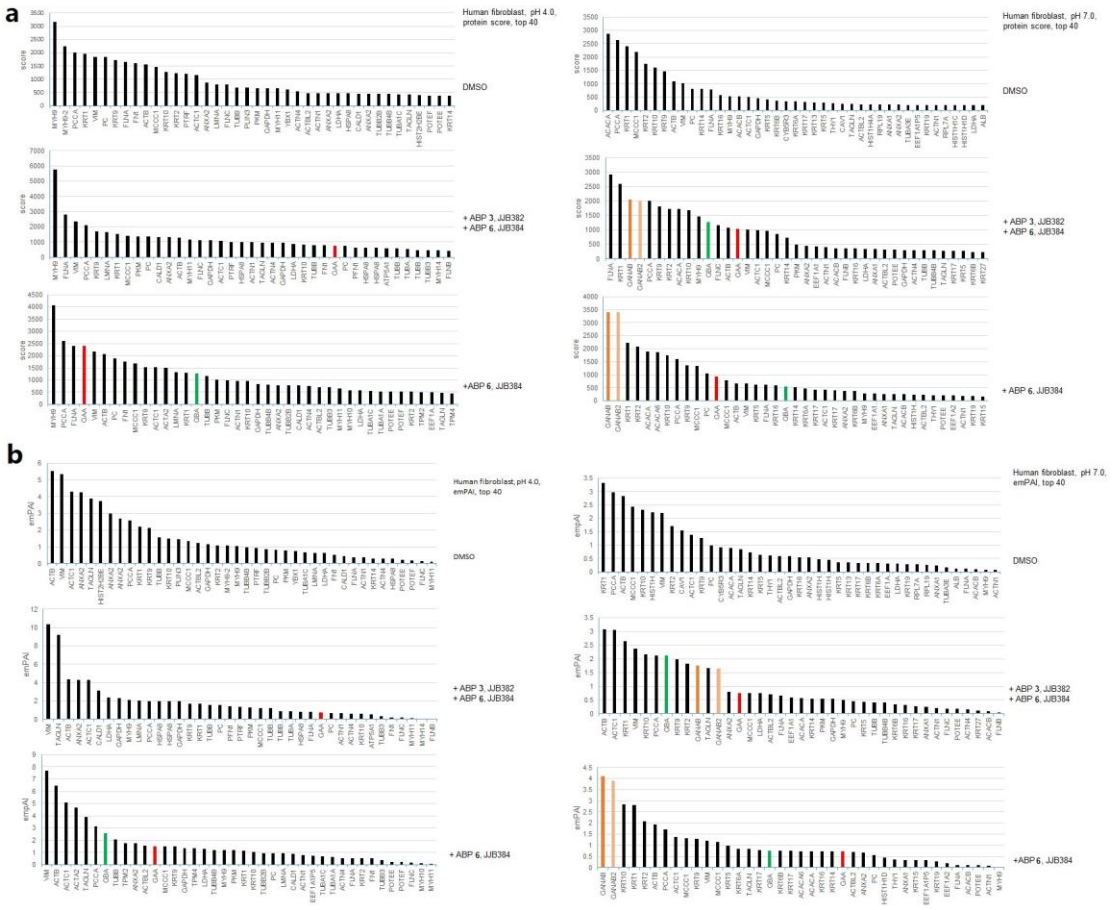
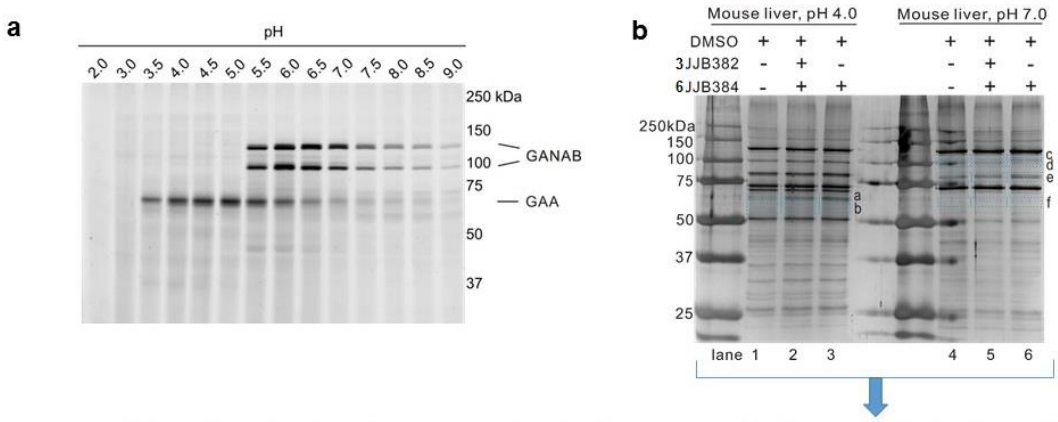


Figure S3. On-bead digest chemical proteomics data of human fibroblast lysate labeling of α-glucosidases. Protein score (**a**) and emPAI value (**b**) of top 40 proteins on-bead digestion analysis using samples of DMSO, ABP **3** competitive inhibition and ABP **6** direct labeling at pH 4.0 and pH 7.0.



Band	Rank	Accession	Protein name	Score	Matches	Coverage(%)	emPAI
2a	2	P70699	Lysosomal alpha-glucosidase, GAA	471	33	18.2	0.44
	4	P17439	Glucosylceramidase, GBA	253	20	19.6	0.47
3a	2	P70699	Lysosomal alpha-glucosidase, GAA	760	40	19.9	0.72
	4	P17439	Glucosylceramidase, GBA	238	21	19.8	0.39
2b	17	P17439	Glucosylceramidase, GBA	139	13	13.4	0.32
3b	13	P17439	Glucosylceramidase, GBA	125	11	15.3	0.18
6c	2	Q8BHN3-2	Isoform 2 of Neutral alpha-glucosidase AB, GANAB-2	909	79	46.8	1.64
6d	2	Q8BHN3	Neutral alpha-glucosidase AB, GANAB	762	62	30.0	1.06
6e	3	Q8BHN3	Neutral alpha-glucosidase AB, GANAB	641	47	23.7	0.82
5f	18	P17439	Glucosylceramidase, GBA	125	11	12.2	0.25
	20	P70699	Lysosomal alpha-glucosidase, GAA	120	8	6.3	0.06
6f	26	P70699	Lysosomal alpha-glucosidase, GAA	67	4	3.8	0.09
1a, 1b, 4c-f, 5c-e	No target enzymes						

Figure S4. (**a**) Mouse liver lysate labeling of α-glucosidases at varying pH *in vitro* with ABP **3**. (**b**) In-gel digestion silver staining gel and identification of target proteins modified by biotin-ABP **6** with mass spectrometry and Mascot search analysis, comparing with DMSO and ABP **3** competitive inhibition samples.

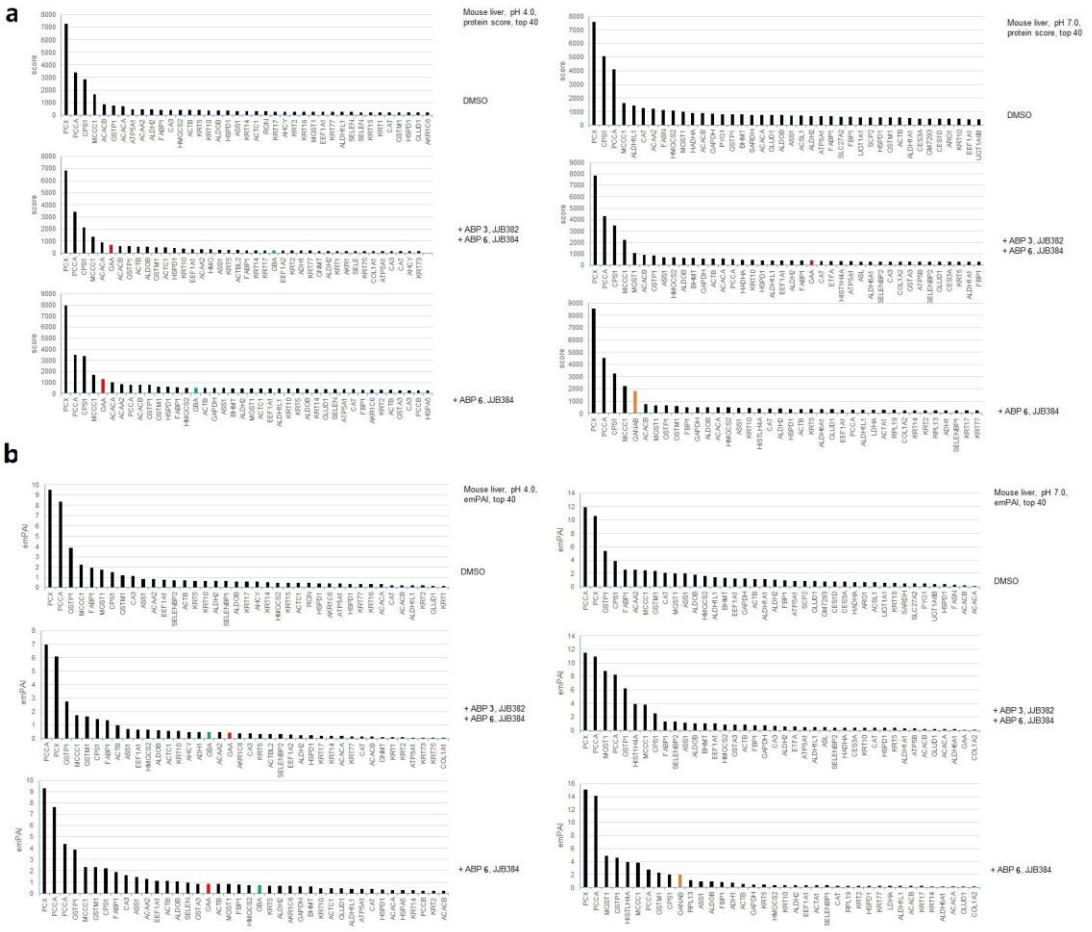
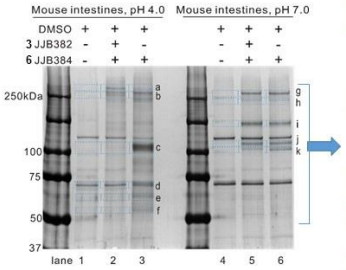


Figure S5. On-bead digest chemical proteomics data of mouse liver lysate labeling of α-glucosidases. Protein score (**a**) and emPAI value (**b**) of top 40 proteins on-bead digestion analysis using samples of DMSO, ABP **3** competitive inhibition and ABP **6** direct labeling at pH 4.0 and pH 7.0.

a



Band	Rank	Accession	Protein name	Score	Matches	Coverage(%)	emPAI
2a	4	B5THE2	Maltase-glucoamylase, MGAM	312	35	12.3	0.18
	30	B5THE3	Sucrase-isomaltase, Sis	43	5	2.5	0.01
3a	19	B5THE2	Maltase-glucoamylase, MGAM	160	12	4.5	0.07
	20	B5THE3	Sucrase-isomaltase, Sis	151	11	4.2	0.06
3b	16	B5THE3	Sucrase-isomaltase, Sis	292	29	12.3	0.13
	31	B5THE2	Maltase-glucoamylase, MGAM	125	11	15.3	0.18
3c	1	B5THE3	Sucrase-isomaltase, Sis	1260	51	14.4	0.37
	33	B5THE2	Maltase-glucoamylase, MGAM	129	9	3.4	0.04
3d	12	B5THE3	Sucrase-isomaltase, Sis	211	19	10.3	0.09
	23	B5THE2	Maltase-glucoamylase, MGAM	124	9	4.9	0.04
	29	A2AFL5	Lysosomal alpha-glucosidase, GAA	55	1	6.2	0.09
3e	12	B5THE3	Sucrase-isomaltase, Sis	309	20	6.5	0.1
	17	B5THE2	Maltase-glucoamylase, MGAM	241	12	4.5	0.07
3f	7	B5THE2	Maltase-glucoamylase, MGAM	505	30	12.8	0.2
	20	B5THE3	Sucrase-isomaltase, Sis	265	19	8	0.1
	72	P17439	Glucosylceramidase, GBA	28	2	2.9	0.05
5g	1	B5THE3	Sucrase-isomaltase, Sis	468	45	13.9	0.28
6g	4	B5THE3	Sucrase-isomaltase, Sis	421	47	15.3	0.3
5h	18	B5THE3	Sucrase-isomaltase, Sis	190	9	3.7	0.06
6h	1	B5THE3	Sucrase-isomaltase, Sis	302	28	10.3	0.18
5i	6	F8VPT3	Protein Lct, Lct	567	54	13.3	0.27
6i	1	F8VPT3	Protein Lct, Lct	549	65	14.1	0.32
5j	11	B5THE3	Sucrase-isomaltase, Sis	553	36	11.8	0.23
	43	B5THE2	Maltase-glucoamylase, MGAM	53	8	2.9	0.03
6j	2	B5THE3	Sucrase-isomaltase, Sis	605	52	16.1	0.3
5k	2	B5THE3	Sucrase-isomaltase, Sis	551	33	11.2	0.24
6k	1	Q8BHN3	Neutral alpha-glucosidase AB, GANAB	683	65	34.4	1.0
	2	B5THE3	Sucrase-isomaltase, Sis	614	39	12.3	0.23
1a-f, 2b-f, 4g-k	No target enzymes						

b

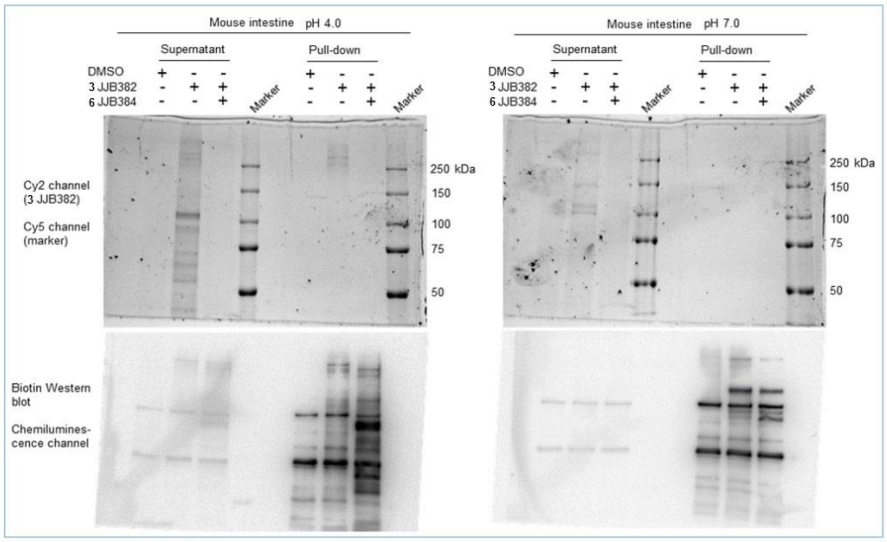


Figure S6. Mouse intestine lysate labeling of α-glucosidases and their identification by proteomics. (**a**) In-gel digestion silver staining gel and identification of target proteins modified by biotin-ABP **6** with mass spectrometry and Mascot search analysis, comparing with DMSO and ABP **3** competitive inhibition samples. (**b**) Pull-down and supernatant samples labeling of α-glucosidases at pH 4.0 and pH 7.0 by fluorescent SDS-PAGE scanning for ABP **3** and Western blotting for ABP **6** analysis

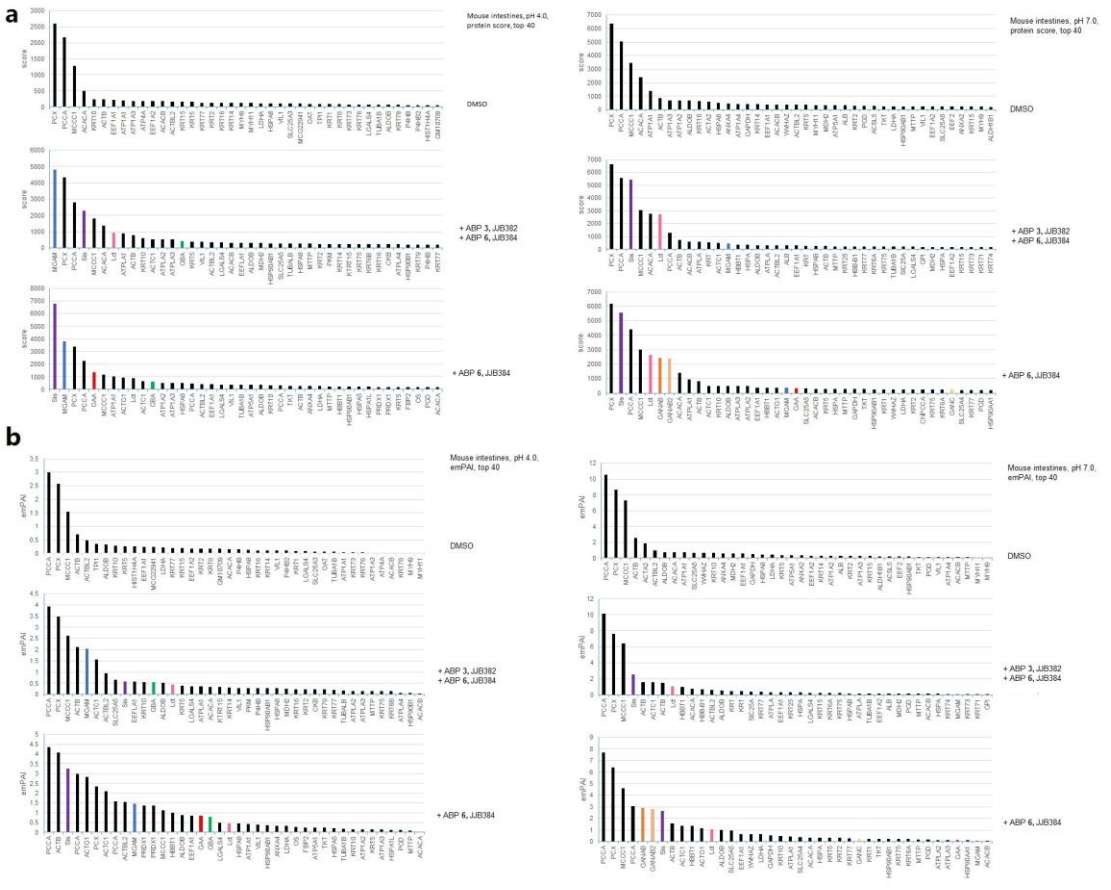


Figure S7. On-bead digest chemical proteomics data of mouse intestine lysate labeling of α-glucosidases. Protein score (**a**) and emPAI value (**b**) of top 40 proteins on-bead digestion analysis using samples of DMSO, ABP **3** competitive inhibition and ABP **6** direct labeling at pH 4.0 and pH 7.0.

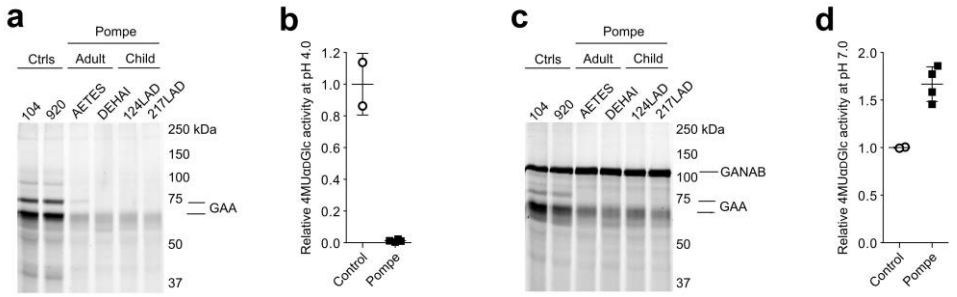


Figure S8. *In vitro* labeling of α-glucosidases in wild type and Pompe fibroblasts (**a**) *In vitro* GAA labeling at pH 4.0 with 1 µM **3** JJB382 in various fibroblast lysates, containing wild type or mutant (Pompe) GAA. (**b**) Hydrolysis of artificial 4-MU-α-D-glucopyranoside at pH 4.0. (**c**) *In vitro* GAA/GANAB labeling at pH 7.0 with 1 µM **3** JJB382 in various fibroblast lysates, containing wild type or mutant (Pompe) GAA. (**d**) Hydrolysis of artificial 4-methylumbelliferyl-α-D-glucopyranoside at pH 7.0.

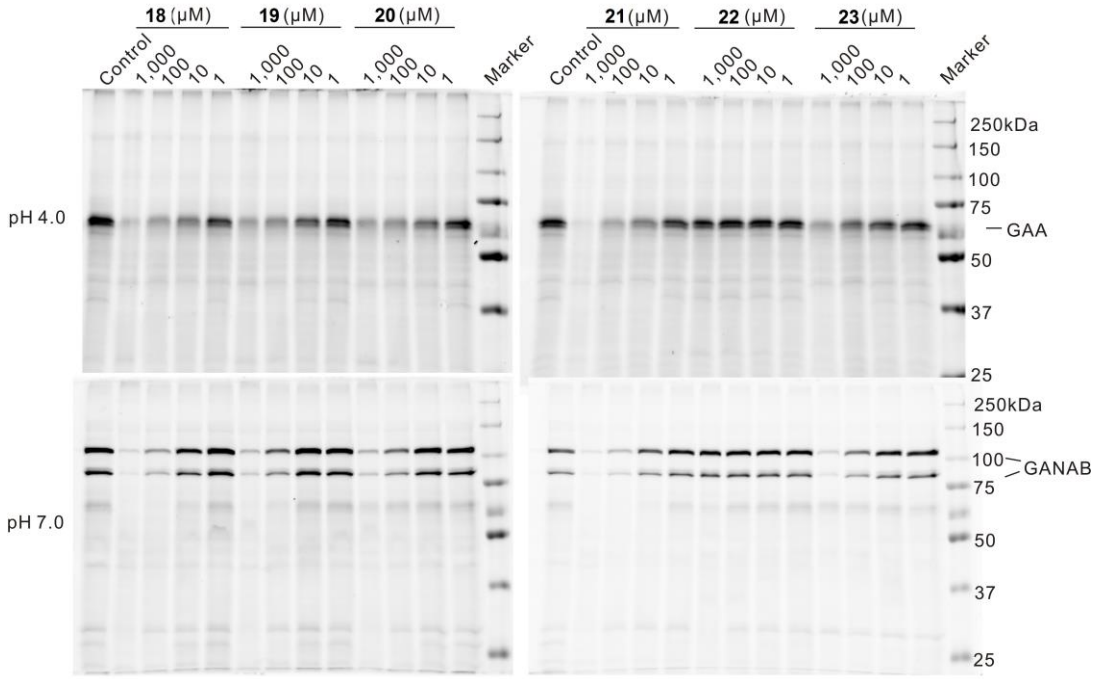


Figure S9. ABP **5** competitive labeling of α-glucosidases in mouse liver extract at pH 4.0 and pH 7.0 with pre-incubation of inhibitors **18-23**.

Experimental methods

Biological assays:

Materials

Chemicals were obtained from Sigma-Aldrich, if not otherwise indicated. Trypsin and Endoproteinase GluC were commercially available from Promega. Recombinant GAA was obtained from Genzyme (Cambridge, MA, USA). Fibroblasts were obtained with consent from donors. Pompe patients were diagnosed on the basis of reduced GAA activity. Cell lines were cultured in HAMF12-DMEM medium (Invitrogen) supplied with 10% (v/v) FCS. Mouse tissue were isolated according to guidelines approved by the ethical committee of Leiden University (DEC#13191). All the cell or tissue lysates were prepared in potassium phosphate lysis buffer(25 mM in pH 6.5, supplemented with protease inhibitor 1x cocktail (Roche)) via homogenization with silent crusher S equipped with Typ 7 F/S head (30 rpm x 1000, 3 × 7 sec) on ice and lysate concentration was determined with BCA Protein Assay Kit (Pierce™). The protein fractions were stored in small aliquots at -80 °C until use.

Enzyme activity assays and IC_{50} measurements.

The α-D-glucosidase activity of lysosomal α-D-glucosidase GAA was assayed at 37 °C by incubating with 3 mM 4-methylumbelliferyl-α-D-glucopyranoside as substrate in 150 mM McIlvaine buffer, pH 5.0, supplemented with 0.1% (w/v) BSA. Activity of rGBA was measured using similar conditions but with 3.75 mM 4-methylumbelliferyl-β-D-glucopyranoside as substrate at pH 5.2, supplemented with 0.1% (v/v) Triton X-100 and 0.2% (w/v) sodium taurocholate. To determine the apparent *in vitro* IC_{50} value, recombinant GAA or rGBA was firstly pre-incubated with a range of inhibitor dilutions for 30 min at 37 °C, prior to addition of the substrate. To determine the influence of pH on the enzymatic activity, enzyme mixtures were firstly pre-incubated for 30 min on ice with McIlvaine buffers of pH 2.0–8.0 whereafter substrate was added, dissolved in Nanopure H_2O. The enzymatic reaction was quenched by adding excess NaOH-glycine (pH 10.3), after which fluorescence of liberated 4-methylumbelliferyl was measured with a fluorimeter LS55 (Perkin Elmer) using λ_{EX} 366 nm and λ_{EM} 445 nm. The *in situ* IC_{50} value was determined by incubating fibroblast cell lines expressing wild-type GAA, grown to confluence, with a range of inhibitor dilutions for 2 h. Hereafter, cells were washed three times with PBS and subsequently harvested by scraping in potassium phosphate buffer (25 mM K_2HPO_4-KH_2PO_4, pH 6.5, supplemented with 0.1% (v/v) Triton X-100 and protease inhibitor cocktail (Roche)). Residual GAA activity was measured using the aforementioned substrate assay. All *in situ* IC_{50} values were determined by replicating each assay twice in duplo in two separate cell lines. Data was corrected for background fluorescence, then normalized to the untreated control condition and finally curve-fitted via one phase exponential decay function (GraphPad Prism 5.0).[1]

ABP 7 pull-down and LC-MS/MS analysis

3mg total protein from human fibroblast lysate or 6 mg total protein from mouse intestines lysate was incubated with either 0.1% (v/v) DMSO, 5 µM **7** JJB384, or firstly with 5 µM **3** JJB382 followed by 5 µM **6** JJB384. For mouse liver lysate, 6 mg total protein and 10µM ABPs **3** and **6** were used. Each step taking 30 min at 37 °C, in a total volume of 0.5 mL McIlvaine buffer of pH 4.0 and pH 7.0, subsequently denatured through the addition of 10% (w/v) SDS 125 µL and boiling for 5 min at 100 °C. From here on, samples were prepared for pull-down with streptavidin beads as published earlier[2]. After pull-down procedure, the samples were divided, 2/3 for on-bead digestion and 1/3 for in-gel digestion. On-bead digestion samples including beads were treated by the trypsin digestion buffer (a mixture containing 100 mM Tris-HCl pH 7.8, 100 mM NaCl, 1 mM $CaCl_2$, 2% acetonitrile and 10 ng/µL trypsin) and the bead suspension was incubated in a shaker at 37 °C overnight. The supernatant containing the

trypsin-digested peptides was desalted using stage tips, followed by evaporation of MeCN and dilution in 70 µL sample solution (H_2O/MeCN/TFA, 95/3/0.1, v/v/v) for LC-MS analysis. The beads containing active-site peptides were treated with endoproteinase Glu-C digestion buffer (100ng/µL in PBS solution); incubated in a shaker at 37 °C overnight after which the supernatant was desalted using stage tips and for LC-MS. In-gel digestion samples were eluted by boiling the beads at 100 °C in 30 µL of 1x Laemmli buffer. The eluted proteins were separated on 10% protein gels at 200 V for 1 h, and the protein gels were silver stained using the Invitrogen kit,[2] and visualized by Bio-rad Chemi-Doc MP Imager using the silver stain channel. Bands were excised with a surgery knife by hand and treated with in-gel digestion buffer (10 mM NH_4HCO_3, 5% ACN, 1mM $CaCl_2$, 10 ng/µL trypsin). The supernatant containing the trypsin-digested peptides was desalted using Stage Tips and prepared for LC-MS. All the peptide samples were analyzed with a 2h gradient of 5-25% ACN on nano-LC, hyphenated to an LTQ-Orbitrap and identified via the Mascot protein search engine, and the Raw data was calculated by MaxQuant program against the Uniprot human or mouse proteome database to present the protein identification list.[2] Mascot identifications were manually validated. The identification results were exported as Excel file including protein accession numbers, Mascot peptide scores, mass of the protein, % coverage of the protein by amino acids identified by LC-MS, peptide matches, miss cleavages, C-terminal peptides and protein emPAI values (see **Supporting Information Data** of Excel in compressed archive file(.zip)).

SDS-PAGE analysis and fluorescence scanning

For labeling procedures see below. All the labeling samples were pre-incubated in 150 mM McIlvaine buffer on ice for 5 minutes. Electrophoresis was performed with sodium dodecylsulfate containing 10% polyacrylamide gels. Wet slab-gels were then scanned for ABP-emitted fluorescence using a Bio-rad ChemiDoc MP imager using green Cy2 (λ_{EX} 470 nm, bandpass 30 nm; λ_{EM} 530 nm, bandpass 28) for **3** JJB382, red Cy3 (λ_{EX} 530 nm, bandpass 28 nm; λ_{EM} 605 nm, bandpass 50) for **4** JJB347, and blue Cy5 (λ_{EX} 625 nm, bandpass 30 nm; λ_{EM} 695 nm, bandpass 55) for **5** JJB383. All samples were denatured with 5× Laemmli buffer (50% (v/v) 1 M Tris−HCl, pH 6.8, 50% (v/v) 100% glycerol, 10% (w/v) DTT, 10% (w/v) SDS, 0.01% (w/v) bromophenol blue), boiled for 5 min at 100 °C, and separated by gel electrophoresis on 10% (w/v) SDS-PAGE gels running continuously at 90 V for 30 min and 200 V for 50min.

***In vitro* labeling of GAA**

The detection limit of each ABP was analyzed by labeling 1 pmol rGAA with 1,000−0.01 fmol **3** JJB382 in 150 mM McIlvaine buffer, pH 5.0, for 30 min at 37 °C. Influence of pH on ABP labeling involved pre-incubation of either 100 fmol rGAA or 10 µg murine liver lysate at pH 2−8 for 30 min on ice, prior to addition of 1 µM **3** JJB382, dissolved in Nanopure H_2O and incubating for 30 min at 37 °C. Assessment of **3** JJB382 labeling kinetics involved pre-cooling of 100 fmol rGAA on ice for 15 min, followed by addition of similarly cooled 100 nM **3** JJB382 solution. After mixing, **3** JJB382 labeling was chased for 0−60 min at either 4 °C or 37 °C, whereafter labeling was stopped by denaturation. For competitive ABPP on rGAA, 1 pmol rGAA was pre-incubated with inhibitors (100 µM **7** JJB307, **1** Chris021, **2** Chris022, 10 µM **4** JJB347, **5** JJB383 or **6** JJB384, 10 mM 4MU-α-D-glucopyranoside, 2.5 M maltose) for 30 min at 37 °C, or boiled for 4 min in 2% (w/v) SDS, prior to labeling with 1 µM **3** JJB382 for 30 min at 37 °C. Labeling of human fibroblast lysates was performed on 100 µg total protein, using 1 µM **3** JJB382 dissolved in either 150 mM McIlvaine buffer, pH 4.0 for labeling of GAA, or pH 7.0 for labeling of GANAB. For pH optimum labeling on lysate of C57Bl6/J mouse liver or human fibrlblast, 50 µg total protein was incubated with 1 µM **3** JJB382 in McIlvaine buffers of pH 2.0−8.0 for 30 min at

37 °C. For competitive ABPP on C57Bl6/J mouse liver, 50 µg of total protein was pre-incubated with compounds **18**-**23** with decreasing concentrations (1000 µM, 100 µM, 10 µM, 1 µM) in 150 mM McIlvaine buffer pH 4.0 and pH 7.0 for 30 min at 37 °C, prior to labeling with 1 µM **3** JJB382 for 30 min at 37 °C.

Visualizing pulled-down proteins by fluorescent detection and Western Blot.

After pull-down of mouse intestinal enzymes using **6** JJB384, the supernatants and eluents (20 µL from both) of all treatment conditions (DMSO, competition, and JJB384, under pH 4.0 or pH 7.0) were subjected to SDS-PAGE and fluorescent detection (Bio-rad ChemiDoc MP Imager; Cy2 channel). Proteins were subsequently transferred to PVDF membranes using a Trans-Blot® Turbo system (BioRad), blocked in 4% (w/v) BSA in TBST, and detected for biotin **6** JJB384 binding with Streptavidin-HRP antibody (Dako) .The blot was developed in the dark using a 10 mL luminal solution, 100 µL ECL enhancer and 3 µL 10% H_2O_2 solution. Chemiluminescence was visualized using the same ChemiDoc imager (BioRad).

Pompe GAA detection by Western blot.

Following fluorescent scanning of SDS-PAGE, proteins on wet slab gel were transferred to PVDF membrane and blocked as described in the previous section. For GAA detection, the membrane was incubated firstly with mouse polyclonal anti GAA and subsequently with goat anti mouse Alexa647 (Life Technologies). Blot was scanned on a Typhoon FLA 9500 Imager (GE Healthcare) using 633 nm laser and LPR filter, and 100 mm as pixel size. Rabbit anti β-actin (Cell Signaling) and goat anti rabbit Alexa647 (Invitrogen) were used for loading control.

***In situ* labeling of GAA**

Fibroblasts were treated with **5** JJB383 (100 nM, for 10 to 240 min) in culture medium. After washing 3 times with PBS, cells were harvested in potassium phosphate buffer (25 mM K_2HPO_4-KH_2PO_4, pH 6.5, supplemented with 0.1% (v/v) Triton X-100 and protease inhibitor cocktail (Roche)), prepared with or without 10 µM **3** JJB382. 20 µg total protein from each cell homogenate of the lysate was subjected to SDS-PAGE. Labeling were visualized by fluorescent detection using Bio-rad ChemiDoc MP Imager under the channels Cy5 for **5** JJB383 and Cy2 for **3** JJB382 with the above described settings. For Chloroquine treatment experiment, cultured fibroblasts were incubated with chloroquine (10-50 µM, for up to 3 h) and subsequently with **6** JJB383 (100 nM, for 3 h) in medium. Cells were harvested with or without 10 mM **3** JJB382 with the above describe methods. Homogenates were denatured, resolved on SDS-PAGE and detected for **5** JJB383 labeling by fluorescent scanning.

***Agd31B* expression and 3-D crystallography**

*A*gd31B expression and purification was carried out as previously described.[3] Protein crystals were obtained using 1.8 M ammonium sulfate, 0.1 M HEPES [7.0], 2% PEG 400 at 20 °C by the sitting drop vapor diffusion method. Crystal complexes with **1** CF021 and **2** CF022 were obtained by soaking in mother liquor containing 5 mM probe for 2 h, before cyroprotecting in 2.0 M lithium sulfate, 0.1 M HEPES [7.0], 2% PEG 400, and flash freezing in liquid N_2 for data collection (**Table S2**).

All data were collected at beamline I04 of the Diamond Light Source, processed using XDS[4] and reduced using Aimless.[5] Complex structures were solved by molecular replacement using MolRep,[6] before subsequent rounds of manual model building and refinement using Coot[7] and REFMAC5[8] respectively. Refinements were carried out using TLS determination of molecular motions.[9] Ligand coordinates were built using jLigand.[10] Crystal structure figures were generated using ccp4mg.[11]

Table S2. Crystal data collection and refinement statistics

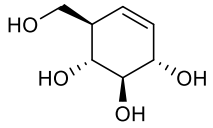
	*Cj*Agd31B-**2** CF022-complex **(PDB: 5I23)**	*Cj*Agd31B- **1** CF021-complex **(PDB: 5I24)**
Data collection		
Space group	P622	P622
Cell dimensions		
a, *b*, *c* (Å)	198.0, 198.0, 103.0	197.3, 197.3, 102.9
α, β, γ (°)	90, 90, 120	90, 90, 120
Resolution (Å)	85.74-1.95 (1.99-1.95)	49.83-1.85 (1.88-1.85)
R_{merge}	0.13 (1.64)	0.10 (1.21)
$I/\sigma I$	19.6 (2.3)	26.8 (3.0)
Completeness (%)	100 (100)	100 (100)
Redundancy	20.0 (20.1)	24.5 (25.0)
Refinement		
Resolution (Å)	85.74-1.95	49.83-1.85
No. reflections	82110	95234
R_{work}/ R_{free}	0.17/0.19	0.17/0.19
No. atoms		
Protein	6288	6264
Ligand/ion	74	70
Water	631	635
B-factors (TLS refinement)		
Protein	35.4	33.9
Ligand/ion	59.3	51.5
Water	40.6	38.8
R.m.s deviations		
Bond lengths (Å)	0.012	0.013
Bond angles (º)	1.57	1.58

Synthesis:

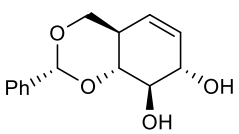
General synthesis

All reagents were of a commercial grade and were used as received unless stated otherwise. Dichloromethane (DCM), tetrahydrofuran (THF) and *N, N*-dimethylformamide (DMF) were stored over 4 Å molecular sieves, which were dried *in vacuo* before use. All reactions were performed under an argon atmosphere unless stated otherwise. Solvents used for flash column chromatography were of pro analysis quality. Reactions were monitored by TLC analysis using Merck aluminum sheets pre-coated with silica gel 60 with detection by UV absorption (254 nm) and by spraying with a solution of $(NH_4)_6Mo_7O_{24}.H_2O$ (25 g/L) and $(NH_4)_4Ce(SO_4)_4.H_2O$ (10 g/L) in 10% sulfuric acid followed by charring at ~150 °C or by spraying with an aqueous solution of $KMnO_4$ (7%) and K_2CO_3 (2%) followed by charring at ~150 °C. Column chromatography was performed using either Baker or Screening Device silica gel 60 (0.04 - 0.063 mm) in the indicated solvents. ^1H NMR and ^{13}C NMR spectra were recorded on Bruker AV-850 (850/214 MHz), Bruker DMX-600 (600/150 MHz) and Bruker AV-400 (400/100 MHz) spectrometer in the given solvent. Chemical shifts are given in ppm relative to the chloroform residual solvent peak or tetramethylsilane (TMS) as internal standard. Coupling constants are given in Hz. All given ^{13}C spectra are proton decoupled. High-resolution mass spectra were recorded with a LTQ Orbitrap (Thermo Finnigan). Optical rotations were measured on an automatic polarimeter (Sodium D-line, λ = 589 nm). LC/MS analysis was performed on an LCQ Adventage Max (Thermo Finnigan) ion-trap spectrometer (ESI^+) coupled to a Surveyor HPLC system (Thermo Finnigan) equipped with a C_{18} column (Gemini, 4.6 mm x 50 mm, 3 µm particle size, Phenomenex) equipped with buffers A: H_2O, B: acetonitrile (MeCN) and C: 1% aqueous TFA, For reversed-phase HPLC purifications an Agilent Technologies 1200 series instrument equipped with a semi preparative Gemini C_{18} column (10 x 250 mm) was used. The applied buffers were A: 50 mM NH_4HCO_3 in H_2O, B: MeCN.

(1R,2R,3S,6R)-6-(Hydroxymethyl)-cyclohex-4-ene-1,2,3-triol



The starting material diol **8** was prepared according to previously reported procedure.[12] Diol compound **8** (680 mg, 2.0 mmol, 1.0 eq.) was co-evaporated 3 times with toluene and then dissolved in dry THF (25 mL) and cooled to -60°C under argon atmosphere. Ammonia (20 mL) was condensed at -60°C under argon atmosphere. Lithium (207 mg, 30.0 mmol, 15 eq.) was added and the mixture was stirred until lithium was completely dissolved. To this solution was slowly added the solution of **8** in THF. The reaction mixture was stirred for 30 min at -60 °C and then quenched by adding of H_2O (30mL). The resulting solution was allowed to come to room temperature and stirred until all ammonia had evolved. Next, the solution was concentrated *in vacuo*, redissolved in H_2O, and neutralized with Amberlite IR-120 H^+. Then, the filtration mixture was concentrated *in vacuo* and purified via silica gel column chromatography (10%→20% MeOH in DCM), affording colorless oil product (180 mg, 1.13 mmol, 57%). TLC (DCM:MeOH, 4/1, v/v): R_f = 0.48; $[\alpha]_D^{20}$ = +105.2 (10 mg/mL in MeOH); ^1H NMR (400 MHz, CD_3OD): δ ppm 5.66 – 5.57 (m, 2H), 4.06 – 4.03 (m, 1H), 3.81 (dd, J = 10.6, 4.1 Hz, 1H), 3.66 - 3.59 (m, 1H), 3.49 - 3.41 (m, 2H), 2.31 - 2.26 (br, 1H); ^{13}C NMR (101 MHz, CD_3OD): δ ppm 130.97, 128.57, 78.81, 73.62, 72.03, 63.45, 47.65; HRMS (*m/z*): calcd. for $C_7H_{12}O_4$ [$M+H^+$], 161.08084; found: 161.08087.

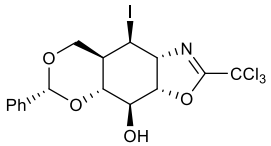


(1R,3R,6R,9S,10S)-9,10-Dihydroxy-3-phenyl-2,4-dioxabicyclo[4.4.0]dec-7-ene (9)

(1R,2R,3S,6R)-6-(Hydroxymethyl)-cyclohex-4-ene-1,2,3-triol (180 mg, 1.13

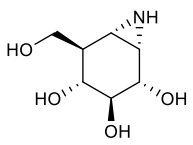
mmol, 1.0 eq.)was dissolved in dry DMF (2.0 mL) and dry MeCN (6.0 mL) in an inert atmosphere. CSA (52.2 mg, 0.25 mmol, 0.2 eq.) was added to the solution, followed by $PhCH(OMe)_2$ (253 µL, 1.69 mmol, 1.5 eq.). After 48 hours, the reaction was quenched with Et_3N (31.0 µL, 0.23 mmol, 0.2 eq.) and concentrated *in vacuo*. The reaction mixture was separated out with EtOAc and water and the aqueous layer was further extracted with EtOAc. The combined organic layers were washed with brine, dried over Na_2SO_4, filtered and concentrated *in vacuo*. The product was purified via silica gel column chromatography (30%→70% EtOAc in pentane), affording **9** (170 mg, 0.69 mmol, 61%). TLC (pentane:EtOAc, 7/3, v/v): R_f = 0.45; $[\alpha]_D^{20}$ = +30.8 (10 mg/mL in $CDCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$): δ ppm 7.52 – 7.47 (m, 2H), 7.37 – 7.31 (m, 3H), 5.57 – 5.52 (m, 2H), 5.28 – 5.25 (m, 1H), 4.24 - 4.18 (m, 2H), 3.86 (dd, *J* = 10.1, 7.5 Hz, 1H), 3.58 – 3.51 (m, 2H), 2.84 (d, *J* = 8.0 Hz,1H), 2.60 – 2.54 (m, 1H); ^{13}C NMR (101 MHz, $CDCl_3$): δ ppm 137.79, 130.49, 130.64, 129.16, 128.34, 126.39, 124.05, 102.16, 80.72, 75.42, 73.74, 69.96, 38.53; HRMS (M/Z): calcd. for $C_{14}H_{16}O_4$ [M+H$^+$], 249.11214; found: 249.11224.

(1R,6R,7R,8R,9S,10R)-10-Hydroxy-7-iodo-3-phenyl-12-trichloromethyl-13-aza-2,4,11-trioxatricyclo[7.4.4.0-8.4.3.0]tridec-12-ene (11)



A solution of **9** (535 mg, 2.15 mmol, 1 eq.) in dry DCM (60 mL) was made under argon atmosphere and cooled to 0 °C. DBU (64.4 µL, 0.431 mmol, 0.2 eq) and trichloroacetonitrile (216 µL, 10.77 mmol, 5 eq.) was added to the solution. After 3 hours, more trichloroacetonitrile (108 µL, 5.39 mmol, 2.5 eq.) was added to the reaction mixture. Starting material was fully conversed to imidate **10** with higher running spot on TLC after 21 hours stirring at room temperature. Then iodine (1.70 g, 6.68 mmol, 3.1 eq), $NaHCO_3$ (1.81 g, 21.5 mmol, 10 eq.) and water (3.8 mL) was added to the reaction mixture. After 26 hours of the first addition of the iodine and $NaHCO_3$, iodine (1.70 g, 6.68 mmol, 3.1 eq.) and $NaHCO_3$ (1.81 g, 21.5 mmol, 10 eq.) was added to the reaction mixture. After approximately 96 hours after the first addition of iodine and $NaHCO_3$, the reaction was quenched by adding $Na_2S_2O_3$ (10% aqueous solution) until the purple color had disappeared. The reaction mixture was separated out and the DCM layer was concentrated *in vacuo*, redissolved in EtOAc and washed with water and brine. The initial aqueous layer was extracted with EtOAc and the combined organic layers were dried ($MgSO_4$), filtered and concentrated *in vacuo*. The product was purified via silica gel column chromatography (0%→16% EtOAc in pentane) affording **11** (455 mg, 0.877 mmol, 41%) was produced. TLC (PE:EtOAc, 4/1, v/v): R_f = 0.57; ^{1}H NMR (400 MHz, $CDCl_3$): δ ppm 7.49 – 7.44 (m, 2H), 7.49 – 7.33 (m, 3H), 5.62 (s, 1H), 5.22 (t, *J* = 7.5 Hz, 1H), 4.85 – 4.83 (m, 1H), 4.72 (s, 1H), 4.21 (dd, *J* = 11.3, 4.7 Hz, 1H), 4.03 (t, *J* = 9.8 Hz, 1H), 3.92 (t, *J* = 10.6 Hz, 1H), 3.83 – 3.79 (m, 1H), 3.08 (s, 1H), 1.20 – 1.13 (m, 1H); ^{13}C NMR (101 MHz, $CDCl_3$): δ ppm 163.72, 137.29, 129.48, 128.51, 126.26, 101.64, 86.69, 77.77, 75.19, 74.93, 72.57, 34.90, 23.69; HRMS (*m/z*): calcd. for $C_{16}H_{15}Cl_3NO_4$ [M+H$^+$], 519.91567; found: 519.91504.

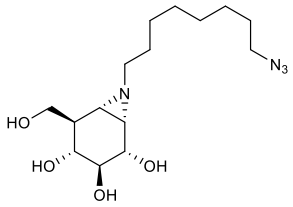
(1S,2R,3S,4R,5R,6R)-5-(Hydroxymethyl)-7-aza-bicyclo-[4.1.0]heptane-2,3,4-triol (1)



A solution of **11** (455 mg, 0.877 mmol, 1.0 eq.) was dissolved in 1, 4-dioxane (9.2 mL) and heated to 60 °C. HCl aqueous (37%, 2.60 mL) was added to the solution. The reaction mixture was left stirring at 60 °C overnight. The reaction mixture was concentrated *in vacuo* and then separated out with EtOAc and water. The aqueous layer was washed with EtOAc and concentrated *in vacuo* and co-evaporated with toluene. The crude product of the free amine intermediate **12** was dissolved in MeOH (30 mL) and $NaHCO_3$ (2.95 g, 35.1 mmol, 40eq.) was added to the solution. The reaction mixture was left stirring at room temperature for 4 days. The reaction mixture filtered and then concentrated *in vacuo*. The

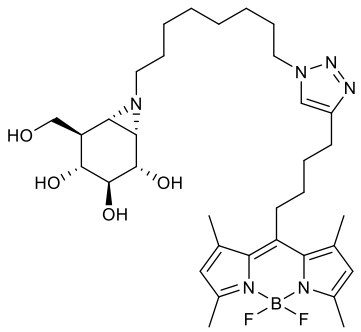
residue was redissolved in water and filtered over a pad of amberlite IR-120 H^+ resin, washed with H_2O and followed by 1M NH_4OH aqueous. The filtrate was concentrated *in vacuo*, affording the aziridine product of **1** (97.3 mg, 0.556 mmol, 63%) as light brown oil. ^{1}H NMR (400 MHz, D_2O): δ ppm 4.11 – 4.04 (m, 2H), 3.97 – 3.90 (m, 1H), 3.55 – 3.39 (m, 2H), 2.82 – 2.80 (m, 1H), 2.57 (d, *J* = 6.4 Hz, 1H), 2.11 – 2.01 (m, 1H). ^{13}C NMR (101 MHz, D_2O): δ ppm 73.46, 71.27, 70.07, 61.50, 44.37, 35.66, 31.69; HRMS (*m*/*z*): calcd. for $C_7H_{13}NO4$ [M+H$^+$], 176.09173; found: 176.09163.

(1S,2R,3S,4R,5R,6R)-2,3,4-trihyrdoxy-5-(hydroxymethyl)-7-(8-azidooctyl)-7-aza-bicyclo[4.1.0]heptan e (2)



A solution of **1** (97.3 mg, 0.556 mmol, 1.0 eq.) in dry DMF (2.2 mL) was made under argon atmosphere and heated to 80 °C. 1-azido-8-iodooctane (234 mg, 0.833 mmol, 1.5 eq) and K_2CO_3 (345.5 mg, 2.50 mmol, 4.5 eq) were added to the solution. After 21 hours stirring, the reaction was quenched with MeOH (0.15 mL) and filtered over a pad of celite. The filtrate was concentrated *in vacuo*. The product was purified via silica gel column chromatography (5%→18% MeOH in DCM), affording a light yellow oil **2** CF022 (72 mg, 0.219 mmol, 39%). TLC (DCM:MeOH, 3/1, v/v): R_f = 0.39; $[\alpha]_D^{21}$ = +25.4 (10 mg/mL in MeOH); ^{1}H NMR (850 MHz, CD_3OD): δ ppm 3.87 (dd, *J* = 10.7, 4.0 Hz, 1H), 3.66 (dd, *J* = 8.6, 3.7 Hz, 1H), 3.63 (dd, *J* = 10.8, 7.1 Hz, 1H), 3.34 – 3.31 (m, 1H), 3.29 (t, *J* = 6.8 Hz, 2H), 3.05 (t, *J* = 10.0 Hz, 1H), 2.36 – 2.33 (m, 1H), 2.17 – 2.14 (m, 1H), 1.86 – 1.82 (m, 2H), 1.68 (d, *J* = 6.5 Hz, 1H), 1.60 – 1.57 (m, 4H), 1.40 – 1.32 (m, 8H); ^{13}C NMR (214 MHz, CD_3OD): δ ppm 75.76, 73.37, 72.45, 63.52, 62.25, 52.46, 46.89, 46.00, 41.95, 30.58, 30.48, 30.19, 29.90, 28.35, 27.77; LC/MS (*m*/*z*): t_R=4.42 min (linear gradient 10%→90% B in 12.5 min), 329.20 [M+H]$^+$; HRMS (*m*/*z*): calcd. for $C_{15}H_{28}N_4O_5$ [M+H$^+$], 329.21833; found: 329.21828.

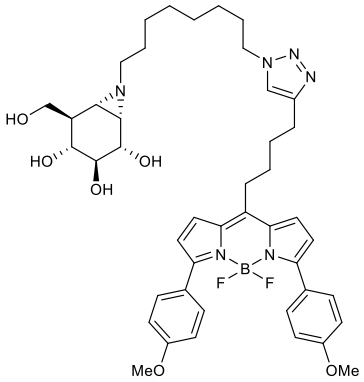
(1S,2S,3S,4R,5R,6S)-7-(8-(4-(4-(5,5-difluoro-1,3,7,9-tetramethyl-5H-4l4,5l4-dipyrrolo[1,2-c:2',1'-f][1,3 ,2]diazaborinin-10-yl)butyl)-1H-1,2,3-triazol-1-yl)octyl)-5-(hydroxymethyl)-7-azabicyclo[4.1.0]heptan e-2,3,4-triol (3)



A solution of azide compound **2** (17 mg, 0.0518 mmol, 1.0eq.) in DMF (1.5 mL) was made under argon atmosphere, green BODIPY compound **13** (19 mg, 0.0579 mmol, 1.1 eq), aqueous solutions of $CuSO_4$ (1M, 10.4 µL, 0.0104 mmol, 0.20 eq.) and sodium ascorbate (1M, 10.8 µL, 0.0108 mmol, 0.21 eq) were added to the solution and the mixture was stirred at room temperature for 12h. The resulting mixture was concentrated under reduced pressure and the crude product was purified by semi-preparative reversed HPLC (linear gradient: 45%→53% B in A, 12 min, solutions used A: 50 mM NH_4HCO_3 in H_2O, B: MeCN) and lyophilized resulting as orange cottony product **3** (13.2 mg, 20.2 µmol, 39%). ^{1}H NMR (850 MHz, CD_3OD) δ ppm 7.73 (s, 1H), 6.11 (s, 2H), 4.34 (t, *J* = 7.0 Hz, 2H), 3.87 (dd, *J* = 8.7, 3.9 Hz, 1H), 3.65 (dd, *J* = 8.5, 3.7 Hz, 1H), 3.61 (dd, *J* = 10.7, 7.2 Hz, 1H), 3.31 – 3.30 (m, 1H), 3.05 (t, *J* = 10.0 Hz, 1H), 3.02 – 2.97 (m, 2H), 2.77 (t, *J* = 7.3 Hz, 2H), 2.43 (s, 6H), 2.37 (s, 6H), 2.34 – 2.26 (m, 1H), 2.15 – 2.11 (m, 1H), 1.91 – 1.81 (m, 6H), 1.67 – 1.62 (m, 3H), 1.57 – 1.53 (m, 2H), 1.36 – 1.22 (m, 8H).; ^{13}C NMR (214 MHz, CD_3OD) δ ppm 154.90, 148.49, 147.88, 142.18, 132.57, 123.36, 122.60, 75.74, 73.34, 72.41, 63.50, 62.19, 51.20, 46.86, 45.99, 41.93, 32.23, 31.24, 30.83, 30.42, 29.88, 29.05, 28.22, 27.32, 25.87, 16.48, 14.44; LC/MS (*m*/*z*): t_R=5.88 min (linear gradient 10%→90% B in

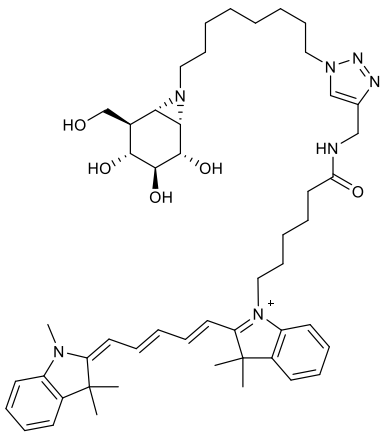
12.5 min), 657.33 $[M+H]^+$, 679.53 $[M+Na]^+$; HRMS (*m/z*): calcd. for $C_{34}H_{51}BF_2N_6O_4$ $[M+H^+]$, 657.41117; found: 657.41033.

(1S,2S,3S,4R,5R,6S)-7-(8-(4-(4-(5,5-difluoro-3,7-bis(4-methoxyphenyl)-5H-4λ⁴,5λ⁴-dipyrrolo[1,2-c:2',1'-f][1,3,2]diazaborinin-10-yl)butyl)-1H-1,2,3-triazol-1-yl)octyl)-5-(hydroxymethyl)-7-azabicyclo[4.1.0]heptane-2,3,4-triol (4)



A solution of azide compound **2** (18 mg, 0.0548 mmol, 1.0eq.) in DMF (1.5 mL) was made under argon atmosphere, red BODIPY compound **14** (29.2 mg, 0.0603 mmol, 1.1 eq), aqueous solutions of $CuSO_4$ (1M, 11.0 µL, 0.0110 mmol, 0.20 eq.) and sodium ascorbate (1M, 11.5 µL, 0.0115 mmol, 0.21 eq) were added to the solution and the mixture was stirred at room temperature for 14h. The resulting mixture was concentrated under reduced pressure and the crude product was purified by semi-preparative reversed HPLC (linear gradient: 51%→57% B in A, 12 min, solutions used A: 50 mM NH_4HCO_3 in H_2O, B: MeCN) and lyophilized resulting as purple cottony product **4** (5.13 mg, 6.32 µmol, 12%). ^{1}H NMR (850 MHz, CD_3OD) δ ppm 7.88 – 7.79 (m, 4H), 7.68 (s, 2H), 7.43 (d, *J* = 4.3 Hz, 3H), 7.02 – 6.94 (m, 4H), 6.69 (d, *J* = 4.3 Hz, 2H), 4.32 (t, *J* = 7.1 Hz, 2H), 3.86 – 3.84(m, 7H), 3.64 (dd, *J* = 8.5, 3.7 Hz, 1H), 3.60 (dd, *J* = 10.7, 7.1 Hz, 1H), 3.32 – 3.31 (m, 1H), 3.07 – 3.03 (m, 3H), 2.78 (t, *J* = 6.9 Hz, 2H), 2.29 – 2.26 (m, 1H), 2.12 – 2.09 (m, 1H), 1.88 – 1.79 (m, 6H), 1.64 (d, *J* = 6.5 Hz, 1H), 1.55 – 1.49 (m, 2H), 1.30 – 1.20 (m, 8H); ^{13}C NMR (214 MHz, CD_3OD) δ ppm 162.20, 158.79, 148.60, 146.78, 137.49, 132.16, 132.14, 132.12, 128.42, 126.51, 123.24, 121.03, 114.85, 114.63, 75.75, 73.36, 72.46, 63.53, 62.18, 55.83, 51.24, 46.86, 45.97, 41.92, 34.13, 31.23, 30.98, 30.40, 30.33, 29.87, 28.20, 27.33, 25.78; LC/MS (*m/z*): t_R=6.82 min (linear gradient 10%→90% B in 12.5 min), 813.33 $[M+H]^+$; HRMS (*m/z*): calcd. for $C_{44}H_{55}BF_2N_6O_6$ $[M+H^+]$, 813.43245; found: 813.43137.

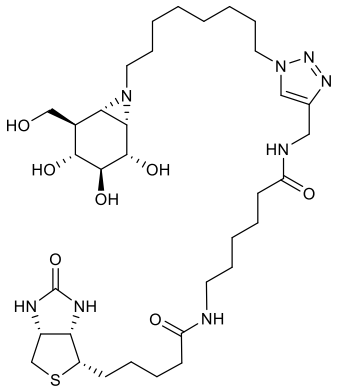
3,3-dimethyl-1-(6-oxo-6-(((1-(8-((1S,2S,3S,4R,5R,6S)-2,3,4-trihydroxy-5-(hydroxymethyl)-7-azabicyclo[4.1.0]heptan-7-yl)octyl)-1H-1,2,3-triazol-4-yl)methyl)amino)hexyl)-2-((1E,3E)-5-((E)-1,3,3-trimethylindolin-2-ylidene)penta-1,3-dien-1-yl)-3H-indol-1-ium (5)



A solution of azide compound **2** (17 mg, 0.0518 mmol, 1.0 eq.) in DMF (1.5 mL) was made under argon atmosphere, Cy5 compound **15** (29.7 mg, 0.0570 mmol, 1.1 eq), aqueous solutions of $CuSO_4$ (1M, 10.4 µL, 0.0104 mmol, 0.20 eq.) and sodium ascorbate (1M, 10.8 µL, 0.0108 mmol, 0.21 eq) were added to the solution and the mixture was stirred at room temperature for 12h. The resulting mixture was concentrated under reduced pressure and the crude product was purified by semi-preparative reversed HPLC (linear gradient: 40%→70% B in A, 12 min, solutions used A: 50 mM NH_4HCO_3 in H_2O, B: MeCN) and lyophilized resulting as dark blue powder product **5** (10.4 mg, 11.7 µmol, 23%). ^{1}H NMR (600 MHz, CD_3OD): δ ppm 8.32 – 8.16 (m, 2H), 7.84 (s, 1H), 7.55 – 7.48 (m, 2H), 7.41 (m, 2H), 7.35 – 7.21 (m, 4H), 6.62 (t, *J* = 12.4 Hz, 1H), 6.29 – 6.27 (m, 2H), 4.41 (s, 2H), 4.36 (t, *J* = 7.1 Hz, 2H), 4.09 (t, *J* = 7.5 Hz, 2H), 3.86 (dd, *J* = 10.7, 3.9 Hz, 1H), 3.69 – 3.57 (m, 5H), 3.33 – 3.30 (m, 1H), 3.10 – 3.00 (m, 1H), 2.34 – 2.29 (m, 1H), 2.25 (t, *J*

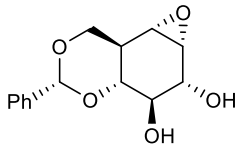
= 7.3 Hz, 2H), 2.15 – 2.11 (m, 1H), 1.93 (s, 6H), 1.90 – 1.77 (m, 6H), 1.73 – 1.57 (m, 14H), 1.60 – 1.51 (m, 2H), 1.50 – 1.42 (m, 2H), 1.33 – 1.27 (m, 8H); ^{13}C NMR (151 MHz, CD_3OD): δ ppm 178.47, 175.73, 175.38, 174.59, 155.54, 155.47, 146.13, 144.23, 143.53, 142.61, 142.51, 129.77, 129.73, 126.60, 126.27, 126.21, 124.15, 123.41, 123.28, 112.01, 111.84, 104.41, 104.23, 75.72, 73.31, 72.39, 63.48, 62.17, 51.33, 50.53, 50.50, 46.87, 46.00, 44.74, 41.92, 36.46, 35.56, 31.50, 31.30, 30.44, 30.41, 29.94, 28.25, 28.12, 27.93, 27.78, 27.37, 27.30, 26.39; LC/MS (*m/z*): t_R=5.18 min (linear gradient 10%→90% B in 12.5 min), 848.60 $[M]^+$; HRMS (*m/z*): calcd. for $C_{50}H_{70}N_7O_5^+$ [M+H$^+$], 849.54653; found: 849.55112.

6-(5-((3aS,4S,6aR)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamido)-N-((1-(8-((1S,2S,3S, 4R,5R,6S)-2,3,4-trihydroxy-5-(hydroxymethyl)-7-azabicyclo[4.1.0]heptan-7-yl)octyl)-1H-1,2,3-triazol-4-yl)methyl)hexanamide (6)



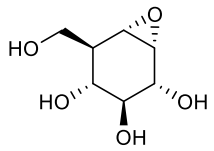
A solution of azide compound **2** (15 mg, 0.0457 mmol, 1.0 eq.) in DMF (1.5 mL) was made under argon atmosphere, biotin compound **16** (19.8 mg, 0.0502 mmol, 1.1 eq), aqueous solutions of $CuSO_4$ (1M, 9.14 µL, 0.00914 mmol, 0.20 eq.) and sodium ascorbate (1M, 9.60 µL, 0.00960 mmol, 0.21 eq) were added to the solution and the mixture was stirred at room temperature for 12h. The resulting mixture was concentrated under reduced pressure and the crude product was purified by semi-preparative reversed HPLC (linear gradient: 19%→28% B in A, 12 min, solutions used A: 50 mM NH_4HCO_3 in H_2O, B: MeCN) and lyophilized resulting as white cottony product **6** (7.60 mg, 10.5 µmol, 23%). ^{1}H NMR (400 MHz, CD_3OD) δ ppm 7.85 (s, 1H), 4.49 (dd, *J* = 7.8, 4.8 Hz, 1H), 4.44 – 4.34 (m, 4H), 4.30 (dd, *J* = 7.9, 4.5 Hz, 1H), 3.86 (dd, *J* = 10.7, 3.9 Hz, 1H), 3.69 – 3.58 (m, 2H), 3.34 – 3.30 (m, 1H), 3.25 – 3.10 (m, 3H), 3.05 (t, *J* = 10.0 Hz, 1H), 2.93 (dd, *J* = 12.7, 5.0 Hz, 1H), 2.74 – 2.64 (m, 1H), 2.37 – 2.29 (m, 1H), 2.26 – 2.10 (m, 5H), 1.95 – 1.79 (m, 4H), 1.79 – 1.40 (m, 14H), 1.37 – 1.30 (m, 10H); ^{13}C NMR (101 MHz, CD_3OD) δ ppm 176.01, 166.11, 161.52, 146.22, 124.16, 75.74, 73.34, 72.41, 63.50, 63.38, 62.20, 61.62, 57.02, 51.35, 46.89, 46.01, 41.95, 41.05, 40.18, 36.81, 36.75, 35.58, 31.29, 30.45, 30.12, 29.95, 29.78, 29.50, 28.26, 27.54, 27.37, 26.93, 26.52; LC/MS (*m/z*): t_R=4.37 min (linear gradient 10%→90% B in 12.5 min), 723.47 $[M+H]^+$; HRMS (*m/z*): calcd. for $C_{34}H_{58}N_8O_7S$ [M+H$^+$], 723.42245; found: 723.42219.

(1aR,2R,3R,3aR,7aS,7bS)-5-phenylhexahydro-2H-oxireno[2',3':3,4]benzo[1,2-d][1,3]dioxine-2,3-diol (17)



A solution of compound **9** (100 mg, 0.40 mmol, 1 eq.) and 3-Chloroperbenzoic acid (mCPBA, 75%) (135 mg, 0.60 mmol, 1.5 eq) in dry DCM (10 mL) were made under argon atmosphere and refluxed at 40°C for 24h. The reaction mixture was washed by saturated $NaHCO_3$ aqueous and brine. The organic layer was dried over Na_2SO_4, filtered and concentrated *in vacuo*. The product was purified via silica gel column chromatography (50%→100% EtOAc in pentane), affording the title compound **17** (47 mg, 0.178 mmol, 44%). TLC (pentane:EtOAc, 3/1, v/v): R_f = 0.30; ^{1}H NMR (400 MHz, $CDCl_3$): δ ppm 7.49 – 7.46 (m, 2H), 7.37 – 7.33 (m, 3H), 5.48 (s, 1H), 4.38 (dd, *J* = 12.0, 4.0 Hz, 1H), 3.97 – 3.95 (m, 1H), 3.74 – 3.68 (m, 1H), 3.64 – 3.60 (m, 1H), 3.32 – 3.25 (m, 2H), 2.78 (d, *J* = 4.0 Hz, 1H), 2.22 – 2.15 (m, 1H); ^{13}C NMR (101 MHz, $CDCl_3$): δ ppm 137.46, 129.38, 128.45, 126.44, 102.14, 79.79, 73.32, 72.76, 68.37, 55.68, 51.71, 38.00; HRMS (*m/z*): calcd. for $C_{14}H_{16}O_5$ [M+H$^+$], 265.10705; found: 265.10720.

(1R,2R,3S,4R,5R,6S)-5-(hydroxymethyl)-7-oxabicyclo[4.1.0]heptane-2,3,4-triol (7)

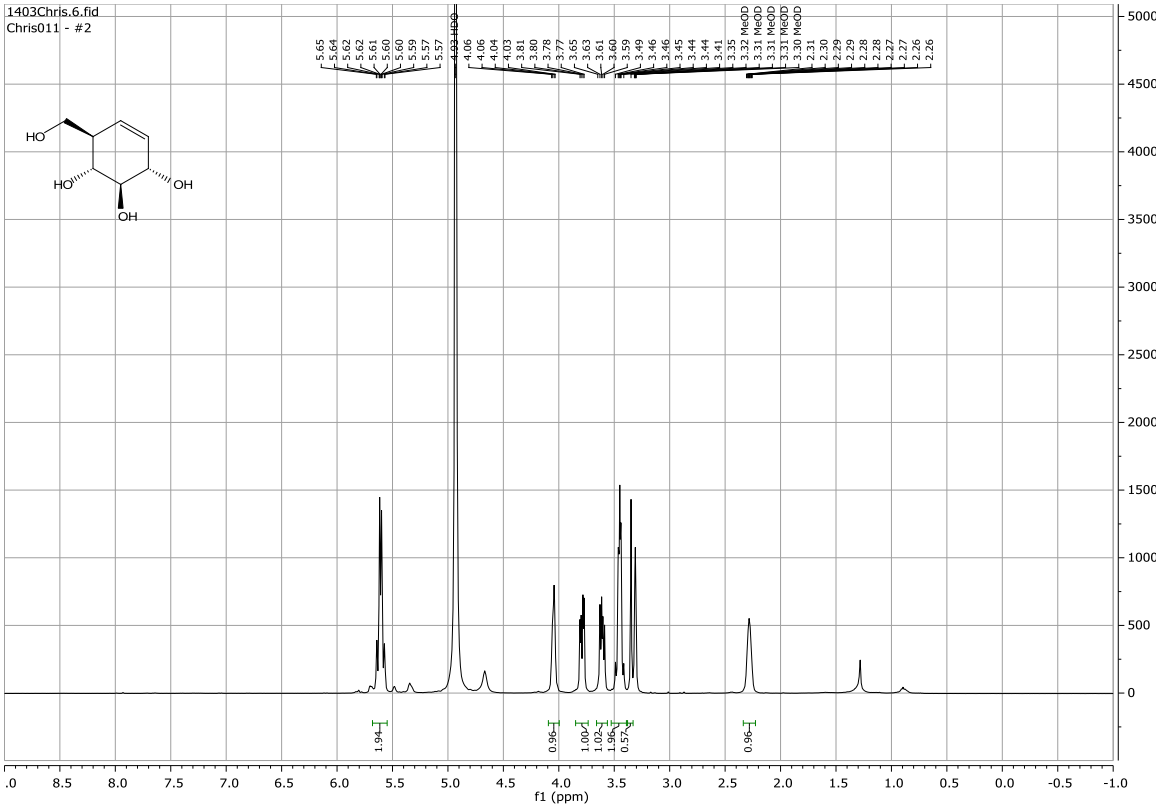


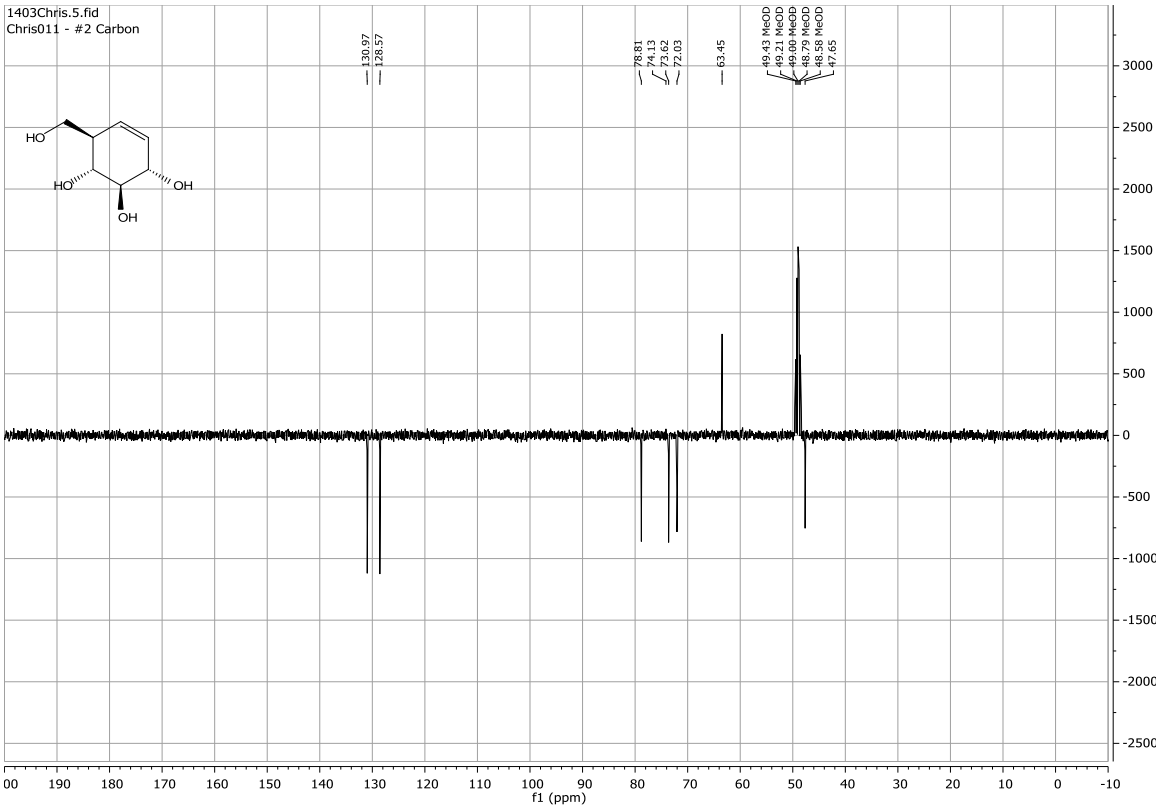
A mixture of product **17** (20 mg, 0.076 mmol, 1.0 eq.) and 15 mg $Pd(OH)_2/C$ (20 wt. % loading(dry basis)) in MeOH (2.0 mL) was stirred at room temperature under hydrogen atmosphere overnight. The catalyst was then filtered off and washed with MeOH. The filtrate and washings were combined and concentrated under reduced pressure. Crude product was purified by semi preparative reversed phase HPLC (linear gradient: 0%-20%, 3 CV, solutions used A: 50mM NH_4HCO_3 in H_2O, B: Acetonitrile) and lyophilization gave the title compound *epi*-cyclophellitol **7** (9.0 mg, 0.051 mmol, 68%) as white solid. TLC (MeOH:$CDCl_3$, 1/3, v/v): R_f = 0.26; ^{1}H NMR (400 MHz, D_2O): δ ppm 3.91 – 3.86 (m, 2H), 3.78 – 3.74 (m, 1H), 3.44 – 3.43 (m, 1H), 3.41 - 3.37 (m, 1H), 3.34 – 3.29 (m, 2H), 2.04 – 1.99 (m, 1H); ^{13}C NMR (101 MHz, D_2O): δ ppm 73.05, 71.23, 69.35, 60.32, 57.47, 55.09, 44.12; HRMS (*m/z*): calcd. for $C_7H_{12}O_5$ [$M+H^+$], 177.07575; found: 177.07571.

References

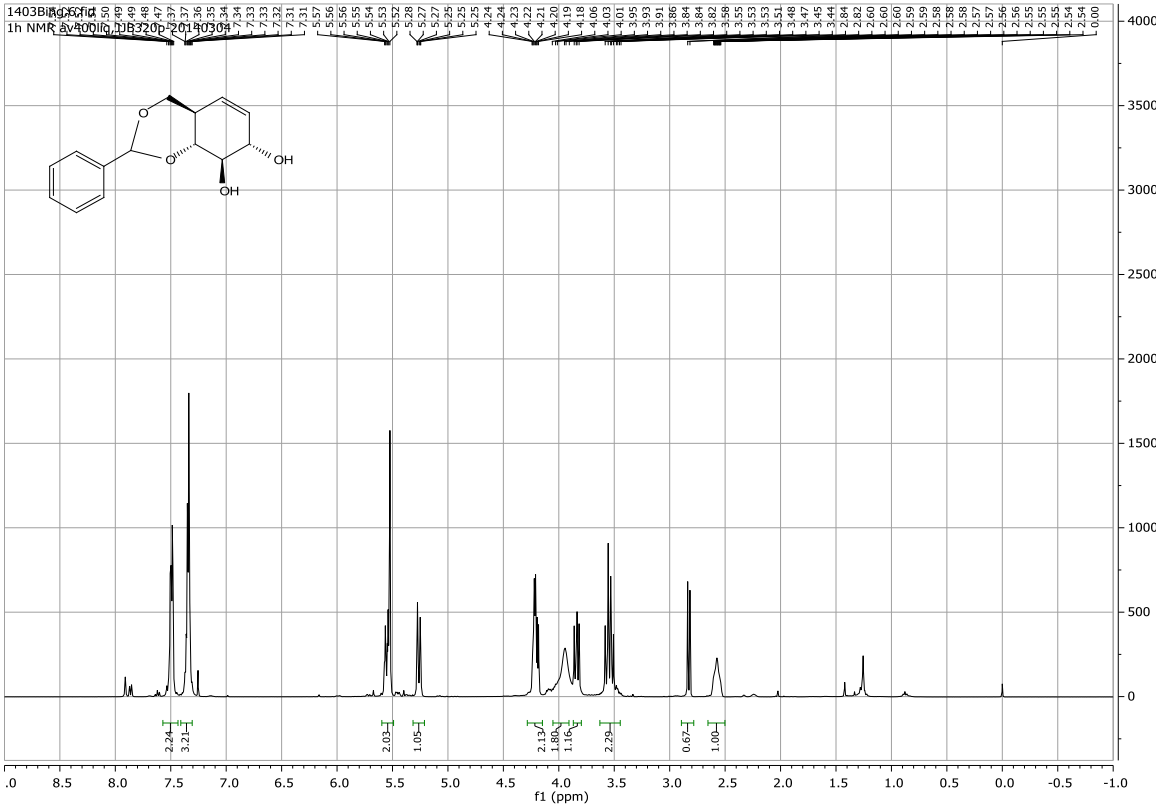
1. Witte, M. D.; Kallemeijn, W. W.; Aten, J.; Li, K.-Y.; Strijland, A.; Donker-Koopman, W. E.; van den Nieuwendijk, A. M. C. H.; Bleijlevens, B.; Kramer, G.; Florea, B. I.; Hooibrink, B.; Hollak, C. E. M.; Ottenhoff, R.; Boot, R. G.; van der Marel, G. A.; Overkleeft, H. S.; Aerts, J. M. F. G. Ultrasensitive in situ visualization of active glucocerebrosidase molecules. *Nat. Chem. Biol.* **2010**, *6*, 907-913.
2. Li, N.; Kuo, C.-L.; Paniagua, G.; van den Elst, H.; Verdoes, M.; Willems, L. I.; van der Linden, W. A.; Ruben, M.; van Genderen, E.; Gubbens, J.; van Wezel, G. P.; Overkleeft, H. S.; Florea, B. I. Relative quantification of proteasome activity by activity-based protein profiling and LC-MS/MS. *Nat. Protocols* **2013**, *8*, 1155-1168.
3. Larsbrink, J.; Izumi, A.; Hemsworth, G. R.; Davies, G. J.; Brumer, H. Structural enzymology of Cellvibrio japonicus Agd31B protein reveals alpha-transglucosylase activity in glycoside hydrolase family 31. *J. Biol. Chem.* **2012**, *287*, 43288-43299.
4. Kabsch, W. XDS. *Acta Crystallogr. D* **2010**, *66*, 125-132.
5. Evans, P.R.; Murshudov, G.N. How good are my data and what is the resolution? *Acta Crystallogr. D* **2013**, *69*, 1204-1214.
6. Vagin, A.; Teplyakov, A. MOLREP: an automated program for molecular replacement. *J Appl. Crystallogr.* **1997**, *30*, 1022-1025.
7. Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D* **2004**, *60*, 2126-2132.
8. Murshudov, G.N.; Vagin, A.A.; Dodson, E.J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **1997**, *53*, 240-255.
9. Painter, J.; Merritt, E.A. Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr. D* **2006**, *62*, 439-450.
10. Lebedev, A. A.; Young, P.; Isupov, M. N.; Moroz, O. V.; Vagin, A. A.; Murshudov, G. N. JLigand: a graphical tool for the CCP4 template-restraint library. *Acta Crystallogr. D* **2012,** *68*, 431-440.
11. McNicholas, S.; Potterton, E.; Wilson, K.S.; Noble, M.E.M. Presenting your structures: the CCP4mg molecular-graphics software. *Acta Crystallogr. D* **2011**, *67*, 386-394.
12. Li, K.-Y.; Jiang, J.; Witte, M. D.; Kallemeijn, W. W.; van den Elst, H.; Wong, C.-S.; Chander, S. D.; Hoogendoorn, S.; Beenakker, T. J. M.; Codée, J. D. C.; Aerts, J. M. F. G.; van der Marel, G. A.; Overkleeft, H. S. Synthesis of Cyclophellitol, Cyclophellitol Aziridine, and Their Tagged Derivatives. *Eur. J. Org. Chem.* **2014***, 2014*, 6030-6043.

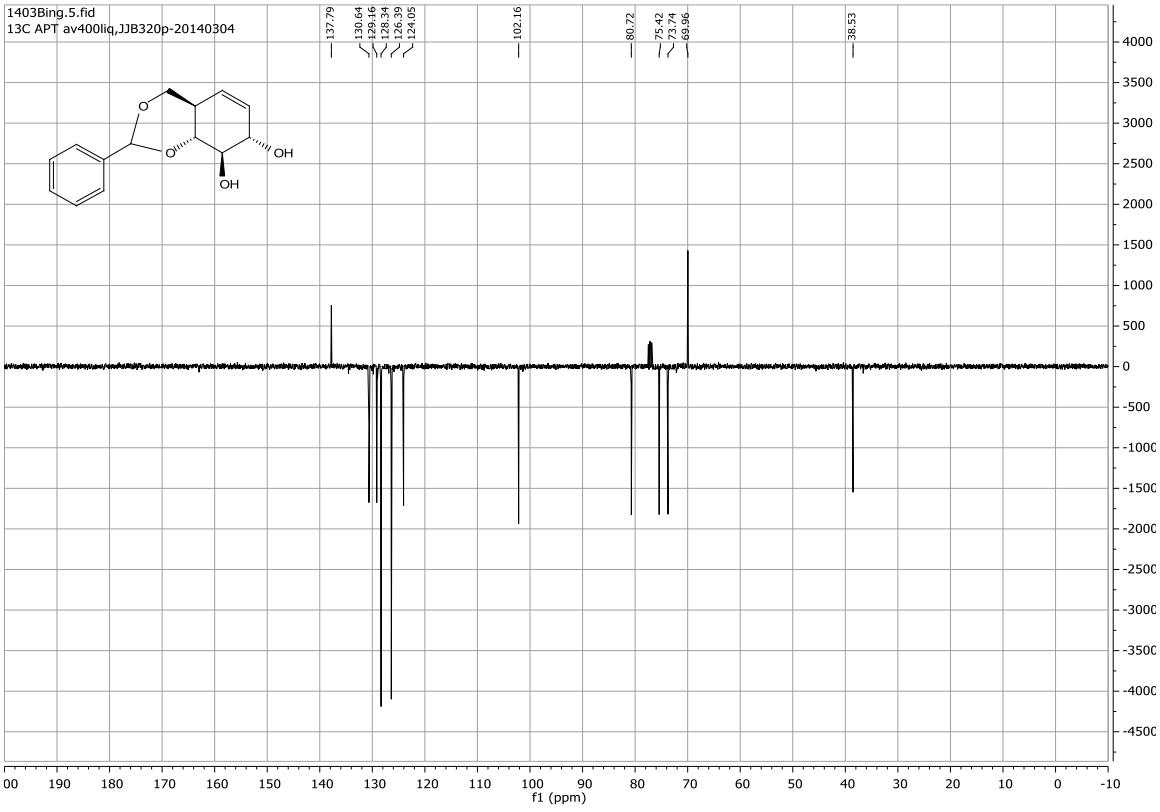
^{1}H and ^{13}C NMR spectra of compound **(1R,2R,3S,6R)-6-(Hydroxymethyl)-cyclohex-4-ene-1,2,3-triol**



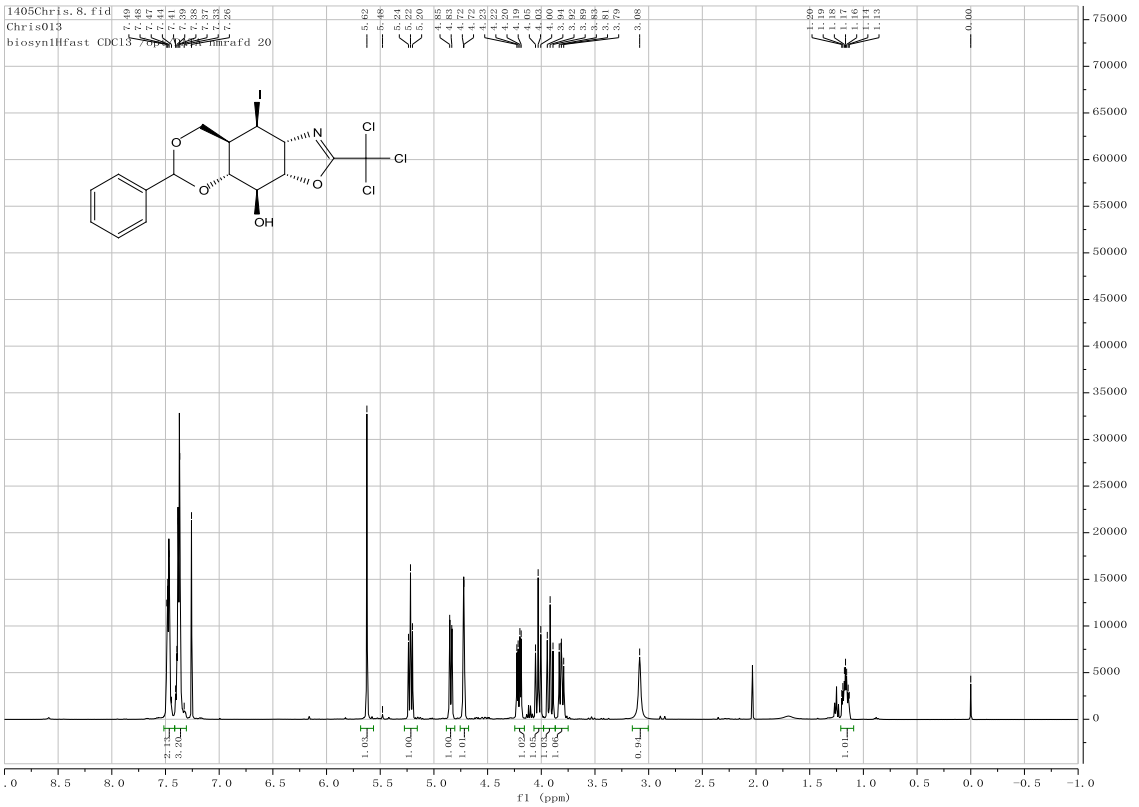


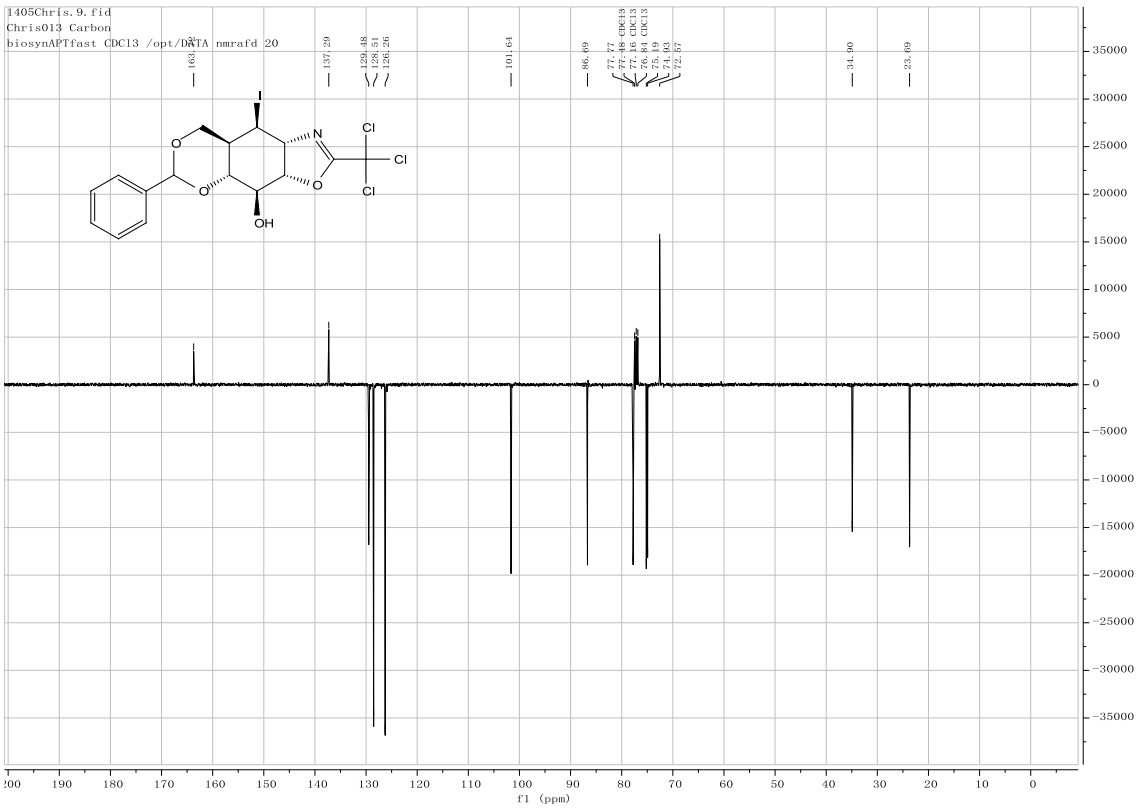
^{1}H and ^{13}C NMR spectra of compound **9**



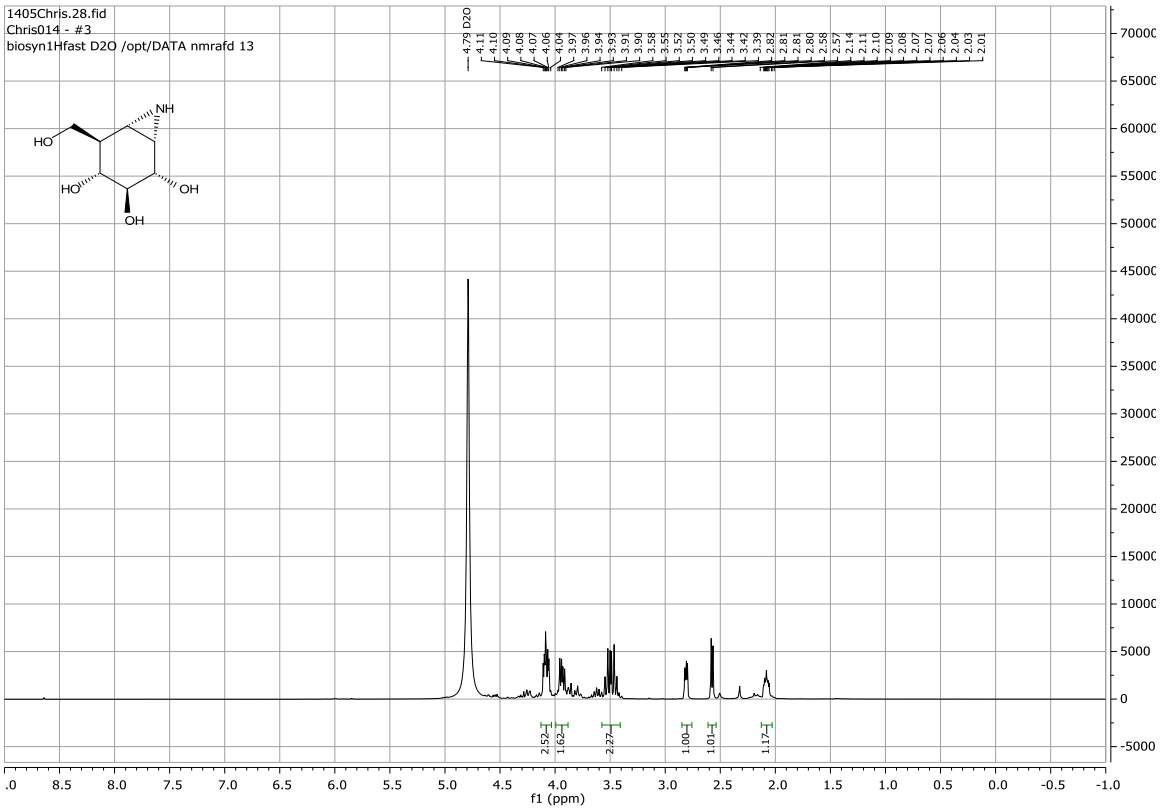


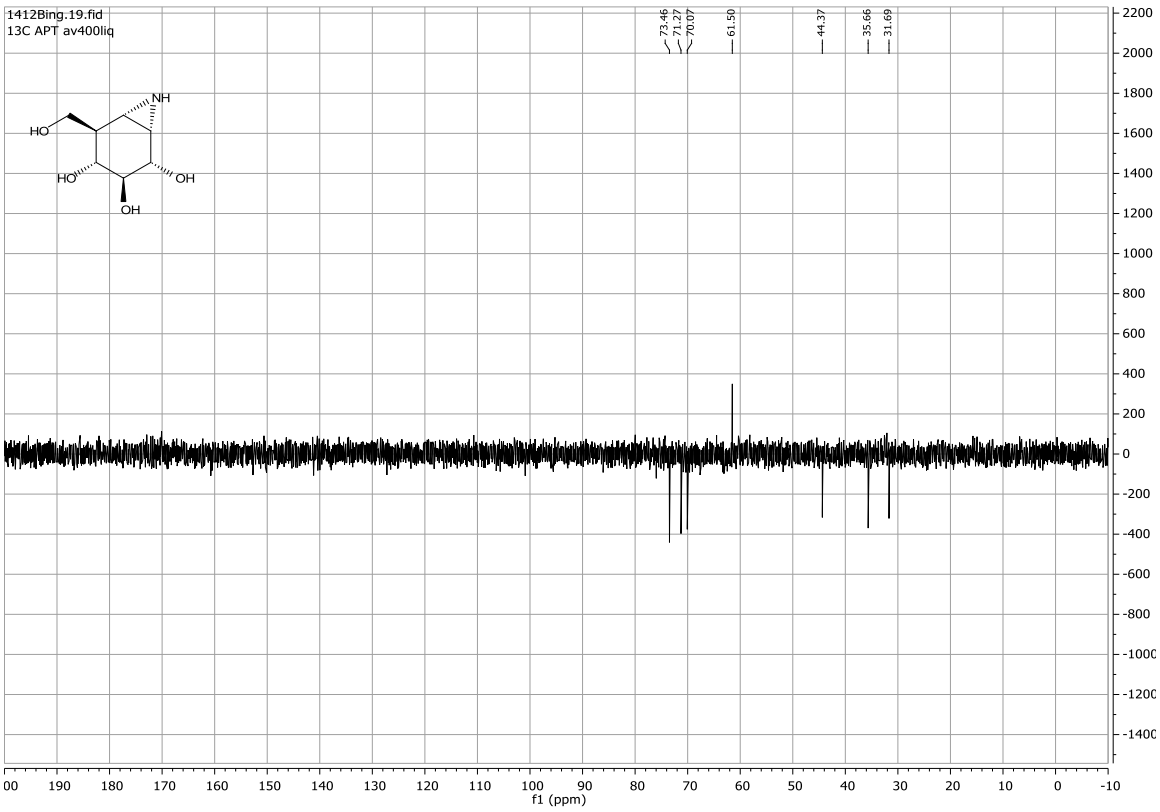
^{1}H and ^{13}C NMR spectra of compound **11**



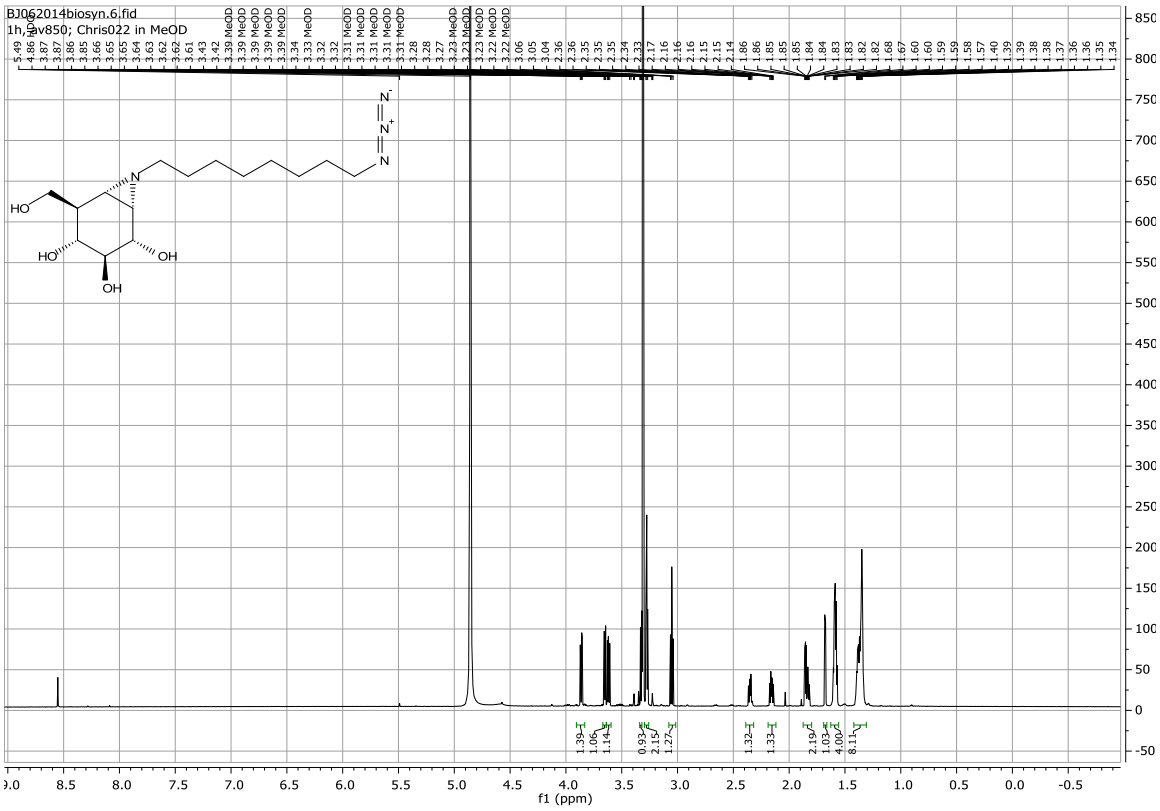


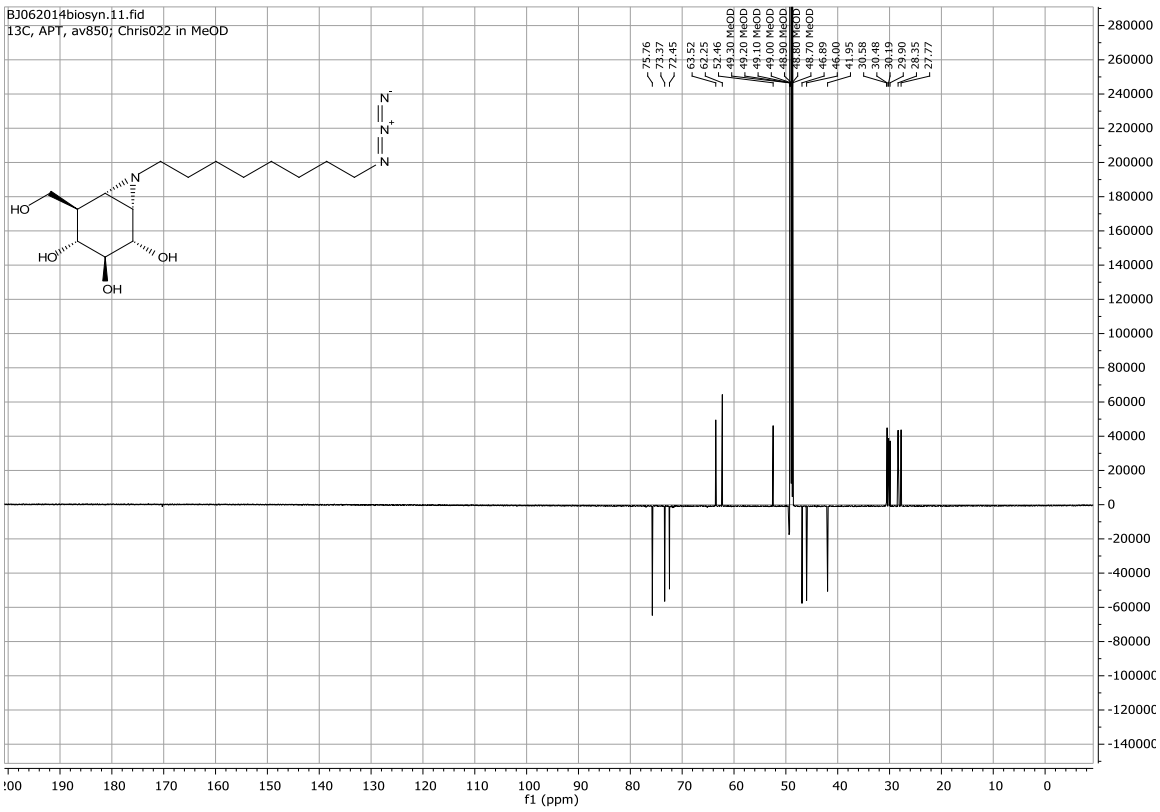
^{1}H and ^{13}C NMR spectra of compound **1**



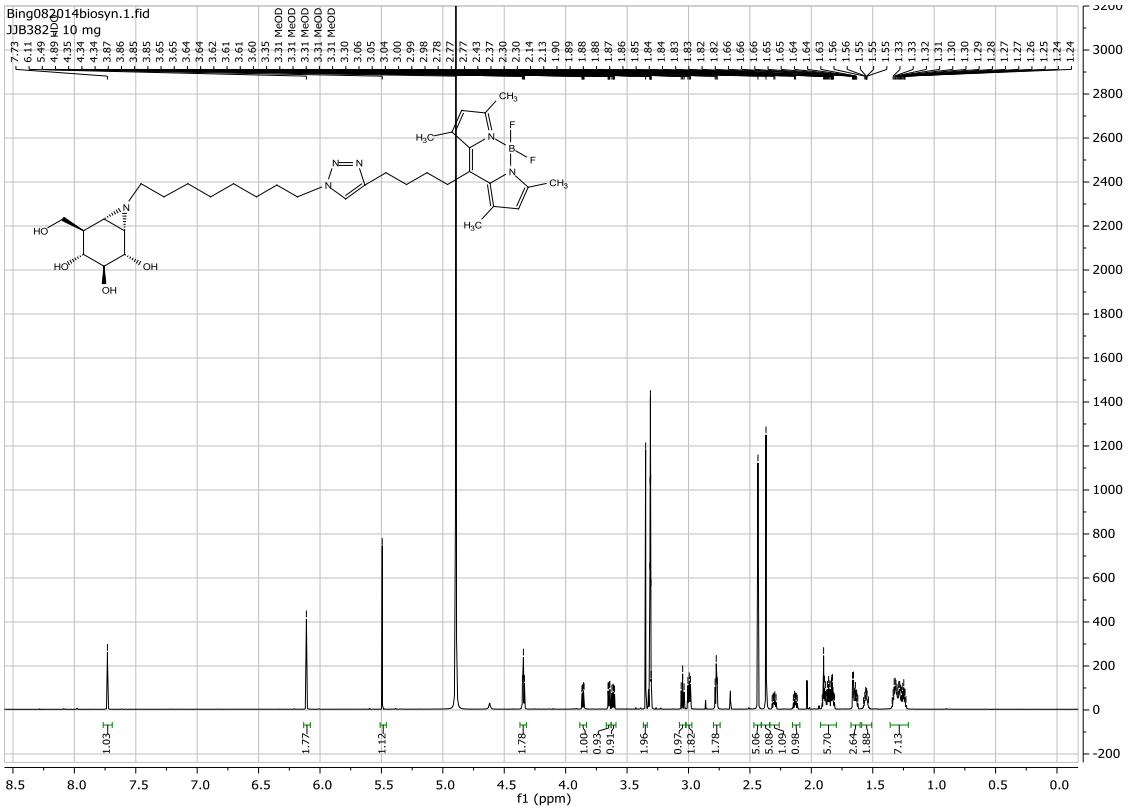


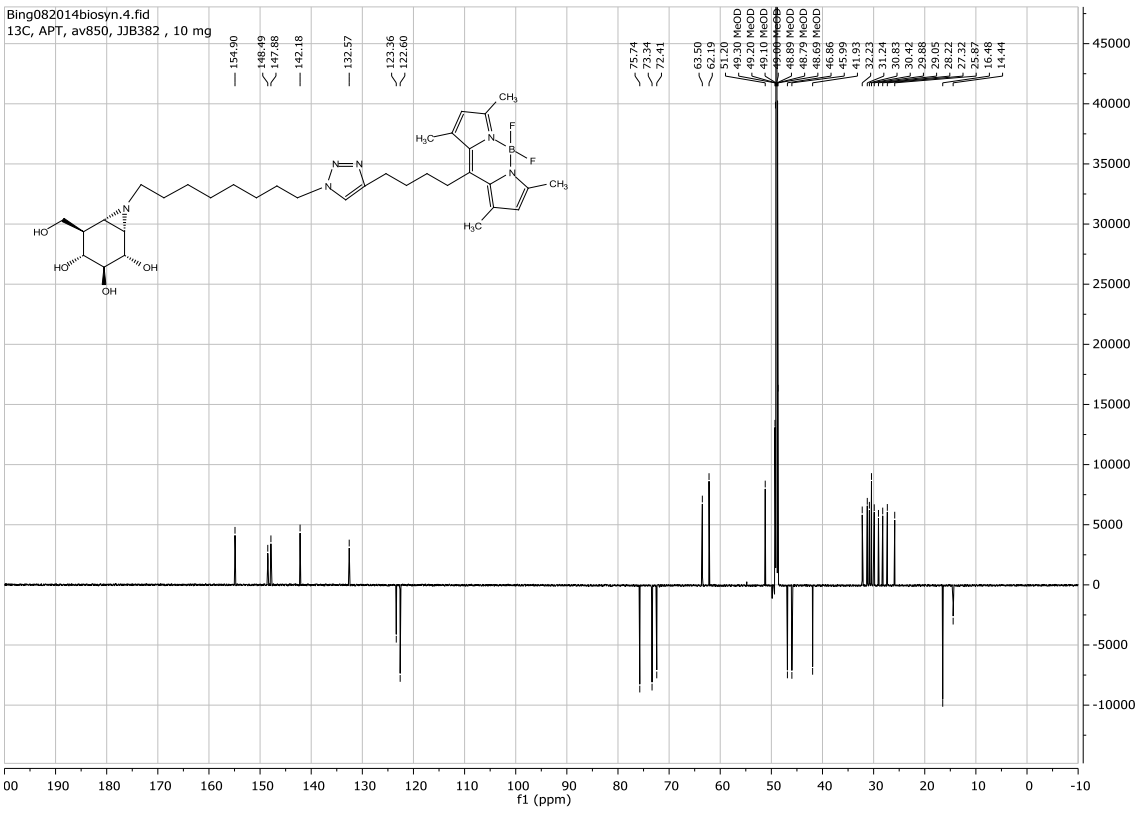
^{1}H and ^{13}C NMR spectra of compound **2**



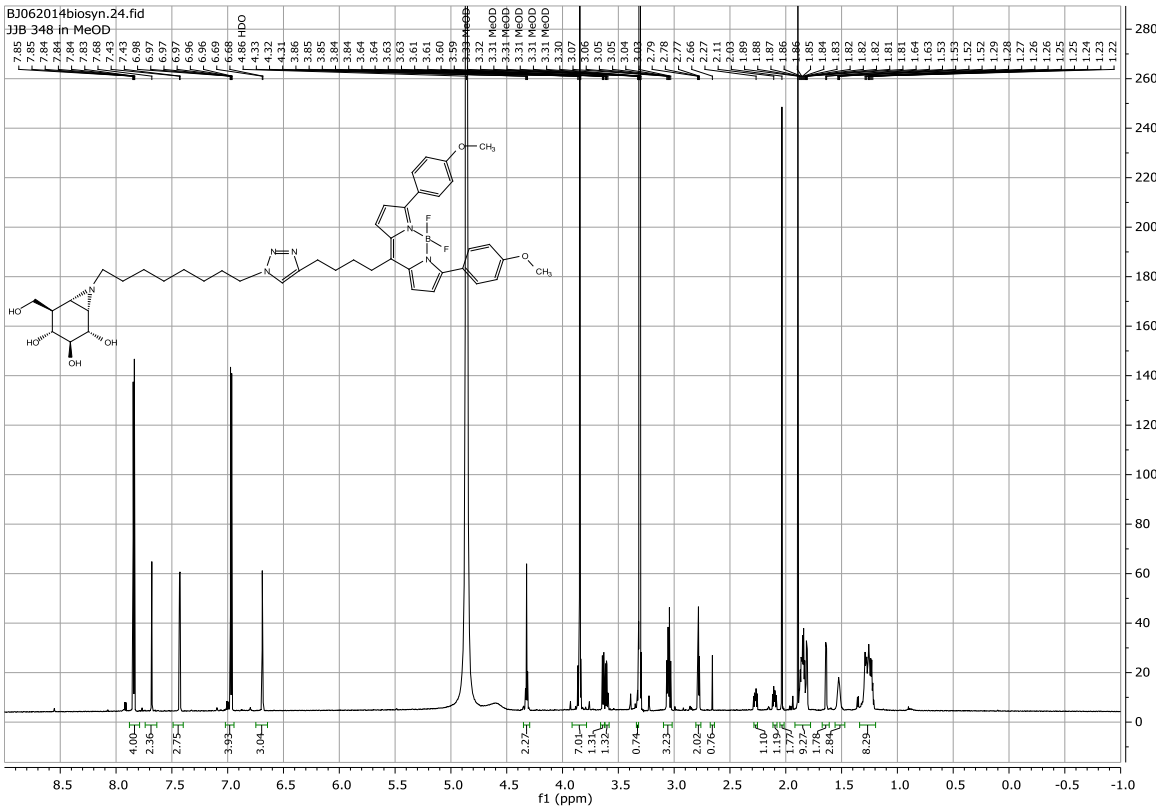


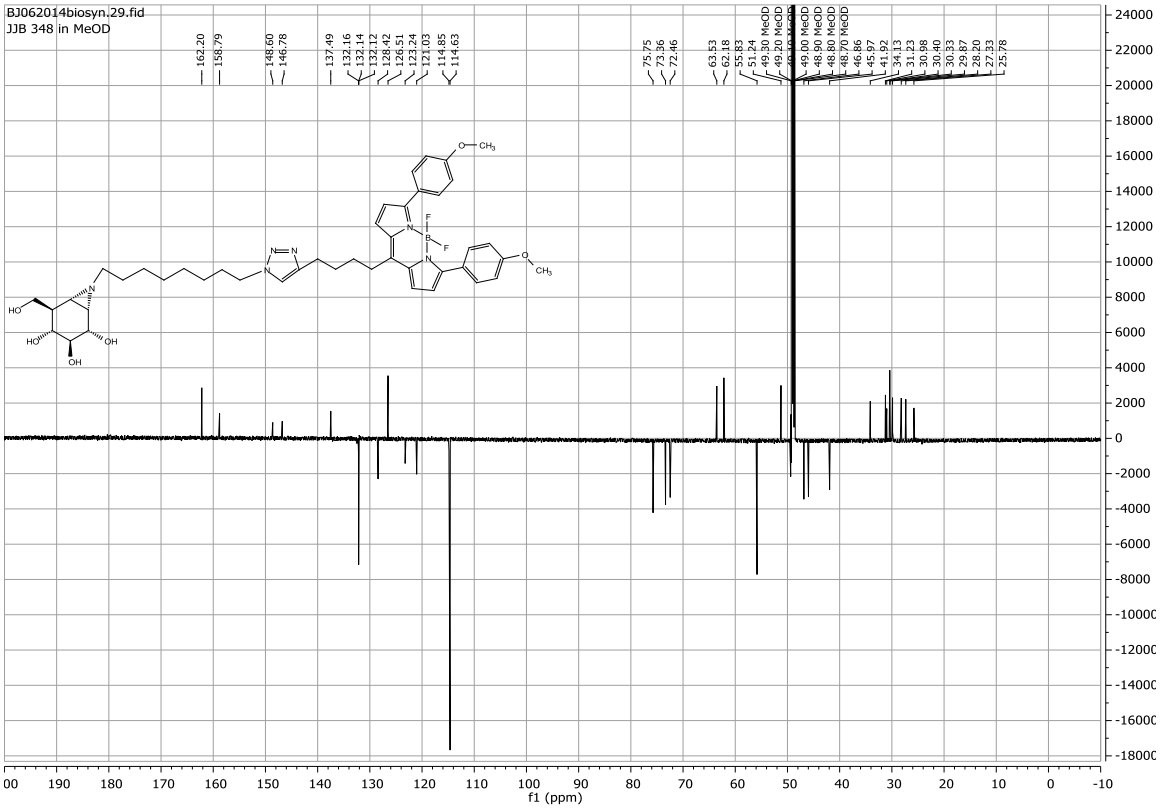
^{1}H and ^{13}C NMR spectra of compound **3**



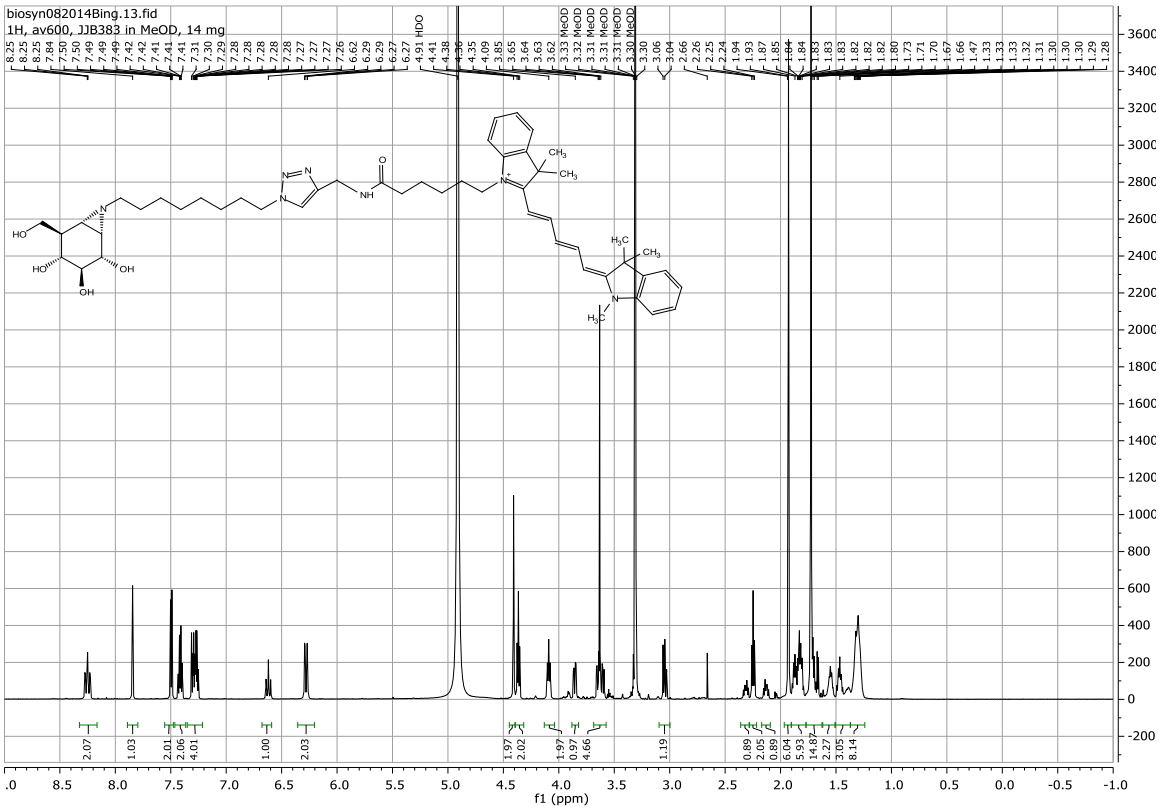


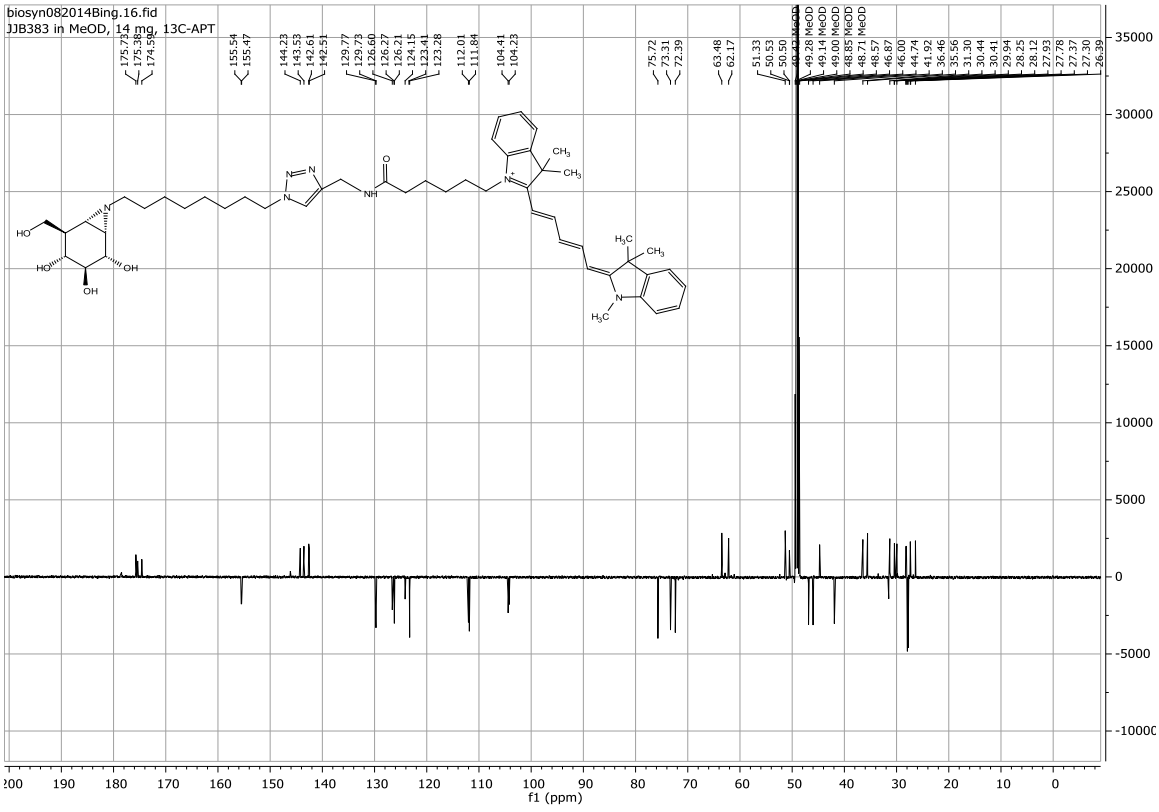
^{1}H and ^{13}C NMR spectra of compound **4**



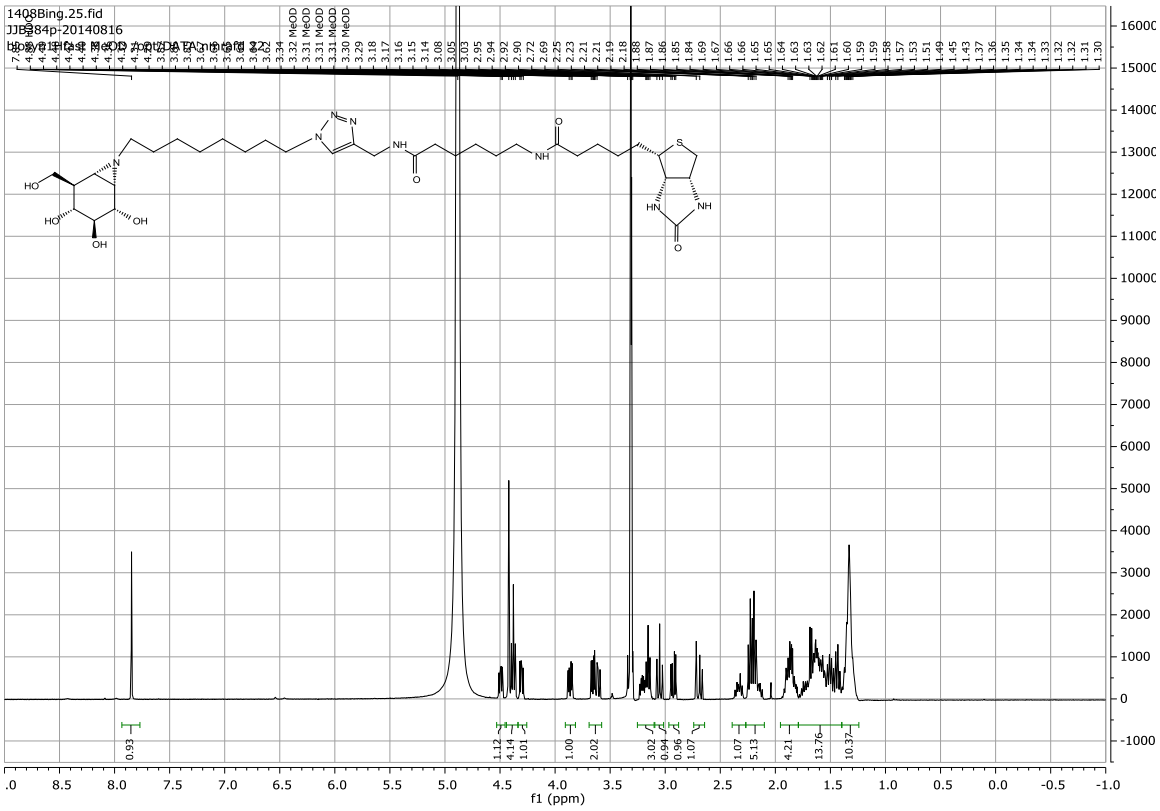


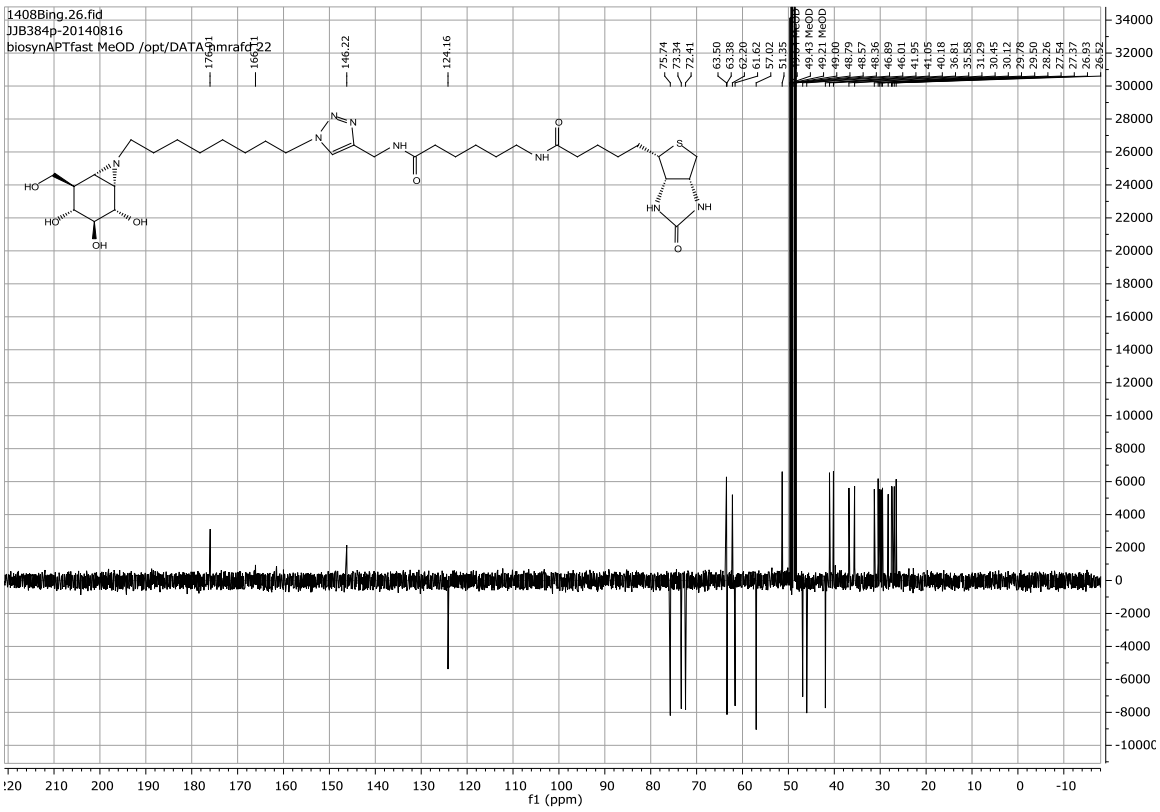
^{1}H and ^{13}C NMR spectra of compound **5**



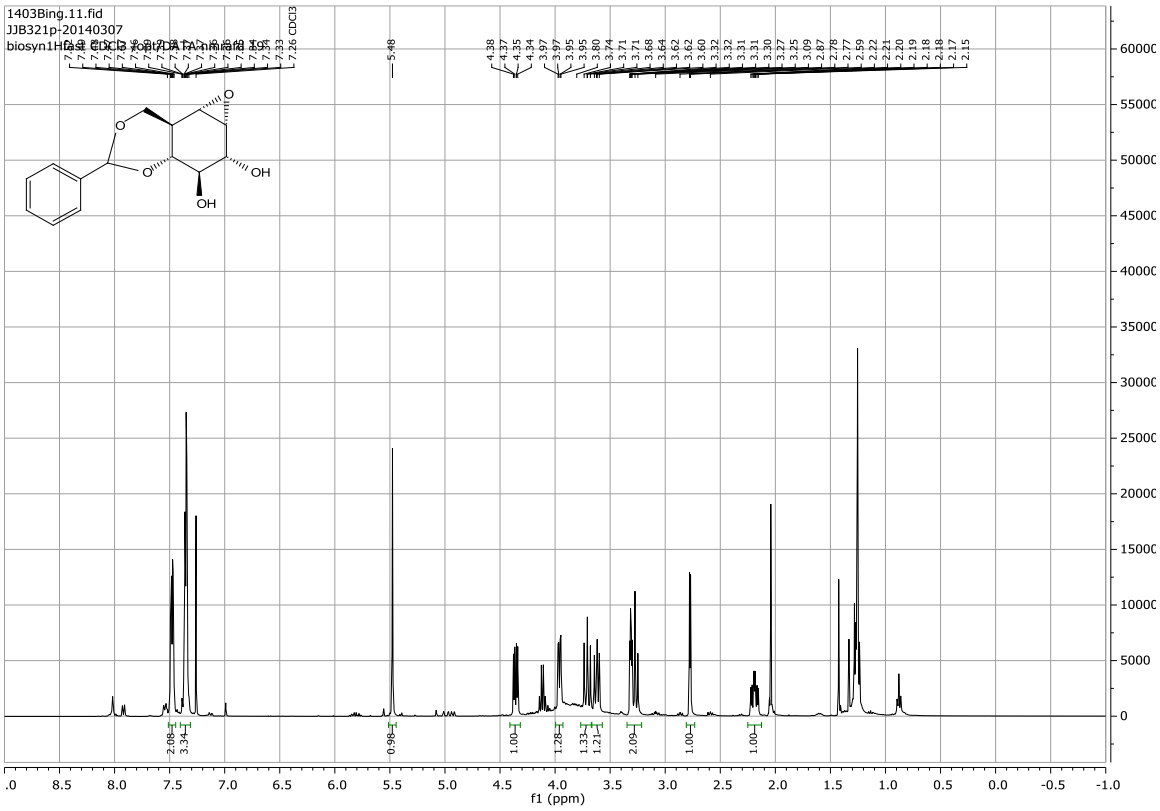


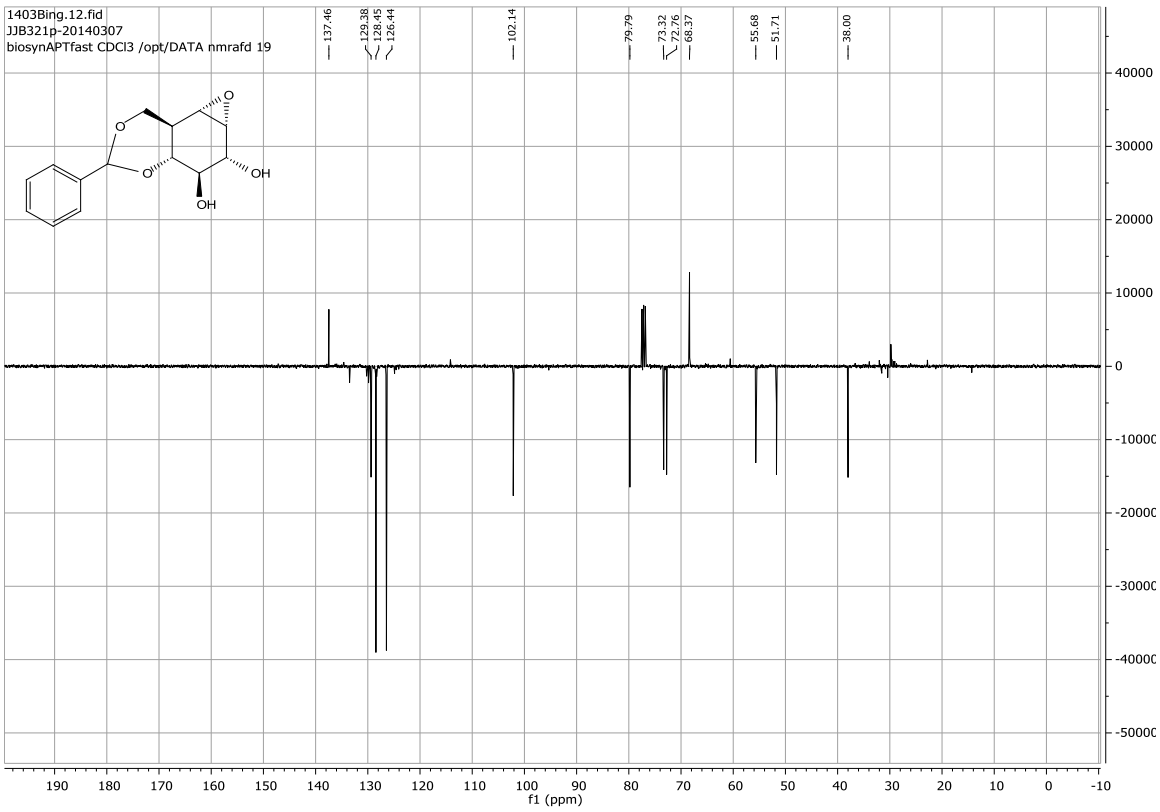
^{1}H and ^{13}C NMR spectra of compound **6**



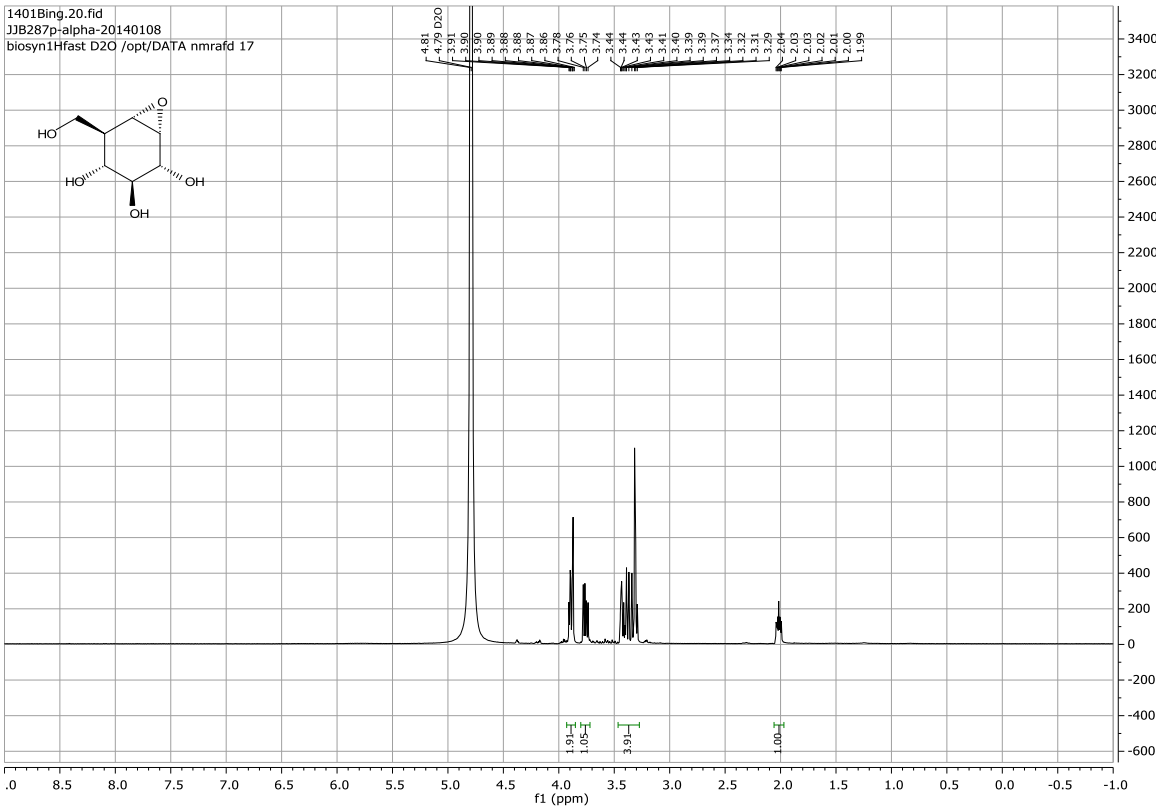


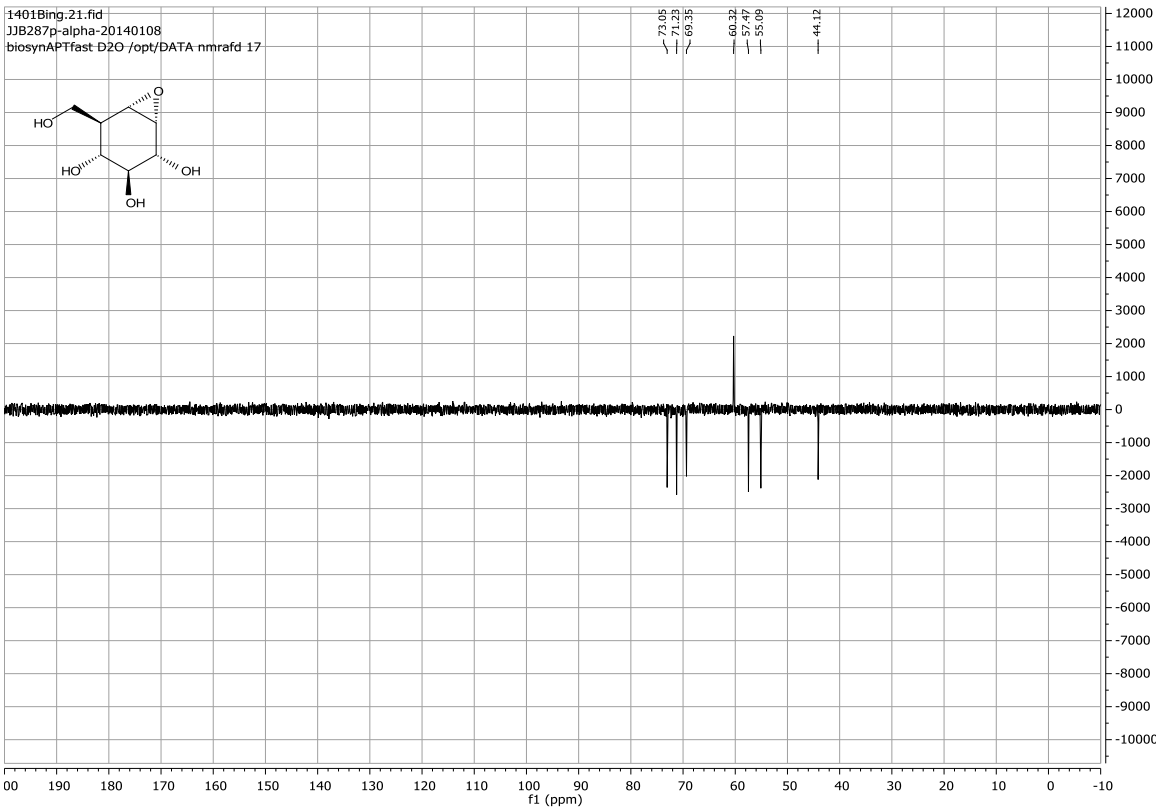
^{1}H and ^{13}C NMR spectra of compound **17**



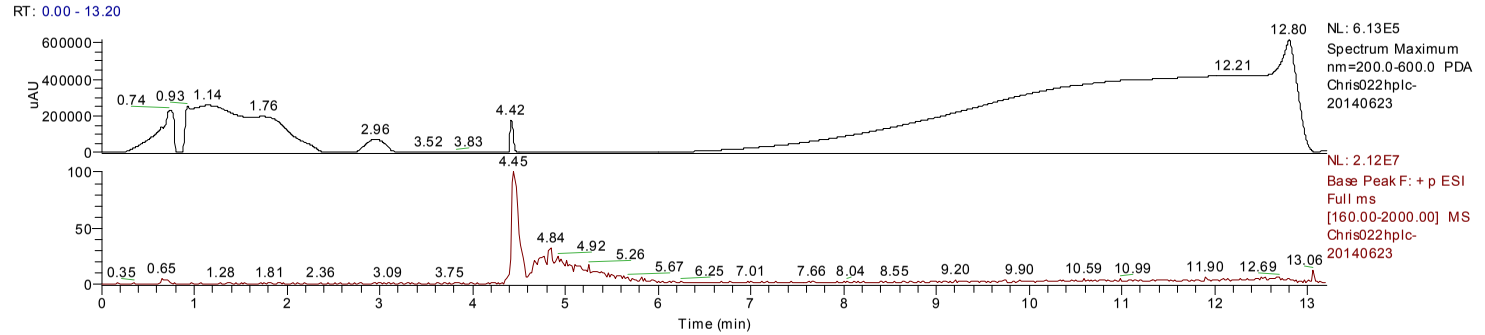


^{1}H and ^{13}C NMR spectra of compound **1**

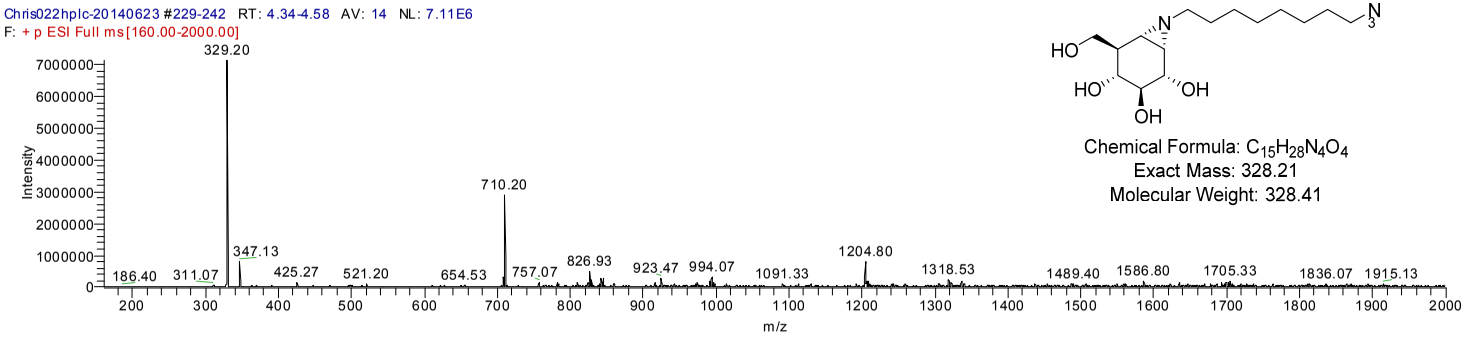




Chris022hplc-20140623 2014/6/23 21:27:10

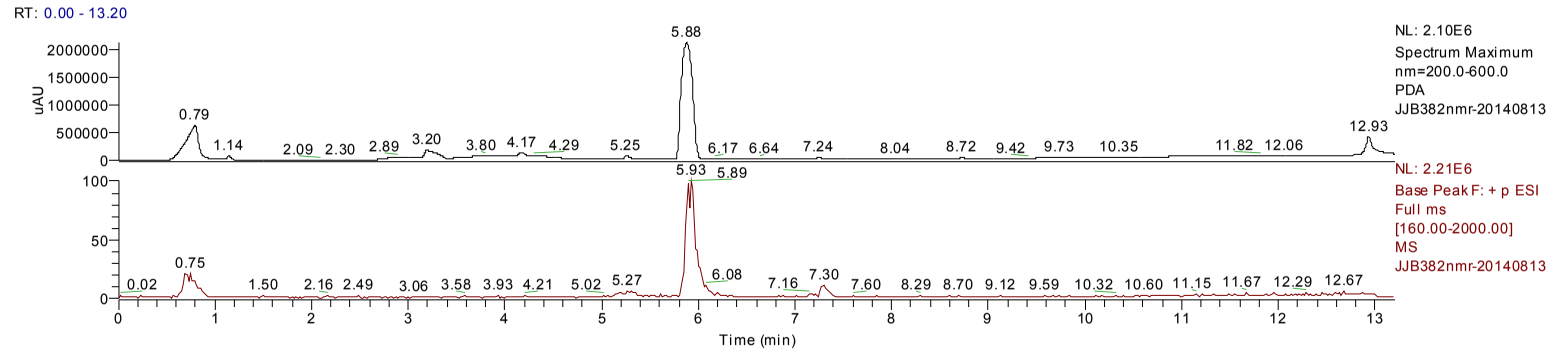


Instrument Method: C:\Xcalibur\methods\General(TFA)\1090%B10_C=TFA_20ul_+pf 12,5min.meth
Processing Method:

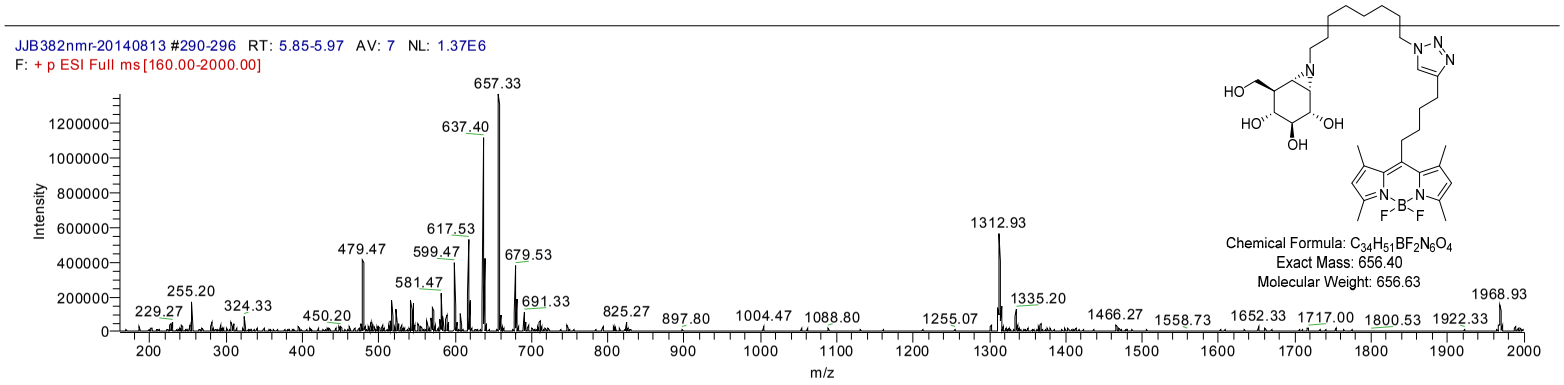


Chemical Formula: $C_{15}H_{28}N_4O_4$
Exact Mass: 328.21
Molecular Weight: 328.41

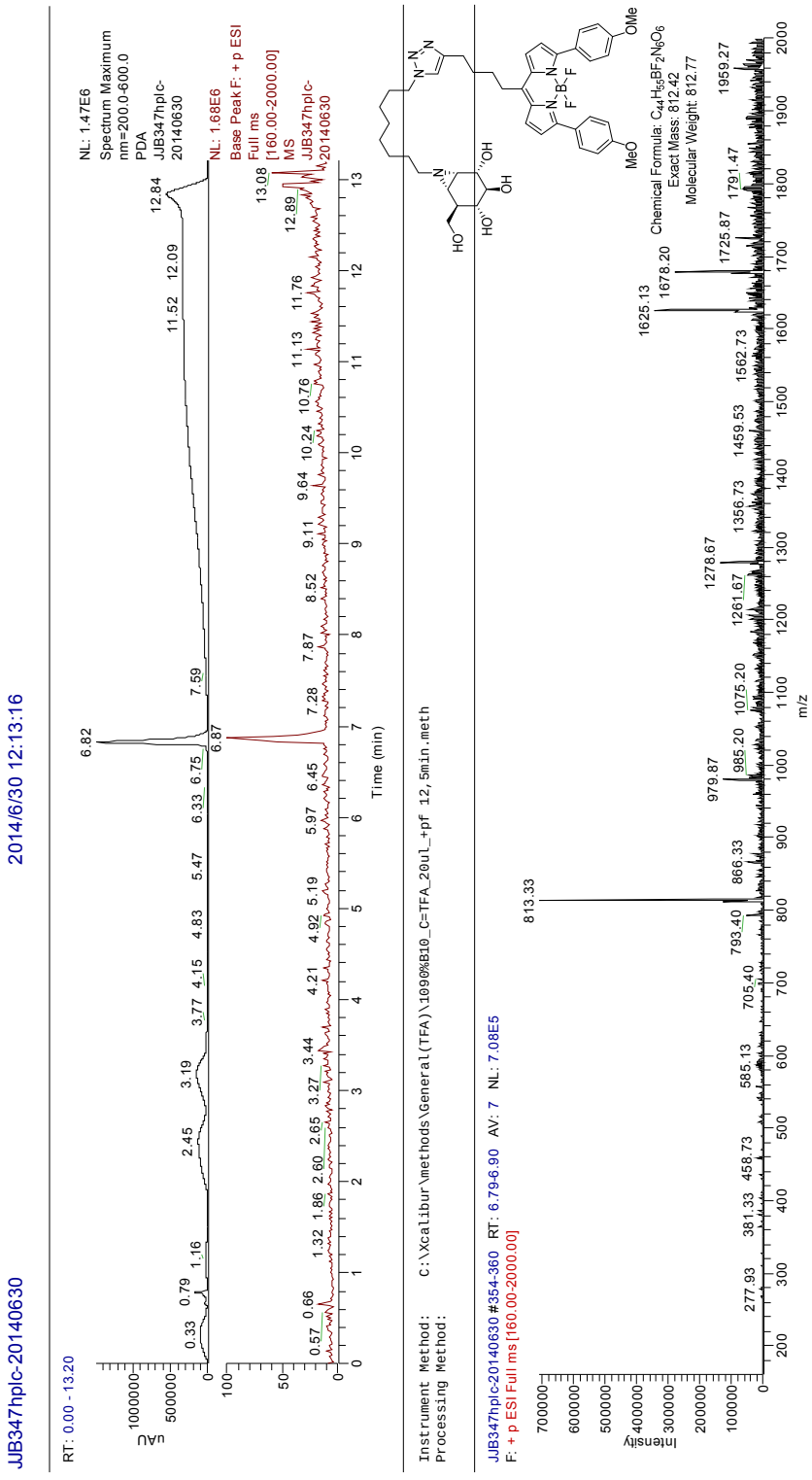
JJB382nmr-20140813 2014/8/13 16:45:49

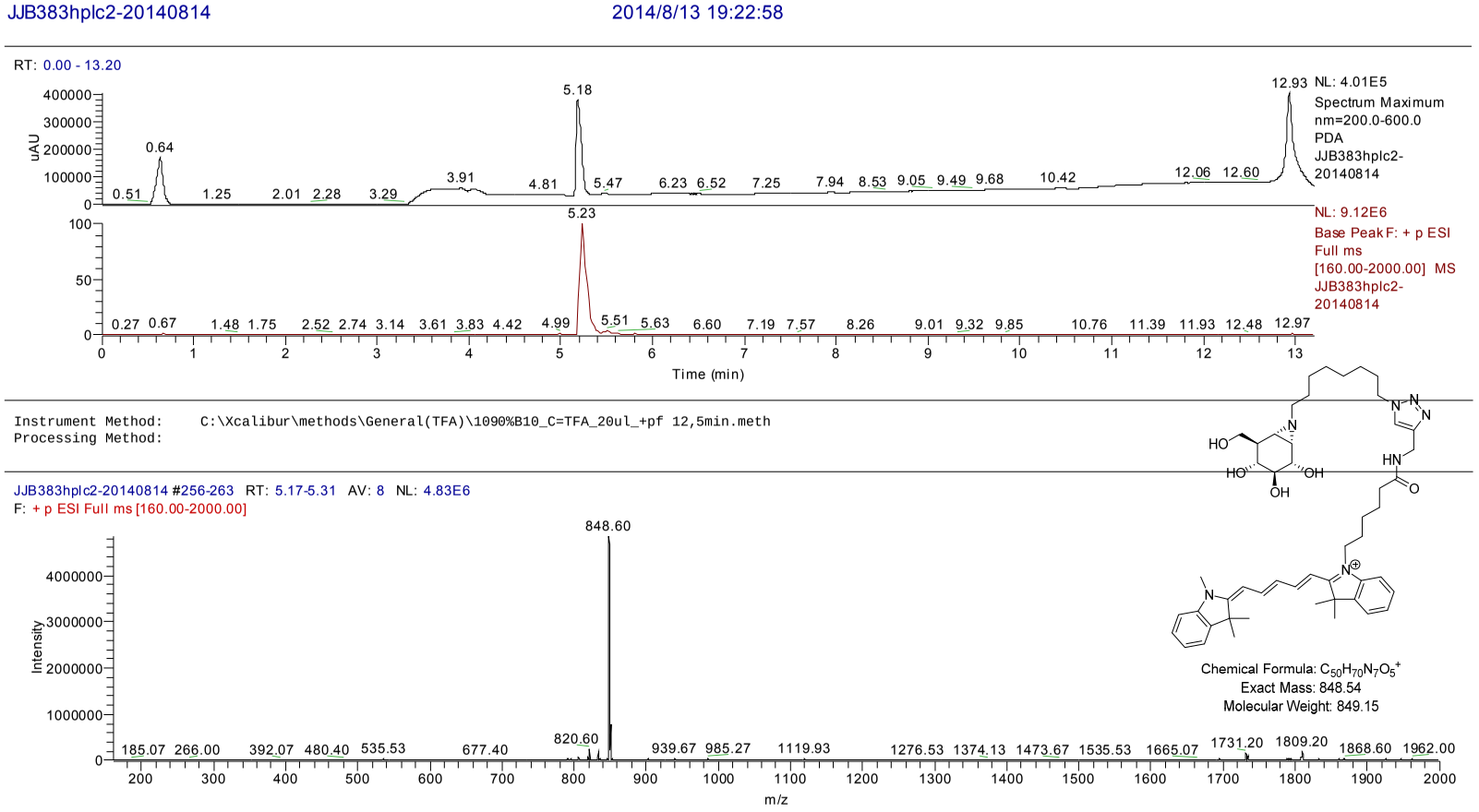


Instrument Method: C:\Xcalibur\methods\General(TFA)\1090%B10_C=TFA_20ul_+pf 12,5min.meth
Processing Method:



LC/MS spectra of compound **4**





LC/MS spectra of compound **5**

LC/MS spectra of compound **6**