

1,6-Cyclophellitol cyclosulfates: a new class of irreversible glycosidase inhibitor

Marta Artola,[†] Liang Wu,[‡] Maria J. Ferraz,[§] Chi-Lin Kuo,[§] Lluís Raich,^{||} Imogen Z. Breen,[‡] Wendy A. Offen,[‡] Jeroen D. C. Codée,[†] Gijsbert A. van der Marel,[†] Carme Rovira,^{||,⊥} Johannes M. F. G. Aerts,[§] Gideon J. Davies,^{*,‡} and Herman S. Overkleeft^{*,†}

[†]Department of Bio-organic Synthesis and [§]Department of Medical Biochemistry, Leiden Institute of Chemistry, Leiden University, P.O. Box 9502, 2300 RA Leiden, the Netherlands

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

^{||}Departament de Química Inorgànica i Orgànica (Secció de Química Orgànica) and Institut de Química Teòrica i Computacional (IQTUB), Universitat de Barcelona, Martí i Franquès 1, 08028 Barcelona, Spain.

[⊥]Fundació Catalana de Recerca i Estudis Avançats (ICREA), Passeig Lluís Companys 23, 08010 Barcelona, Spain.

*Corresponding authors: gideon.davies@york.ac.uk; h.s.overkleeft@chem.leidenuniv.nl

TABLE OF CONTENTS

1. Supporting Figures and Tables	S2
2. Materials and Methods	S11
2.1. Biochemical and Biological Methods	S11
2.2. Molecular Modeling	S14
2.3. Crystallographic data collection and refinement statistics	S15
2.4. Chemical Synthesis	S17
2.4.1. General Experimental Details	S17
2.4.2. Synthesis and Characterization Data of Compounds 5–6	S18
3. NMR Spectra	S22
4. References	S29

1. Supporting Figures and Tables

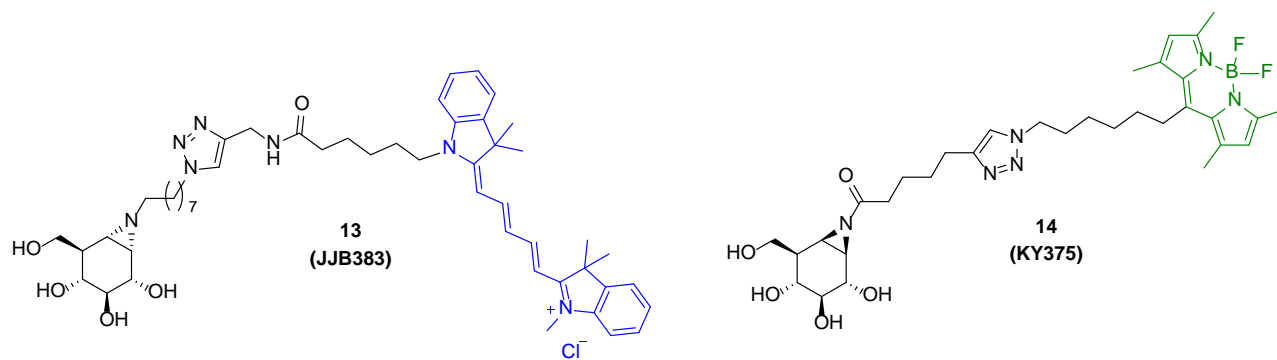


Figure S1. Chemical structures of additional compounds used in this work.

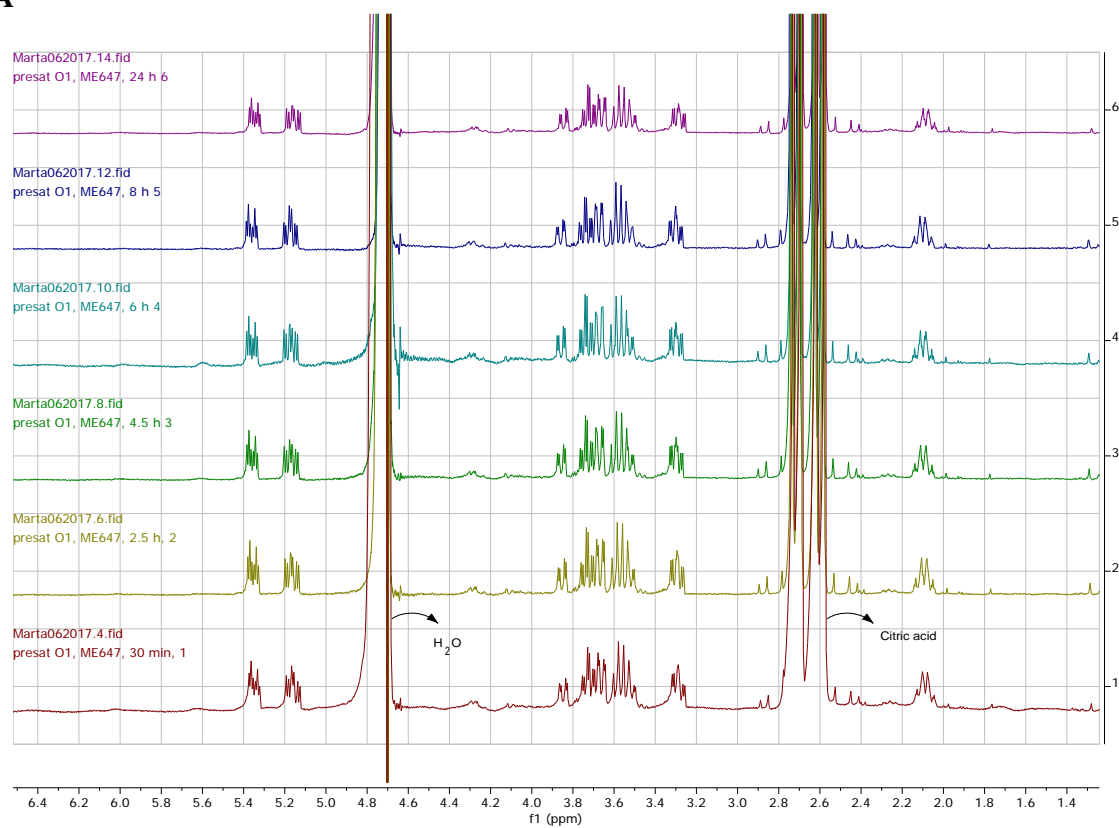
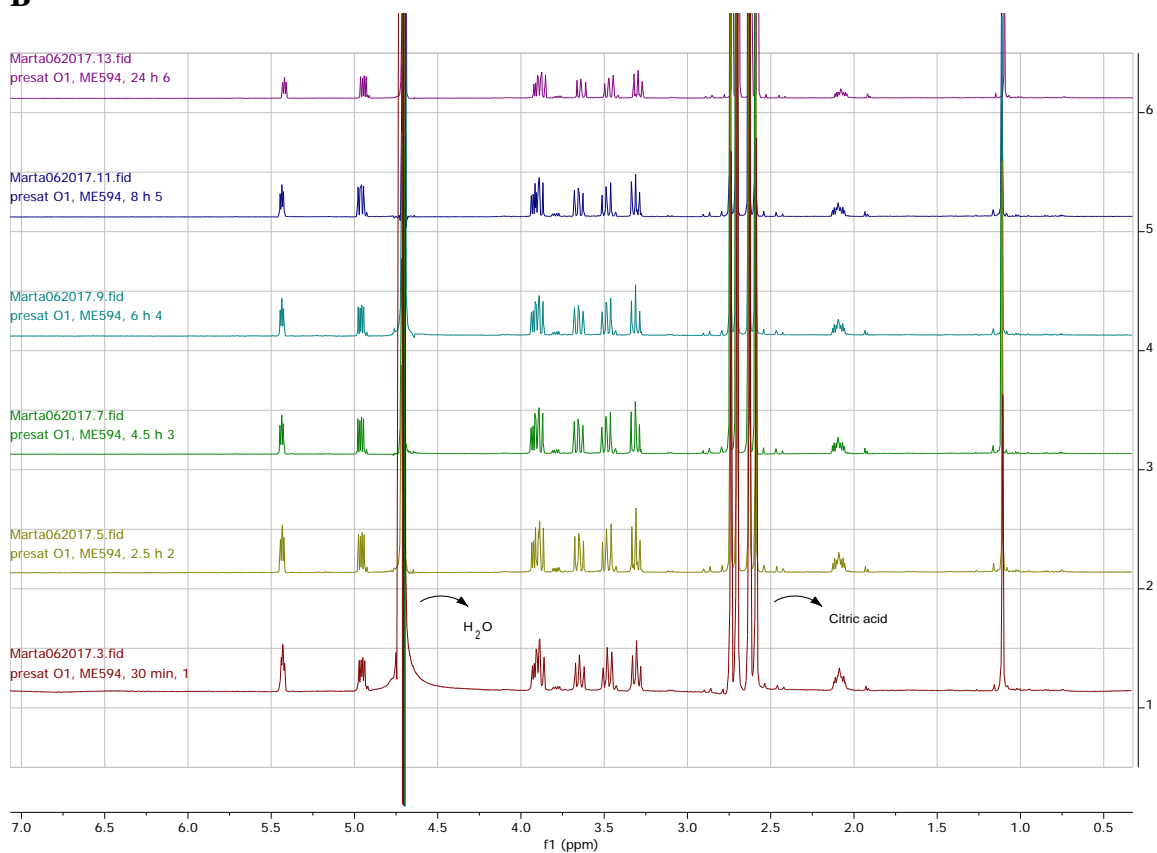
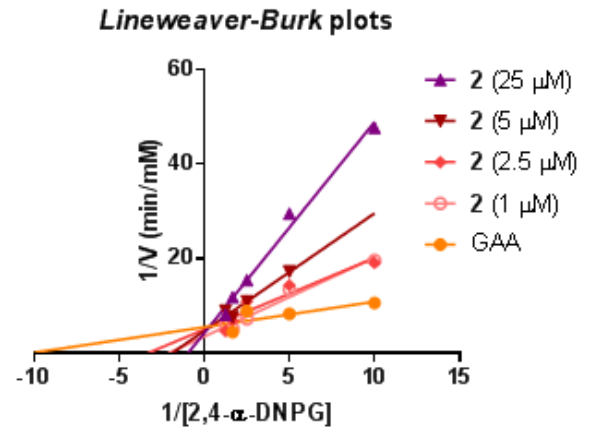
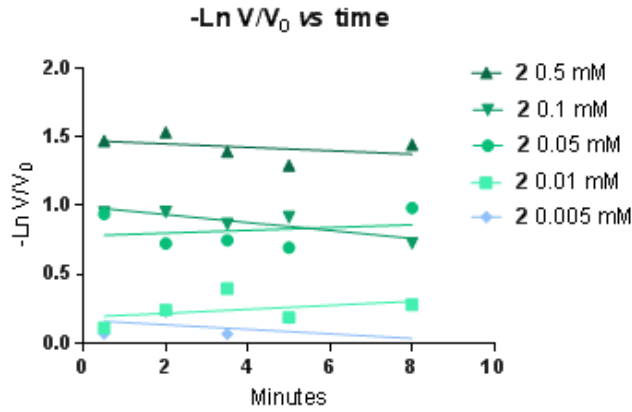
A**B**

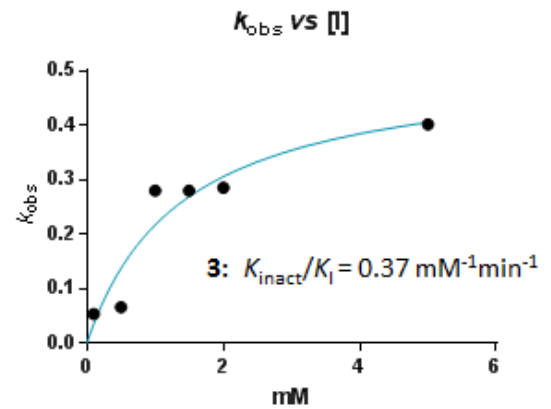
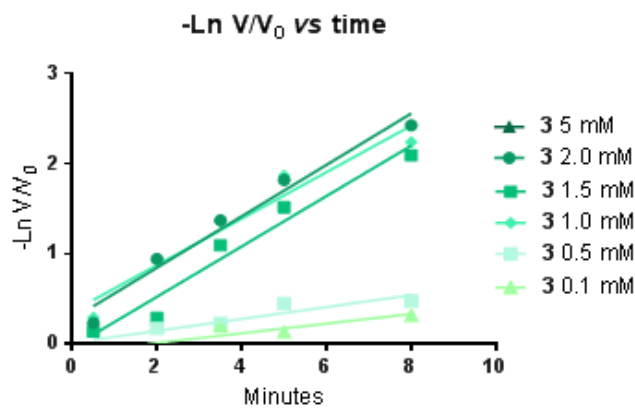
Figure S2. ¹H-NMR spectra of cyclosulfate **5** (A) and **6** (B) in 10% D₂O and 90% 150 mM McIlvaine buffer pH 4.0 registered after 0.5, 2.5, 4.5, 6, 8 and 24 h.

A GAA Inhibition ($K_m = 0,098 \text{ mM}$)

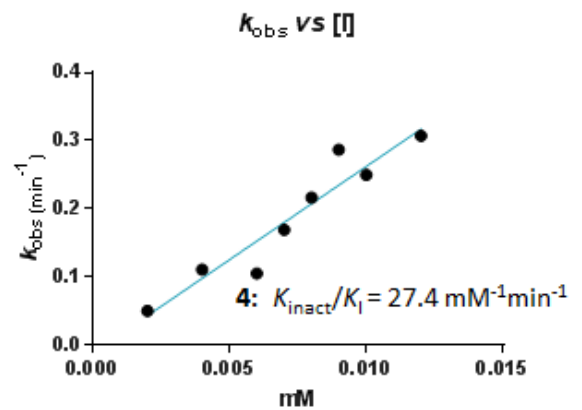
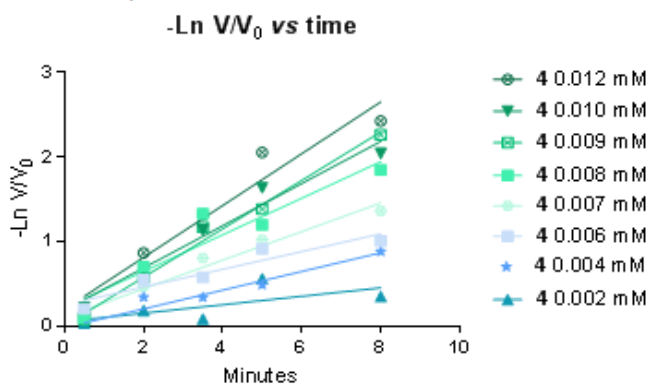
2: $K_i = 9,63 \mu\text{M}$ (Non-Covalent Inhibition)



3: $K_i = 1.39 \text{ mM}$; $K_{\text{inact}} = 0,518 \text{ min}^{-1}$

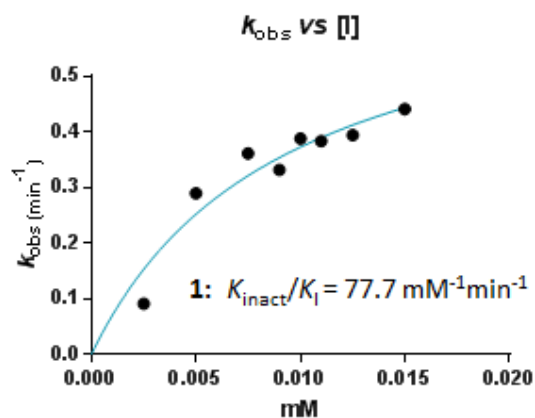
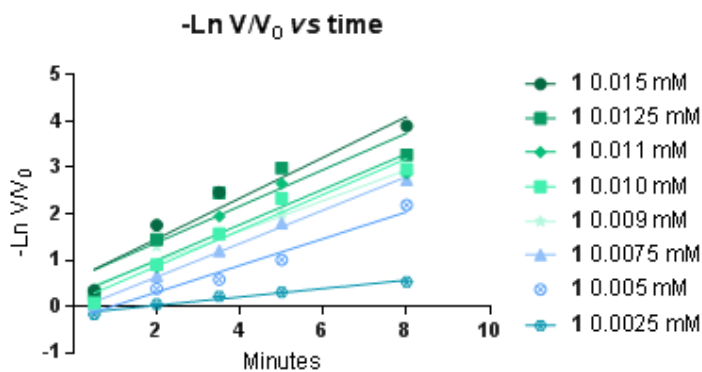


4: $K_{\text{inact}}/K_i = 27.4 \text{ mM}^{-1}\text{min}^{-1}$

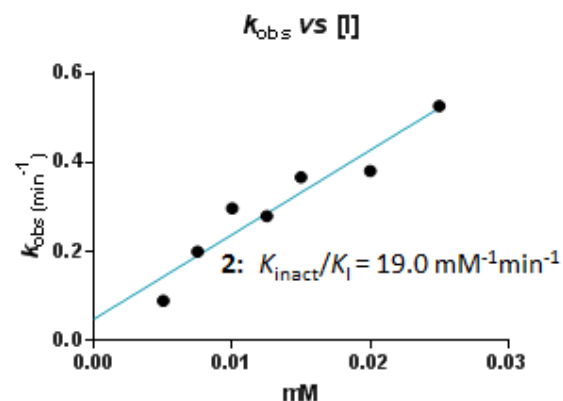
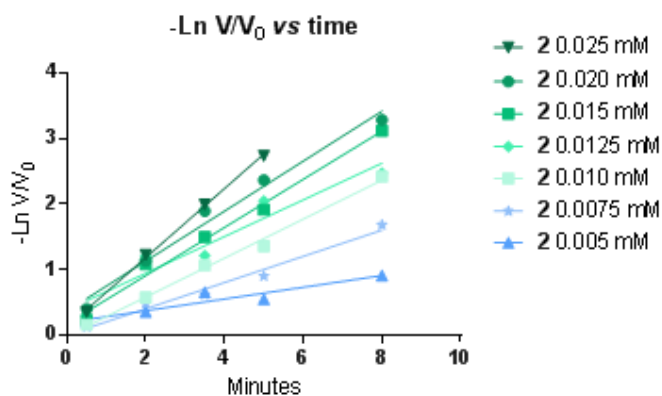


B GBA1 Inhibition ($K_m = 0,168 \text{ mM}$)

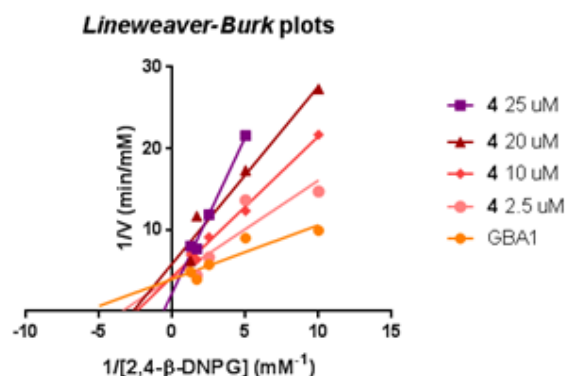
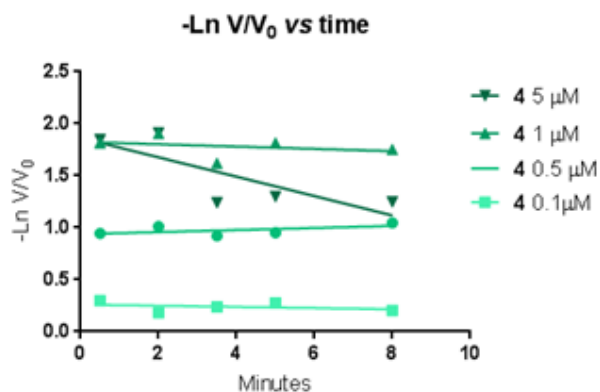
1: $K_I = 9.2 \text{ } \mu\text{M}$; $K_{inact} = 0.716 \text{ min}^{-1}$



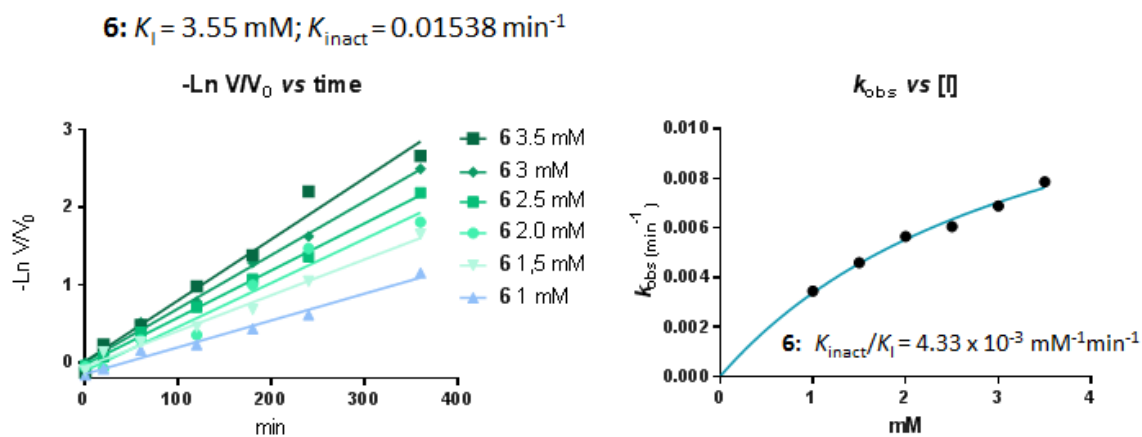
2: $K_{inact}/K_I = 19.0 \text{ mM}^{-1}\text{min}^{-1}$



4: $K_I = 6.83 \text{ } \mu\text{M}$ (Non-Covalent Inhibition)



B GBA1 Inhibition ($K_m = 0.168$ mM)



C *TmGH1* Inhibition ($K_m = 0.24$ mM)

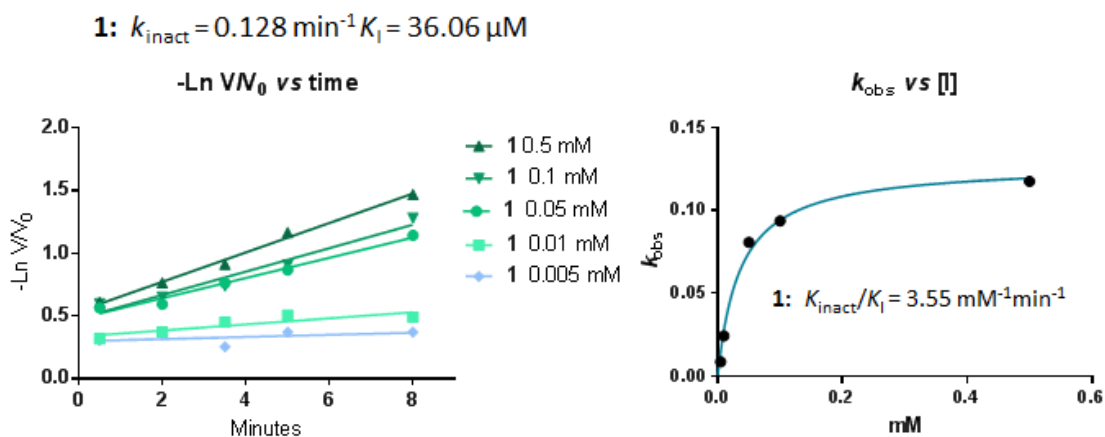
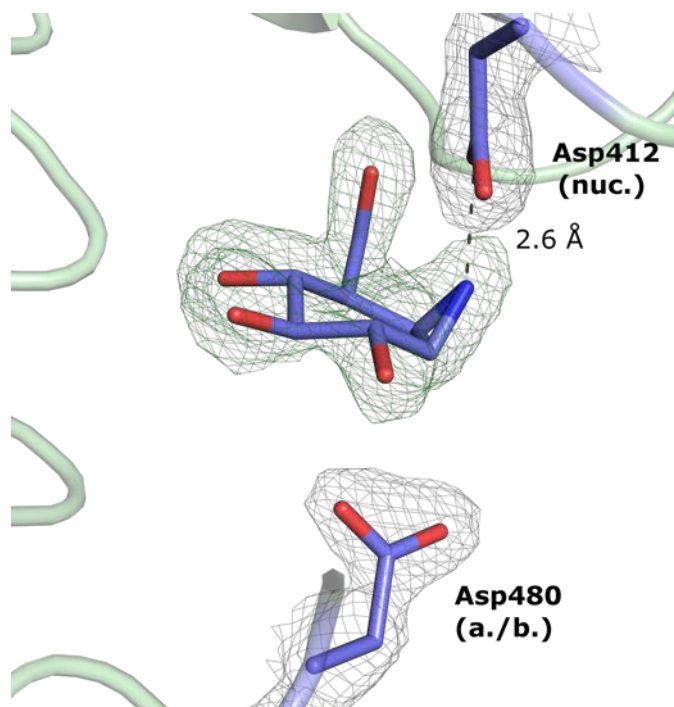


Figure S3. Inhibition kinetics for compounds 1–6 against representative α - and β - glucosidases. **a** Inhibition of α -glucosidase GAA. Plots of residual GAA activity vs time at different concentrations of inhibitors **2**, **3** and **4**. Parameters for irreversible inhibition by **3** and **4** is modeled by plotting measured inactivation rate constants (k_{obs}) vs [I] (see also: Main text **Figure 3c**). K_i for reversible inhibition by **2** is modeled using a set of *Lineweaver-Burk* plots in the absence and presence of varying concentrations of inhibitor. **b** Inhibition of β -glucosidase GBA1 by **1**, **2**, **4** and **6**. Plots as described for GAA. **c** Irreversible inhibition of *TmGH1* by **1**.



Off target binding of **2** E_3 conformation

Figure S4. Unreacted **2** in complex with *CjAgd31B*. Cyclophellitol aziridine **2** binds to the active site of α -glucosidase *CjAgd31B*, forming a H-bonding interaction with the catalytic nucleophile residue Asp412. **2** adopts an E_3 conformation in the active site of *CjAgd31B*, rather than the 4H_3 conformation seen for the ‘correct’ aziridine **4**. Electron density for protein sidechains is REFMAC maximum-likelihood/ σ_A -weighted $2F_o - F_c$ contoured to 0.55 electrons/ \AA^3 . Electron density for **2** is $F_o - F_c$, calculated just prior to building in ligand, contoured to 0.36 electrons/ \AA^3 . nuc. = nucleophile, a./b. = acid/base.

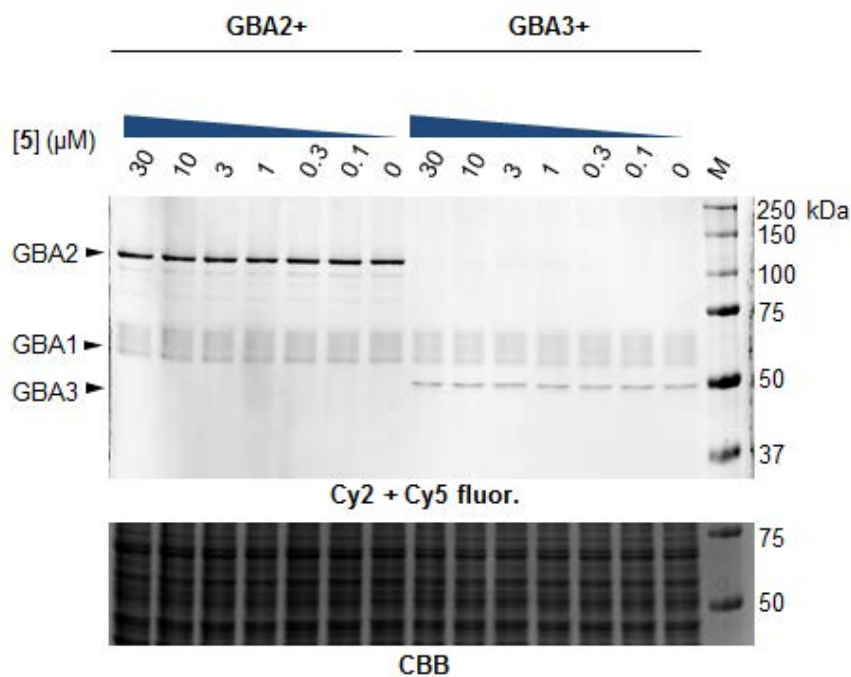
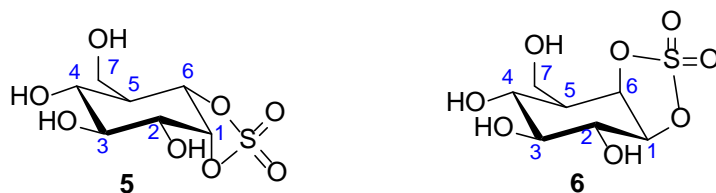


Figure S5. *In vitro* selectivity of α -cyclosulfate **5** against β -glucosidases. Incubation of **5** (from 30 to 0 μ M) in GBA2-expressing Hek293T lysates (GBA2+) and GBA3-expressing Hek293T lysates (GBA3+), followed by labeling of β -glucosidases by 500 nM broad spectrum cyclophellitol-aziridine-BODIPY probe **14** (KY375)¹. No abrogation of labeling by **14** is observed. Overlay image of Cy2 channel used for **14** and Cy5 channel for the marker. CBB: coomassie brilliant blue stain used for loading control.

Table S1. Experimental coupling constants of cyclosulfates **5** and **6**.

Coupling	<i>J</i> type 5	<i>J</i> exp. 5 (Hz)	<i>J</i> type 6	<i>J</i> exp. 6 (Hz)
H1-H6	<i>J</i> eq-ax	4.2	<i>J</i> ax-eq	4.7
H1-H2	<i>J</i> eq-ax	3.9	<i>J</i> ax-ax	8.5
H2-H3	<i>J</i> ax-ax	n.d.	<i>J</i> ax-ax	9.9
H3-H4	<i>J</i> ax-ax	9.5	<i>J</i> ax-ax	9.5
H4-H5	<i>J</i> ax-ax	9.5	<i>J</i> ax-ax	11.1
H5-H6	<i>J</i> ax-ax	10.2	<i>J</i> ax-eq	3.3
H5-H7a	-	2.5	-	4.8
H5-H7b	-	2.8	-	n.d.
H7a-H7b	-	11.2	-	10.6

Coupling constants were determined by ¹H NMR experiments (exp.). n.d.: values could not be observed because of peak overlap or small couplings.

Table S2. Enzyme inhibition efficacy of compounds **1–9** against bacterial glucosidases. Apparent IC₅₀ values for *in vitro* inhibition of *TxGH116*, *TmGH1*, and *CjAgd31B*. Values are means ± standard deviations from three technical replicates.

Compound	IC ₅₀ (μM)		
	β-glucosidase		α-glucosidase
	<i>TmGH1</i>	<i>TxGH116</i>	<i>CjAgd31B</i>
1	0.37 ± 0.02	0.009 ± 0.007	>100
2	0.13 ± 0.01	0.007 ± 0.0004	>100
3	>100	>100	>100
4	30.6 ± 3.12	0.46 ± 0.01	1.42 ± 0.06
5	>100	77.5 ± 2.38	0.49 ± 0.02
6	>100	>100	>100

Table S3. Crystal data collection and refinement statistics.

	<i>Cj</i> Agd31B-5 complex (5NPB)	<i>Cj</i> Agd31B(D412N)-5 complex (5NPC)	<i>Cj</i> Agd31B(D412N)-4 complex (5NPD)	<i>Cj</i> Agd31B-2 complex (5NPE)	<i>Tx</i> GH116-6 reacted complex (5NPF)	<i>Tx</i> GH116-6 unreacted complex (5OOS)
Data collection						
Space group	P622	P622	P622	P622	P2 ₁ 2 ₁ 2	P2 ₁ 2 ₁ 2
Cell dimensions						
<i>a</i> , <i>b</i> , <i>c</i> (Å)	197.4, 197.4, 102.9	197.1, 197.1, 102.8	197.3, 197.3, 102.7	197.0, 197.0, 103.0	178.2, 53.8, 83.3	178.2, 53.7, 83.1
α , β , γ (°)	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 90	90, 90, 90
Resolution (Å)	49.85-1.90 (1.93-1.90)	65.67-1.96 (2.01-1.96)	47.39-1.95 (1.99-1.95)	44.43-1.95 (1.99-1.95)	53.81-1.38 (1.42-1.38)	60.78-1.16 (1.18-1.16)
<i>R</i> _{merge}	0.08 (0.76)	0.21 (3.00)	0.10 (1.41)	0.07 (0.46)	0.069 (0.97)	0.058 (1.63)
<i>I</i> / σ <i>I</i>	30.2 (5.5)	11.4 (1.2)	21.7 (2.6)	28.8 (6.8)	15.1 (1.3)	13.9 (1.1)
Completeness (%)	100 (100)	99.9 (99.9)	100 (100)	100 (100)	99.3 (93.5)	100 (99.9)
Redundancy	24.1 (24.1)	20.0 (19.2)	20.1 (20.3)	20.0 (19.0)	6.1 (3.4)	7.8 (7.5)
Refinement						
Resolution (Å)	49.85-1.90	65.67-1.96	47.39-1.95	44.43-1.95	51.56-1.38	60-78-1.16
No. reflections	88031	79999	81290	81221	155502	261679
<i>R</i> _{work} / <i>R</i> _{free}	0.20/0.23	0.19/0.23	0.17/0.21	0.19/0.24	0.13/0.16	0.13/0.16
No. atoms						
Protein	6302	6283	6302	6307	6423	6393
Ligand/ion	60	123	60	69	69	53
Water	502	362	502	568	567	550
B-factors						
Protein	32.2	38.8	32.2	29.7	18.6	18.8
Ligand/ion	50.7	62.5	50.7	54.1	30.8	26.1
Water	35.6	39.6	35.6	34.5	33.2	31.8
R.m.s deviations						
Bond lengths (Å)	0.014	0.012	0.013	0.015	0.017	0.016
Bond angles (°)	1.66	1.57	1.60	1.70	1.77	1.72

2. Materials and Methods

2.1. Biochemical and Biological Methods

IC₅₀ measurements

Enzyme preparations used for IC₅₀ and kinetics measurements were as follows: Recombinant human GBA1 (Cerezyme) and recombinant human GAA (Myozyme) were obtained from Genzyme, USA. For GBA2 - cellular homogenates of HEK293T over-expressing GBA2 pre-incubated for 30 min with an inhibitor of GBA1 (1 μM MDW933²). For GANAB - fibroblasts of Pompe patients diagnosed on the basis of absence GAA. Bacterial enzymes *TmGH1*³, *TxGH116*⁴ and *CjAgd31B*⁵ were expressed as previously described. All cell or tissue lysates were prepared in KPI buffer (25 mM Potassium Phosphate pH 6.5, supplemented with protease inhibitor 1x cocktail (Roche)) via homogenization on ice with silent crusher S equipped with Typ 7 F/S head (30 rpm x 1000, 3 × 7 sec). Lysate protein concentrations were determined with BCA Protein Assay Kit (Pierce). Lysates and proteins were stored in small aliquots at -80 °C until use.

***In vitro* apparent IC₅₀:** To determine *in vitro* apparent IC₅₀s, 12.5 μL of enzyme mixture was pre-incubated with 12.5 μL of inhibitor for 30 min, in the following buffers: GBA1 in 150 mM McIlvaine buffer pH 5.2, 0.2% Taurocholate (w/v), 0.1% Triton X-100 (v/v), 0.1% Bovine Serum albumin (BSA) (w/v). GBA2 in 150 mM McIlvaine pH 5.8, 0.1% BSA (w/v). GAA in 150 mM McIlvaine buffer pH 4.0, 0.1% BSA (w/v). GANAB in 150 mM McIlvaine buffer pH 7.0, 0.1% BSA (w/v). *TmGH1* and *TxGH116* in 50 mM NaHPO₄ pH 6.8, 0.1% BSA (w/v). *CjAgd31B* in 50 mM citrate buffer pH 6.5, 0.1% BSA (w/v).

Following pre-incubation, 25 μL of this E + I mix was added to 100 μL of substrate solution in the same buffer. GBA1 residual activity was measured using final 0.7 nM concentration of enzyme (cerezyme) and 3.0 mM 4-methylumbeliferone(4MU)-β-D-glucopyranoside, for 30 min at 37 °C. GBA2 residual activity was measured using cellular homogenates of HEK293T over-expressing GBA2 pre-incubated for 30 min with an inhibitor of GBA1 (1 μM MDW933²), and further incubation with 3.0 mM 4-methylumbeliferone(4MU)-β-D-glucopyranoside for 1 h at 37 °C. GAA activity was measured using final concentrations of 47 nM and 2.4 mM 4MU-α-D-glucopyranoside, for 1 h at 37 °C. GANAB activity was measured using fibroblasts of Pompe patients diagnosed on the basis of absence GAA activity and 2.4 mM 4MU-α-D-glucopyranoside, for 2 h at 37 °C. *TmGH1* and *TxGH116* residual activity was measured using final concentrations of 0.1 nM and 0.2 nM respectively and 3.0 mM 4MU-β-D-glucopyranoside, for 30 min at 25 °C. *CjAgd31B* residual activity was measured using final concentrations of 6.5 μM and 2.4 mM 4MU-α-D-glucopyranoside, for 1 h at 25 °C. Finally, all enzyme reactions were quenched with 200 μL 1M NaOH-Glycine (pH 10.3), and liberated 4MU fluorescence measured with a LS55 fluorescence spectrophotometer (Perkin Elmer; λ_{EX} 366 nm, λ_{EM} 445 nm). Values plotted for [I] are those in the final reaction mixture, containing E + I + S. *In vitro* IC₅₀ values were determined in technical triplicate.

***In situ* apparent IC₅₀s:** The *in situ* apparent IC₅₀ value was determined by incubating fibroblast cell lines expressing wild-type GAA, grown to confluence, with a range of inhibitor **5** dilutions for 2 or 24 h. Hereafter, cells were washed three times with PBS and subsequently harvested by scraping in potassium phosphate buffer (25 mM K₂HPO₄-KH₂PO₄, pH 6.5, supplemented with 0.1% (v/v) Triton X-100 and protease inhibitor cocktail (Roche)). Residual GAA activity was measured using lysates containing 4 µg total protein from one duplicate equilibrated to pH 4 or 7 in McIlvaine buffer (150 mM citrate/phosphate) for 5 min on ice, and incubated for 4 h at 37°C with 3.0 mM 4MU- α -glucopyranoside at pH 4.0 or 7.0 in a total volume of 125 µL in a 96-well plate. Reaction was stopped with 200 µL 1 M Glycine-NaOH (pH 10.3), and the microplates were subjected to fluorescent scan (LS-55, PerkinElmer, λ_{EX} = 366 nm , λ_{EM} = 455 nm). 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).

Kinetic studies

Enzyme and relevant inhibitor dilutions were pre-incubated for 0.5, 2, 3.5, 5 and 8 min for fast inhibitors or 0.5, 20, 60, 120, 180, 240, and 360 min for slow inhibitors in the following conditions: 0.47 µM final concentration of GAA at 37 °C in 150 mM McIlvaine buffer pH 4.8, supplemented with 0.1% (w/v) BSA; 0.14 µM final concentration of GBA1 at 37 °C in 150 mM McIlvaine buffer pH 5.2 supplemented with 0.2% Taurocholate (w/v), 0.1% Triton X-100 (v/v), 0.1% Bovine Serum albumin (BSA) (w/v); 37 nM final concentration of *TmGH1* at 25 °C in 50 mM sodium phosphate pH 6.8 with 0.1% BSA (w/v).

Following inhibitor pre-incubations, reactions were started by adding 5 µL of the enzyme-inhibitor mixture to 95 µL of a substrate mix containing 200 µM 2,4-dinitrophenyl- α -D-glucopyranoside for GAA, and 200 µM 2,4-dinitrophenyl- β -D-glucopyranoside for GBA1 and *TmGH1*. Buffers for the final reaction mix were the same as for the pre-incubation mix for each enzyme.

Release of 2,4-dinitrophenolate was monitored via absorbance at 400 nm every 2s for 60 s to determine the hydrolysis rate in the presence of inhibitor (V_i), and in the absence of inhibitor (V_o). Pseudo-first order rate constants (k_{obs}) for each value of [I] were obtained from the gradient of a plot of $-\ln V_i/V_o$ against time. k_{obs} were then plotted against [I], and fitted to the equation $k_{obs} = (k_{inact}[I]/K_I + [I])$ in GraphPad Prism 7. For cases of fast inhibition at high [I] (> 50% inhibition after 30s), a combined k_{inact}/K_I ratio was determined using the approximation $k_{obs} = k_{inact}[I]/K_I$, where k_{inact}/K_I is the slope of a linear fit of k_{obs} vs [I]. For irreversible inhibition kinetics, values plotted for [I] are those in the initial inhibition mixture, containing only E + I.

Where reversible inhibition was observed (no variation of $-\ln V_i/V_o$ with time), inactivation kinetics were reassessed by *Lineweaver-Burk* plot analysis. GBA1 and GAA were pre-incubated with a range of inhibitor dilutions for 15 min at 37 °C (buffers as above), before addition of a range (0.1 – 1.0 mM) of concentrations of 2,4-dinitrophenyl-glucopyranoside substrate to the enzyme-inhibitor mixture (α -glucopyranoside for GAA, β -glucopyranoside

for GBA1). Release of 2,4-dinitrophenolate was monitored via absorbance at 400 nm every 2s for 60 s to determine the hydrolysis rate. K_I values of reversible inhibition were determined by taking the slopes from each Lineweaver-Burk plot ($(K_m(1+[I]/K_I))/k_{cat}$) and plotting them against inhibitor concentration. After fitting the data into a straight line by linear regression, the X-intercept gives the value of $-K_I$. For competitive inhibition, values plotted for [I] are those in the final reaction mixture, containing E + I + S.

SDS-PAGE analysis and fluorescence scanning

Detailed ABP labeling protocols are given below. Following labeling, all samples were denatured with 5x 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 50 min.

Wet slab-gels were scanned for ABP-emitted fluorescence using a Typhoon FLA9500 imager (GE) with the following settings: $\lambda_{EX} \geq 473$ nm, $\lambda_{EM} \geq 510$ nm, for **14**, and $\lambda_{EX} \geq 635$ nm, $\lambda_{EM} \geq 665$ nm for **13**.

Labeling protocols

Generation of stable GBA2- and GBA3-expressing Hek293T: The PCR-amplified human GBA2 (GBA2, acc. nr: NM_020944.2) coding sequence (using the following oligonucleotides: sense 5'- GGGGACAAGTTTGTACAAAAAAGCAGGCTTAACCACCA TGGGGACCCAGGATCCAG-3' and antisense 5'- GGGGACCACTTTGTACAAGAAAGC TGGGTTTCACTCTGGGCTCAGGTTTG-3') was cloned into pDNOR-221 and sub-cloned in pLenti6.3/TO/V5-DEST using the Gateway system (Invitrogen). Correctness of the construct was verified by sequencing. To produce lentiviral particles Hek293T cells were transfected with pLenti6.3-GBA2 in combination with the envelope and packaging plasmids pMD2G, pRRE and pRSV. Subsequently, culture supernatant containing viral particles was collected and used for infection of Hek293T cells. Selection using blasticidin for several weeks rendered cells stably expressing human GBA2 as determined by activity assays and activity based probes (data not shown).

For human GBA3 the PCR-amplified GBA3 (GBA3, acc. Nr: NM_020973.4) coding sequence using the following oligonucleotides: sense. 5'- GAATTCGCCGCCACCATGGCTTTCCTGCAGGATTTG-3' and antisense 5'- GCGGCCGCAGATGTGCTTCAAGGCCATTG-3') was cloned in pcDNA3.1/Zeo and transfected into Hek293T cells using FuGENE® 6 Transfection Reagent (Promega Benelux, Leiden, The Netherlands). Selection using Zeocin for several weeks rendered cells stably expressing human GBA3 as determined by activity assays and Activity based probes (data not shown).

***In vitro* selectivity of α -cyclosulfate 5 vs β -glucosidases:** Cells were pelleted and lysed in Kpi buffer (25 mM K_2HPO_4 - KH_2PO_4 , pH 6.5, 0.1% (v/v) Triton X-100 and protease inhibitor cocktail (Roche, (version 12)) for 1 h on ice and vortexing. Protein concentrations were

measured by BCA assay (Pierce BCA kit, Thermo Fisher). For labeling, 20 μg total protein from the lysate were equilibrated with 750 mM McIlvaine buffer to pH 5.8 (GBA2 labeling) or 6.0 (GBA3 labeling) in a total volume of 10 μL . The samples were treated with **5** at various concentrations for 1h at 37 $^{\circ}\text{C}$ (prepared as 5x stock solutions in 2.5 μL McIlvaine buffer pH 5.8 or 6.0), followed by incubation with 500 nM of probe **14** for 30 min at 37 $^{\circ}\text{C}$ (prepared as 6x stock solution in 2.5 μL McIlvaine buffer pH 5.8 or 6.0). Samples were boiled with 5x Laemmli's sample buffer for 5 min at 98 $^{\circ}\text{C}$, resolved with 10% SDS-PAGE, and scanned for Cy2 ($\lambda_{\text{EX}} \geq 473$ nm, $\lambda_{\text{EM}} \geq 510$ nm) and Cy5 ($\lambda_{\text{EX}} \geq 635$ nm, $\lambda_{\text{EM}} \geq 665$ nm) fluorescence on a Typhoon FLA9500 Imager (GE). Coomassie Brilliant Blue was used for loading control, and the stained gel was imaged by a ChemiDoc MP Imager (Bio-Rad).

Stability of cyclosulfate **5** and **6**

Compounds **5** (1.32 mg) and **6** (1.66 mg) were dissolved in a total volume of 0.5 mL of 10% D2O and 90% 150 mM McIlvaine buffer pH 4.0 and left at room temperature. 1H NMR was registered after 0.5, 2.5, 4.5, 6, 8 and 24 h, showing no acidic hydrolysis after 24 h (**Figure S3**).

Molecular Modeling

The free energy landscapes of the four cyclophellitol sulfates (1,6-epi-cyclophellitol cyclosulfate **5** and cyclophellitol cyclosulfate **6**) were obtained by means of quantum mechanical calculations, using Density Functional Theory-based molecular dynamics (MD), according to the Car-Parrinello method.⁶ All systems were enclosed in an isolated cubic box of 12.0 \AA x 12.0 \AA x 12.0 \AA , using a fictitious electron mass of 700 au and a time step of 0.12 fs. The Kohn-Sham orbitals were expanded in a plane wave (PW) basis set with a kinetic energy cutoff of 70 Ry. Ab initio pseudopotentials generated within the Troullier-Martins scheme were employed.⁷ The Perdew, Burke and Ernzerhoff generalized gradient-corrected approximation (PBE)⁸ was selected in view of its good performance in previous work on isolated sugars,^{9, 10} glycosidases¹¹ and glycosyltransferases.¹²

The metadynamics algorithm^{13, 14} was used to explore the conformational free energy landscape of all systems, taking as collective variables θ and ϕ of the puckering coordinates of Cremer and Pople.¹⁵ Initially, the height of these Gaussian terms was set at 0.6 kcal $\cdot\text{mol}^{-1}$ and a new Gaussian-like potential was added every 30 fs. Once the whole free energy space was explored, the height of the Gaussian terms was reduced to the half (0.3 kcal $\cdot\text{mol}^{-1}$) and a new Gaussian-like potential was added every 60 fs. The width of the Gaussian terms was set to 0.10 \AA . The free energy landscape of each molecule was fully explored in less than 60 ps and the simulations were further extended up to 140 ps, when convergence was achieved according to the invariance of the energy differences between the principal wells (standard deviation < 1 kcal $\cdot\text{mol}^{-1}$, considering the last 20 ps of simulation).

2.2. Crystallographic data collection and refinement

Agd31B wt and D412N mutant

Wild type *CjAgd31B* expression and purification was carried out as previously described.⁵ Expression of plasmid for the *CjAgd31B* D412N nucleophile mutant was obtained by site directed mutagenesis of the wild type *CjAgd31B* plasmid using a PCR based method¹⁶ with primers Fwd: 5'-GGTGGGGCAATCTGGGTGAACCGGAAATG-3' and Rev: 5'-TCACCCAGATTGCCCCACCAGCCCGCGAC-3'. Expression and purification of *CjAgd31B* D412N mutant were carried out as for the wild type enzyme.

Crystals for both wild type and mutant *CjAgd31B* were obtained by the sitting drop vapor diffusion method at 20 °C using 1.8 M ammonium sulfate, 0.1 M HEPES pH 7.0, 2% PEG400. Reacted complexes with **5** were obtained by soaking wt *CjAgd31B* crystals with 1 mM **5** in cryoprotectant solution (2.0 M lithium sulfate, 0.1 M HEPES pH 7.0, 2% PEG400) for 2 h at 20 °C, before flash freezing in liquid N₂ for data collection. Complexes with **2** were obtained by soaking wt *CjAgd31B* with **2** (at 5 mM) in cryoprotectant solution for 5-10 min at room temperature, followed by flash freezing in liquid N₂ for data collection. Complexes of unreacted **5** and **4** were obtained by soaking *CjAgd31B* D412N crystals with **5** or **4** (at 5 mM) in cryoprotectant solution for 20 min at room temperature, followed by flash freezing in liquid N₂ for data collection.

Data were collected at beamlines I02 and I04 of the Diamond light source, UK. Reflections were autoprocessed with the xia2 pipeline¹⁷ of the CCP4 software suite, or manually processed using XDS¹⁸ and Aimless¹⁹. Complexes were solved by molecular replacement with MolRep²⁰ using *apo CjAgd31B* (4B9Y) as a search model, followed by subsequent rounds of manual model building and refinement using Coot²¹ and REFMAC5²² respectively. For mutant *CjAgd31B* complexes with unreacted **4** and **5**, ligands were modeled at 0.7 occupancy due to weaker ligand binding to the mutant enzyme. For the mutant *CjAgd31B* complex with unreacted **4**, the displayed F_c-F_o map was calculated by refinement without a bulk solvent mask, which improved difference density for the ligand. All ligand coordinates were built using jLigand²³. Ligand ring conformations were analyzed using Privateer²⁴. Crystal structure figures were generated using Pymol.

TxGH116

TxGH116 expression and purification was carried out as previously described⁴.

Crystals of TxGH116 were obtained by the sitting drop vapor diffusion method at 20 °C using 0.1 M BisTris pH 5.5 - 6.7, 0.2 M ammonium sulfate, 20% PEG 3350. Unreacted complexes with **6** were obtained by soaking TxGH116 crystals with **6** (at 5 mM) in mother liquor solution for 10-20 min at 20 °C, before transferring to cryoprotectant solution (mother liquor supplemented with 25% ethylene glycol) and flash freezing in liquid N₂ for data collection. Reacted complexes with **6** were obtained by soaking TxGH116 crystals with **6** (at 5 mM) in mother liquor solution for ~24 h at 20 °C, before transferring cryoprotectant solution and flash freezing in liquid N₂ for data collection.

Data were collected at beamline I02 of the Diamond light source. Processing, model building and refinements were carried out as described above for *CjAgd31B*.

2.3. Chemical synthesis

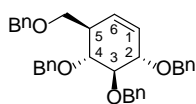
General Experimental Details

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. Triethylamine was dried over KOH and distilled before using. 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 analytical thin-layer chromatography (TLC) using Merck aluminum sheets pre-coated with silica gel 60 with detection by UV absorption (254 nm) and by spraying with a solution of $(\text{NH}_4)_6\text{Mo}_7\text{O}_{24}\cdot\text{H}_2\text{O}$ (25 g/L) and $(\text{NH}_4)_4\text{Ce}(\text{SO}_4)_4\cdot\text{H}_2\text{O}$ (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 manually using either Baker or Screening Device silica gel 60 (0.04 - 0.063 mm) or a Biotage Isolera™ flash purification system using silica gel cartridges (Screening devices SiliaSep HP, particle size 15-40 µm, 60A) in the indicated solvents. ^1H NMR and ^{13}C NMR spectra were recorded on 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. The following abbreviations are used to describe peak patterns when appropriate: s (singlet), d (doublet), t (triplet), qt (quintet), m (multiplet), br (broad), ar (aromatic), app (apparent). 2D NMR experiments (HSQC, COSY and NOESY) were carried out to assign protons and carbons of the new structures and assignment follows the general numbering shown in cyclohexene **8**. High-resolution mass spectra (HRMS) of intermediates were recorded with a LTQ Orbitrap (Thermo Finnigan) and final compounds were recorded with a apex-QE instrument (Bruker). Optical rotations were measured on a Anton Paar MCP automatic polarimeter (Sodium D-line, $\lambda = 589$ nm). LC/MS analysis was performed on an LCQ Advantage Max (Thermo Finnigan) ion-trap spectrometer (ESI+) coupled to a Surveyor HPLC system (Thermo Finnigan) equipped with a C18 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 or 50 mM NH_4HCO_3 in H_2O .

Cyclophellitol (**1**)²⁵, cyclophellitol aziridine (**2**)²⁵, 1,6-*epi*-cyclophellitol (**3**)²⁶, 1,6-*epi*-cyclophellitol aziridine (**4**)²⁶, (1*R*,2*R*,5*S*,6*S*)-5,6-bis(benzyloxy)-2-(hydroxymethyl)cyclohex-3-en-1-ol (**7**)²⁵, JJB383²⁶ (**13**), KY358 (**14**)¹, 2,4-dinitrophenyl-β-D-glucopyranoside²⁷ and 2,4-dinitrophenyl-α-D-glucopyranoside²⁸ were synthesized following procedures previously described and their spectroscopic data are in agreement with those previously reported.

Synthesis and Characterization Data of Compounds 5 and 6.

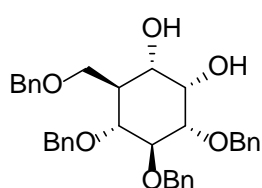
[{(1*R*,2*R*,5*S*,6*R*)-5,6-bis(benzyloxy)-2-[(benzyloxy)methyl]cyclohex-3-en-1-yl}oxy)methyl] benzene (**8**).



Catalytic amount of TBAI (54 mg, 0.15 mmol), BnBr (0.87 mL, 7.34 mmol) and sodium hydride (60% dispersion in mineral oil, 294 mg, 7.3 mmol) were added to a solution of **7** (500 mg, 1.47 mmol) in DMF (60 mL) at 0 °C and the reaction mixture was stirred at room temperature overnight. The reaction mixture was cooled to 0 °C and quenched by the addition of MeOH. The mixture was further diluted with water and subsequently extracted with EtOAc. The organic phase was washed with brine, dried over MgSO₄, filtered, and concentrated *in vacuo*. The crude was purified by silica column chromatography (from pentane to pentane/EtOAc 9:1) to afford the desired product **11** (538 mg, 70%) as a colourless oil. $[\alpha]_D^{20} = +113.6$ ($c = 1$, CHCl₃). ¹H NMR (400 MHz, CDCl₃): $\delta = 7.38 - 7.25$ (m, 18H, CH Ar.), 7.20 – 7.17 (m, 2H, CH Ar.), 5.72 (dt, $J = 10.2, 2.2$ Hz, 1H, CH=CH), 5.67 (dt, $J = 10.2, 1.8$ Hz, 1H, CH=CH), 4.95 – 4.88 (m, 3H, 3/2CH₂), 4.70 (s, 2H, CH₂), 4.48 – 4.38 (m, 3H, 3/2CH₂), 4.29 – 4.22 (m, 1H, CH-2), 3.81 (dd, $J = 10.1, 7.8$ Hz, 1H, CH-3), 3.67 (t, $J = 9.8$ Hz, 1H, CH-4), 3.52 (d, $J = 4.0$ Hz, 2H, CH₂), 2.60 – 2.46 (m, 2H, CH-5). ¹³C NMR (100 MHz, CDCl₃): $\delta = 139.0, 138.7, 138.6, 138.3$ (4C Ar.), 129.3, 128.5, 128.2, 128.1, 128.0, 127.9, 127.8, 127.7, 127.0 (20CH Ar.), 2CH=CH), 85.5 (CH-3), 81.0 (CH-2), 78.5 (CH-4), 75.5 (2CH₂), 73.2, 72.2, 69.3 (3CH₂), 44.5 (CH-5). HRMS: calcd. for [C₃₅H₃₇O₄]⁺ 521.26918; found 521.26890. HRMS: calcd. for [C₃₅H₃₆NaO₄]⁺ 543.25113; found 543.24993.

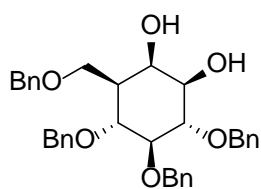
(1*S*,2*S*,3*S*,4*S*,5*R*,6*S*)-3,4,5-tris(benzyloxy)-6-[(benzyloxy)methyl]cyclohexane-1,2-diol (**9**) and (1*R*,2*R*,3*S*,4*S*,5*R*,6*S*)-3,4,5-tris(benzyloxy)-6-[(benzyloxy)methyl]cyclohexane-1,2-diol (**10**).

A solution of sodium metaperiodate (319 mg, 1.49 mmol) and catalytic amount of ruthenium trichloridetrihydrate (18.2 mg, 0.07 mmol) in water (8 mL) was added to a vigorously stirred ice-cooled solution of diene **8** (518 mg, 1.0 mmol) in EtOAc/ACN 1:1 (30 mL). The stirring was continued at 0°C for 2h. The reaction was then quenched by addition of 10% aqueous Na₂S₂O₃ solution (10 mL). The aqueous layer was separated and extracted three times with ethyl acetate (3 x 20 mL). The combined organic extracts were washed with brine, dried with MgSO₄ and concentrated under reduced pressure. The crude was purified by silica column chromatography (Pentane/EtOAc from 9:1 to 3:7) to afford **9** (215 mg, 0.39 mmol, 39%) and **10** (142 mg, 0.26 mmol, 26%).



α -cis diol 9: $[\alpha]_D^{20} = +11.0$ ($c = 1$, CHCl₃). ¹H NMR (400 MHz, CDCl₃): $\delta = 7.43 - 7.24$ (m, 20H, CH Ar.), 5.00 (d, $J = 10.8$ Hz, 1H, 1/2CH₂), 4.94 (d, $J = 10.8$ Hz, 1H, 1/2CH₂), 4.89 (d, $J = 10.8$ Hz, 1H, 1/2CH₂), 4.77 (s, 2H, CH₂), 4.61 – 4.45 (m, 3H, 3/2CH₂), 4.20 (t, $J =$

2.6 Hz, 1H, CH-1), 4.03 (t, $J = 9.4$ Hz, 1H, CH-3), 3.90 (dd, $J = 8.9, 2.6$ Hz, 1H, 1/2CH₂), 3.77 – 3.69 (m, 2H, CH-6, 1/2CH₂), 3.53 – 3.42 (m, 2H, CH-2, CH-4), 3.17 (br s, 1H, OH), 2.83 (br s, 1H, OH), 2.23 (tdd, $J = 10.9, 4.8, 2.6$ Hz, 2H, CH-5). ¹³C NMR (100 MHz, CDCl₃): $\delta = 138.9, 138.6, 138.1, 138.0$ (4C Ar.), 128.6, 128.5, 128.0, 127.8, 127.7, 127.6 (20CH Ar.), 83.0 (CH-3), 80.1 (CH-2), 77.7 (CH-4), 75.7, 75.4, 73.4, 72.6 (4CH₂), 70.4 (CH-1), 69.2 (CH-6), 67.6 (CH₂), 43.3 (CH-5). HRMS: calcd. for [C₃₅H₃₉O₆]⁺ 555.27466; found 555.27448. HRMS: calcd. for [C₃₅H₃₈NaO₆]⁺ 577.25661; found 577.25580.



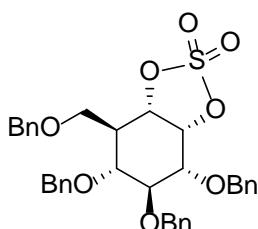
β -cis diol 10: Obtained as white crystals by recrystallization from MeOH:Et₂O. $[\alpha]_D^{20} = +16.8$ ($c = 1$, CHCl₃). ¹H NMR (400 MHz, CDCl₃): $\delta = 7.38 - 7.12$ (m, 20H, CH Ar.), 4.97 – 4.85 (m, 4H, 2CH₂), 4.79 (d, $J = 11.1$ Hz, 1H, 1/2CH₂), 4.53 – 4.42 (m, 3H, 3/2CH₂), 4.23 (t, $J = 2.5$ Hz, 1H, CH-6), 3.94 – 3.80 (m, 3H, CH-2, CH-4, 1/2CH₂), 3.71 (dd, $J = 9.0, 3.1$ Hz, 1H, 1/2CH₂), 3.61 – 3.46 (m, 2H, CH-1, CH-3), 3.40 (br s, 1H, OH), 2.55 (br s, 1H, OH), 1.78 – 1.69 (m, 1H, CH-5). ¹³C NMR (101 MHz, CDCl₃): $\delta = 138.8, 138.7, 138.4, 137.6$ (4C Ar.), 128.6, 128.5, 128.1, 128.0, 127.9, 127.8, 127.6 (20CH Ar.), 86.6 (CH-3), 82.4 (CH-2), 77.4 (CH-4), 75.7, 75.6, 75.5 (3CH₂), 74.6 (CH-1), 73.6 (CH₂), 70.9 (CH-6), 68.8 (CH₂), 43.5 (CH-5). HRMS: calcd. for [C₃₅H₃₉O₆]⁺ 555.27466; found 555.27450. HRMS: calcd. for [C₃₅H₃₈NaO₆]⁺ 577.25661; found 577.25587.

General Procedure for the Synthesis of Cyclosulfates.

Thionyl chloride (3.5 equiv. for cis and 7 equiv. for trans diol) was added over 5 min to a solution of diol (1 equiv.) and triethylamine (4 equiv. for cis and 8 equiv. for trans diol) in DCM (50 mL/mmol) at 0 °C. The reaction mixture was diluted with cold diethyl ether and washed with cold water and brine. The organic phase was dried (MgSO₄), filtered, concentrated under reduced pressure, and the residual triethylamine was removed under high vacuum (1 h).

The resulting oil was dissolved in CCl₄ (40 mL/mmol) and ACN (40 mL/mmol), and the solution was cooled to 0 °C in an ice-bath. A solution of catalytic amount of RuCl₃·3H₂O (0.1 equiv.) and NaIO₄ (2 equiv.) in water (40 mL/mmol) was added and the reaction mixture was stirred at 0 °C for 3 h. Diethyl ether was added and the two layers were separated. The aqueous phase was extracted again with diethyl ether and the combined organic extracts were washed with brine and dried over MgSO₄. The crude was concentrated under reduced pressure and purified by silica column chromatography (from Pentane to Pentane/EtOAc to 8:2) to afford the desired intermediates.

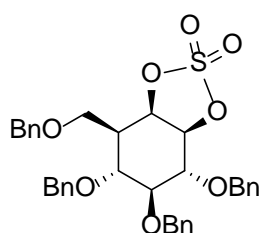
(3*a*R,4*R*,5*S*,6*R*,7*R*,7*a*S)-4,5,6-tris(benzyloxy)-7-[(benzyloxy)methyl] hexahydro-1,3,2-benzodioxathiole 2,2-dioxide (**11**).



Obtained as a colorless oil from **9** (190 mg, 0.34 mmol) in 59% yield (125 mg, 0.20 mmol). $[\alpha]_D^{20} = -0.6$ ($c = 1$, CHCl₃). ¹H NMR (400 MHz, CDCl₃): $\delta = 7.38 - 7.23$ (m, 18H, CH Ar.), 7.21 – 7.14 (m, 2H, CH Ar.), 5.12 – 5.04 (m, 2H, CH-1, CH-6), 4.83 – 4.67 (m, 5H,

5/2CH₂), 4.51 – 4.36 (m, 3H, 3/2CH₂), 3.91 (t, *J* = 8.0 Hz, 1H, CH-2), 3.87 (dd, *J* = 9.5, 2.2 Hz, 1H, 1/2CH₂), 3.71 (dd, *J* = 7.8, 3.2 Hz, 1H, CH-3), 3.60 – 3.49 (m, 2H, CH-4, 1/2CH₂), 2.57 – 2.47 (m, 1H, CH-5). ¹³C NMR (101 MHz, CDCl₃): δ = 138.0, 137.9, 137.6, 137.0 (4C Ar.), 128.8, 128.6, 128.5, 128.4, 128.2, 128.1, 128.0, 127.9, 127.8 (20CH Ar.), 81.8 (CH-2), 81.0, 80.1 (CH-1, CH-6), 75.7 (CH-3), 75.4 (CH-4), 75.2, 75.0, 73.7, 73.4, 64.1 (5CH₂), 43.4 (CH-5). HRMS: calcd. for [C₃₅H₃₆NaO₈S]⁺ 639.20286; found 639.20186.

(3*a*S,4*R*,5*S*,6*R*,7*R*,7*a*R)-4,5,6-tris(benzyloxy)-7-[(benzyloxy)methyl] hexahydro-1,3,2-benzodioxathiole 2,2-dioxide (**12**).

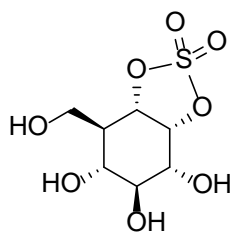


Obtained as a colorless oil from **10** (122 mg, 0.22 mmol) in 62% yield (84 mg, 0.14 mmol). [α]_D²⁰ = +10.8 (*c* = 1, CHCl₃). ¹H NMR (400 MHz, CDCl₃): δ = 7.36 – 7.26 (m, 18H, CH Ar.), 7.17 – 7.13 (m, 2H, CH Ar.), 5.41 (dd, *J* = 4.7, 3.3 Hz, 1H, CH-6), 4.90 – 4.76 (m, 6H, CH-1, 5/2CH₂), 4.50 – 4.42 (m, 3H, 3/2CH₂), 4.19 (dd, *J* = 9.5, 8.1 Hz, 1H, CH-2), 3.79 (dd, *J* = 8.8, 4.4 Hz, 1H, 1/2CH₂), 3.61 – 3.47 (m, 2H, CH-3, CH-4), 3.39 (dd, *J* = 10.1, 9.0 Hz, 1H, 1/2CH₂), 2.23 (ddt, *J* = 10.5, 7.6, 3.8 Hz, 1H, CH-5). ¹³C NMR (101 MHz, CDCl₃): δ = 138.0, 137.6, 137.5, 137.4 (4C Ar.), 128.7, 128.6, 128.5, 128.3, 128.2, 128.1, 128.0, 127.9, 127.8 (20CH Ar.), 87.1 (CH-1), 83.6 (CH-3), 81.4 (CH-6), 80.6 (CH-2), 75.9 (CH-4), 75.9 (CH₂), 75.8 (2CH₂), 73.6, 66.2 (2CH₂), 42.1 (CH-5). HRMS: calcd. for [C₃₅H₃₆NaO₈S]⁺ 639.20286; found 639.20191.

General Procedure for the Deprotection of Cyclosulfates.

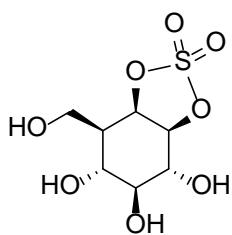
10% Palladium on carbon (0.4 equiv.) was added to a solution of intermediates **11** and **12** (1 equiv.) in MeOH (50 mL/mmol) under argon atmosphere. The mixture was hydrogenated with a hydrogen balloon and stirred under hydrogen atmosphere for 18 h. Then, the reaction mixture was filtered through a celite plug and the solvent was evaporated under reduced pressure. The crude was purified by chromatography (from DCM to DCM/MeOH 9:1) to afford the desired final products **5** and **6**.

(3*a*R,4*R*,5*S*,6*R*,7*R*,7*a*S)-4,5,6-trihydroxy-7-(hydroxymethyl)hexahydro-1,3,2-benzodioxathiole 2,2-dioxide (**5**).



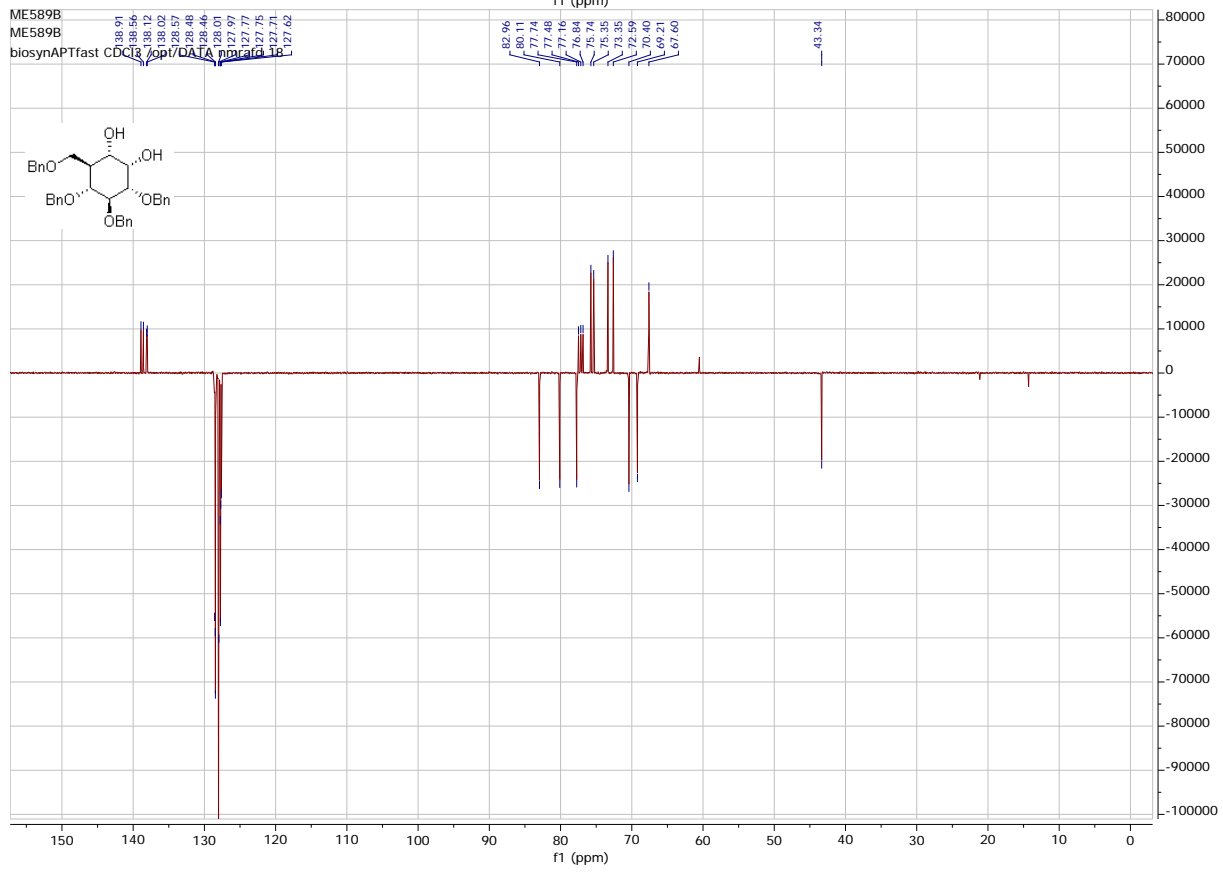
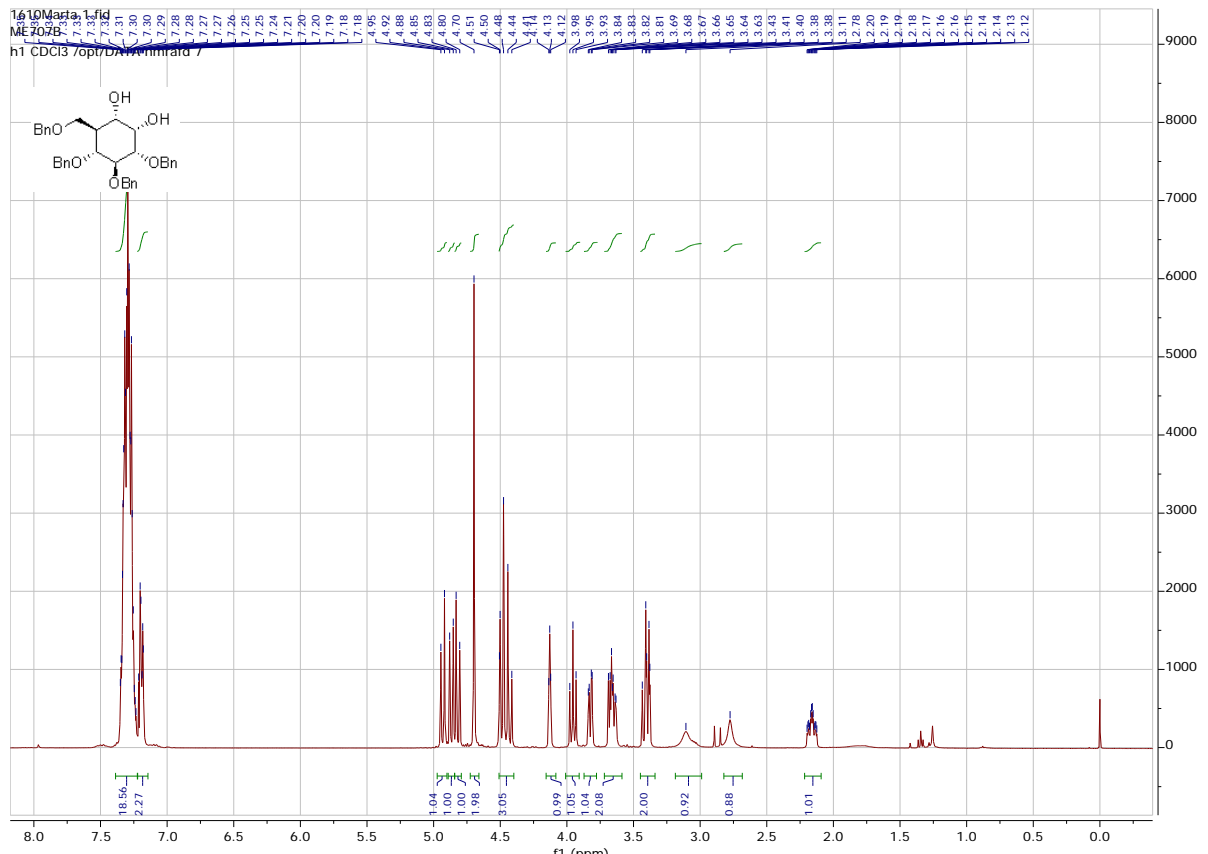
Obtained as a white solid from **11** (95 mg, 0.15 mmol) in 71% yield (28 mg, 0.11 mmol). [α]_D²⁰ = -22.5 (*c* = 0.2, CH₃OH). ¹H NMR (400 MHz, Methanol-*d*₄): δ 5.29 (t, *J* = 3.9 Hz, 1H, CH-1), 5.18 (dd, *J* = 10.2, 4.4 Hz, 1H, CH-6), 4.03 (dd, *J* = 11.2, 2.3 Hz, 1H, 1/2CH₂), 3.72 – 3.62 (m, 2H, CH-2, 1/2CH₂), 3.58 (t, *J* = 9.5 Hz, 1H, CH-4), 3.39 – 3.33 (m, 1H, CH-3 and MeOD), 2.10 (app ddt, *J* = 11.2, 10.6, 2.5 Hz, 1H, CH-5). ¹³C NMR (101 MHz, Methanol-*d*₄): δ = 86.2 (CH-1), 83.0 (CH-6), 74.7 (CH-4), 70.3 (CH-2), 68.7 (CH-3), 57.1 (CH₂), 46.7 (CH-5). HRMS: calcd. for [M-H]⁻ [C₇H₁₁O₈S]⁻ 255.01801; found 255.01659.

3a*S*,4*R*,5*S*,6*R*,7*R*,7a*R*)-4,5,6-trihydroxy-7-(hydroxymethyl)hexahydro-1,3,2-benzodioxathiole 2,2-dioxide (**6**).

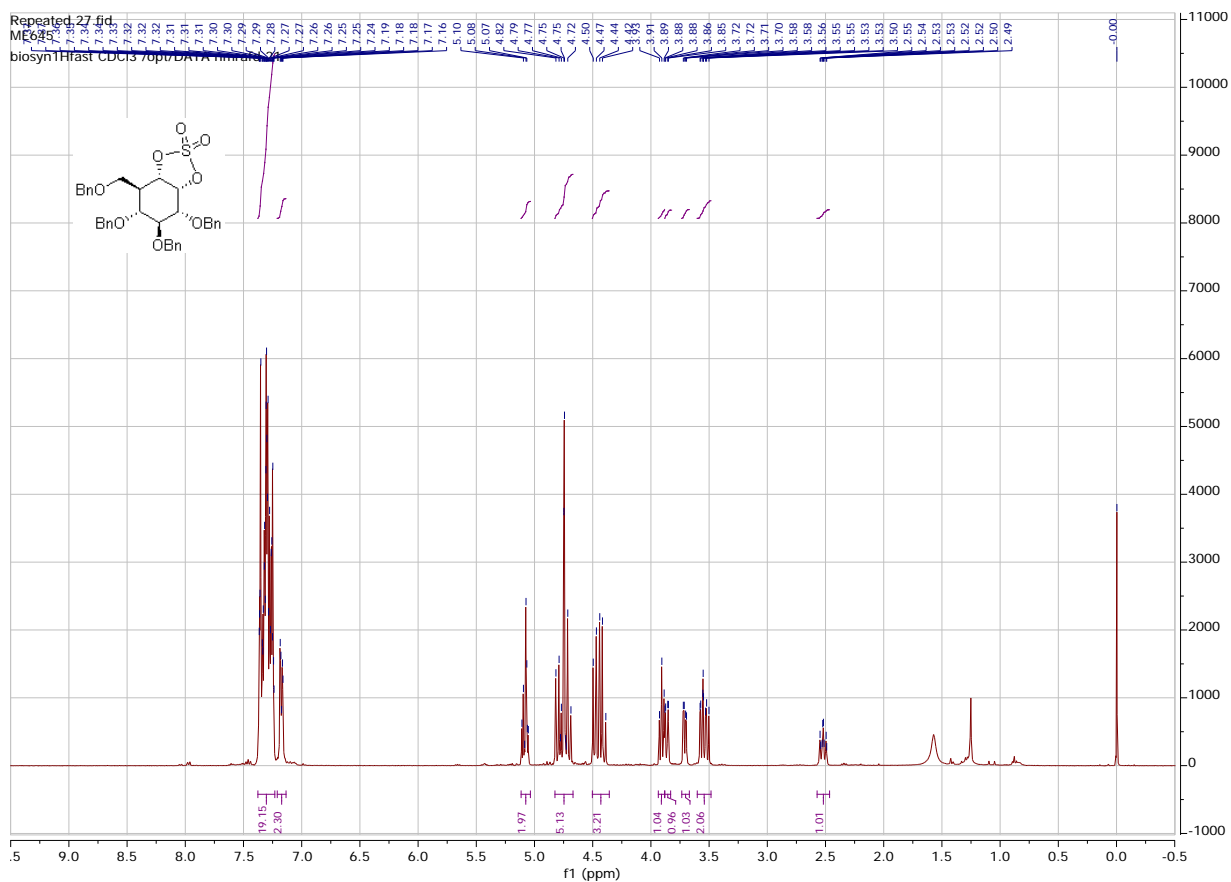


Obtained as a white solid from **12** (28 mg, 0.05 mmol) in 72% yield (9.2 mg, 0.036 mmol). $[\alpha]_{\text{D}}^{20} = -10.5$ ($c = 0.2$, CH_3OH). ^1H NMR (400 MHz, Methanol- d_4): $\delta = 5.42$ (dd, $J = 4.7, 3.4$ Hz, 1H, CH-6), 4.90 – 4.86 (m, 1H, CH-1 and H_2O), 4.06 (dd, $J = 10.8, 4.8$ Hz, 1H, $1/2\text{CH}_2$), 3.91 (dd, $J = 10.0, 8.5$ Hz, 1H, CH-2), 3.63 (t, $J = 10.4$ Hz, 1H, $1/2\text{CH}_2$), 3.39 (dd, $J = 11.1, 9.3$ Hz, 1H, CH-4), 3.24 (t, $J = 9.7$ Hz, 1H, CH-3), 2.13 – 2.02 (m, 1H, CH-5). ^{13}C NMR (101 MHz, Methanol- d_4): $\delta = 89.7$ (CH-1), 83.6 (CH-6), 76.3 (CH-3), 73.7 (CH-2), 69.9 (CH-4), 60.2 (CH_2), 45.2 (CH-5). HRMS: calcd. for $[\text{M-H}]^- [\text{C}_7\text{H}_{11}\text{O}_8\text{S}]$ 255.01801; found 255.01728.

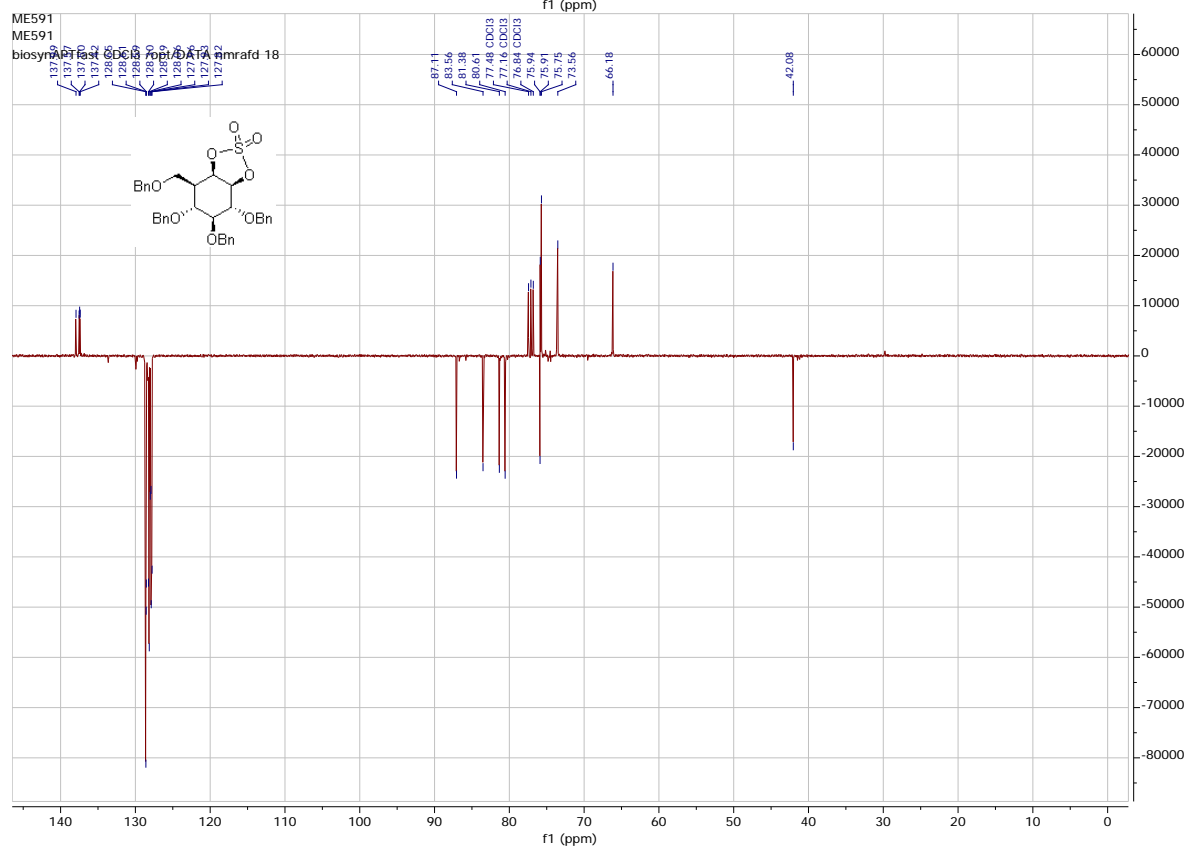
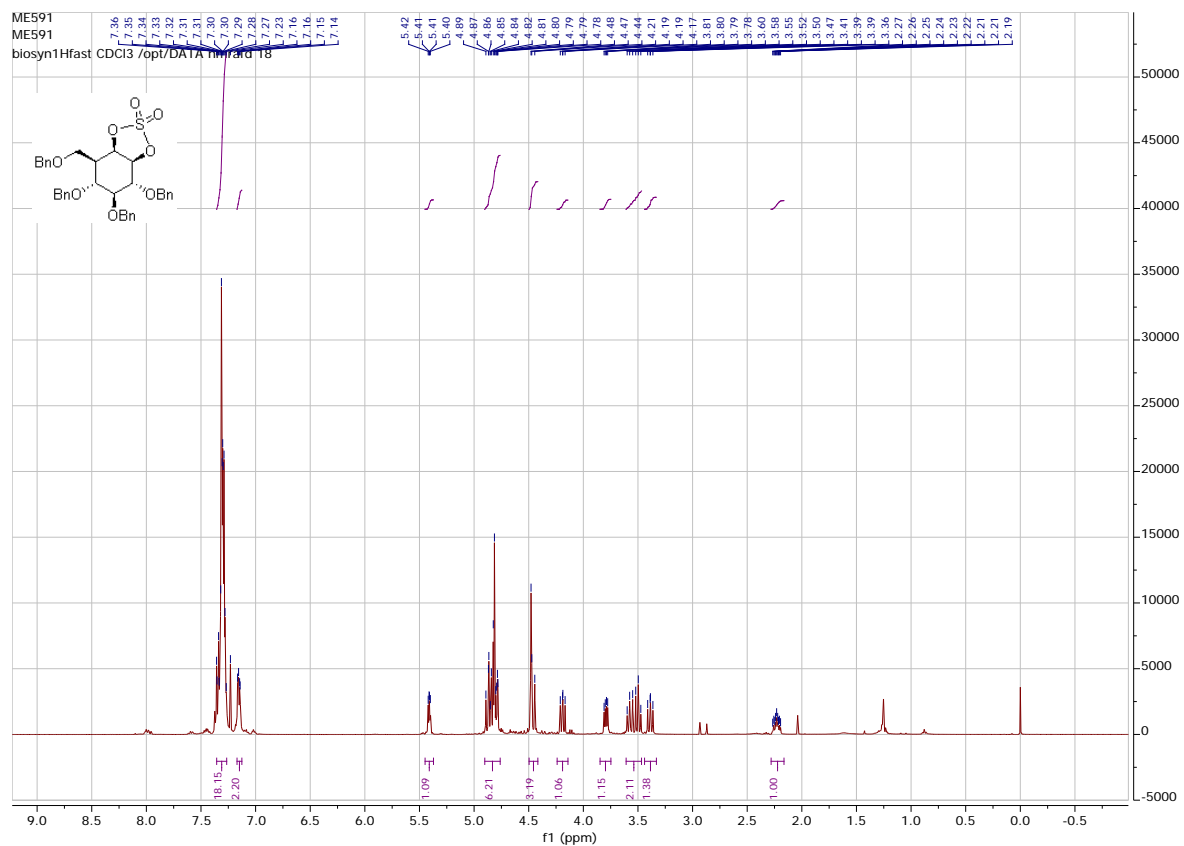
¹H-NMR and ¹³C-NMR spectra of **9** in CDCl₃



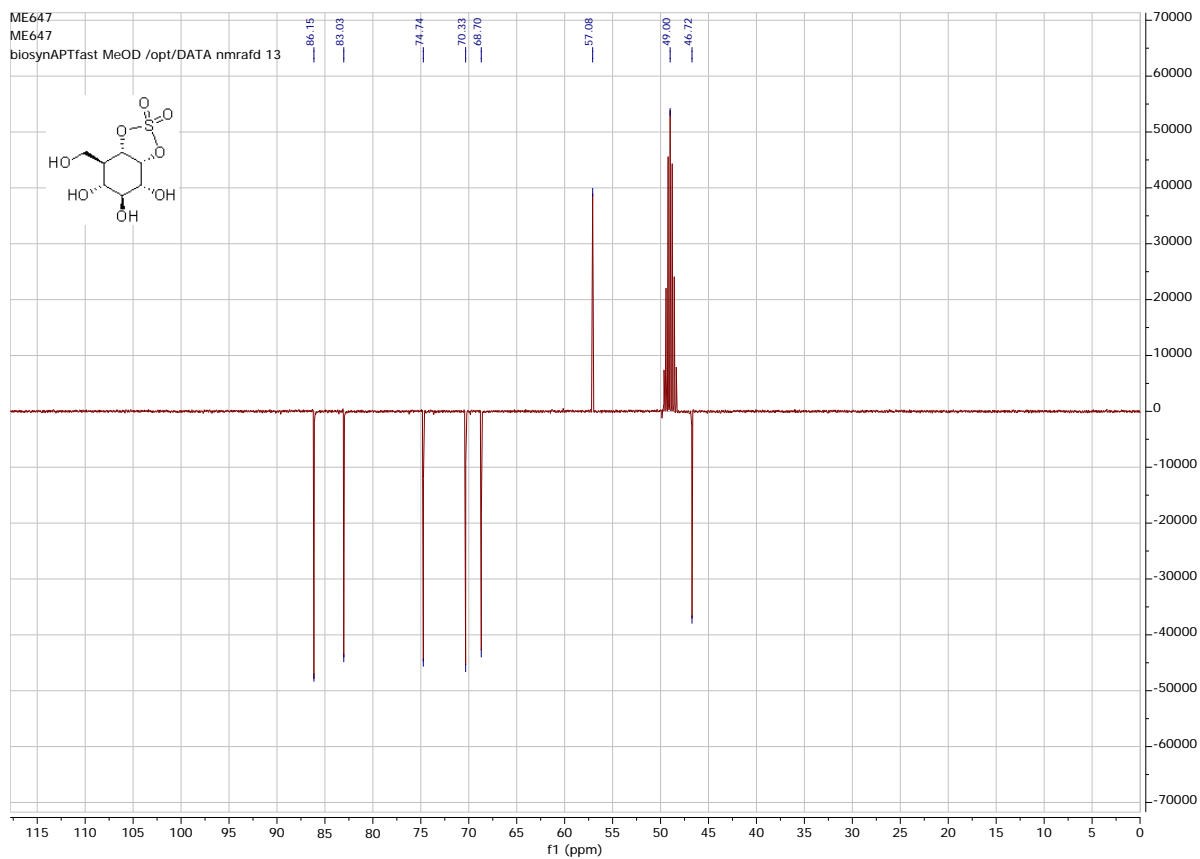
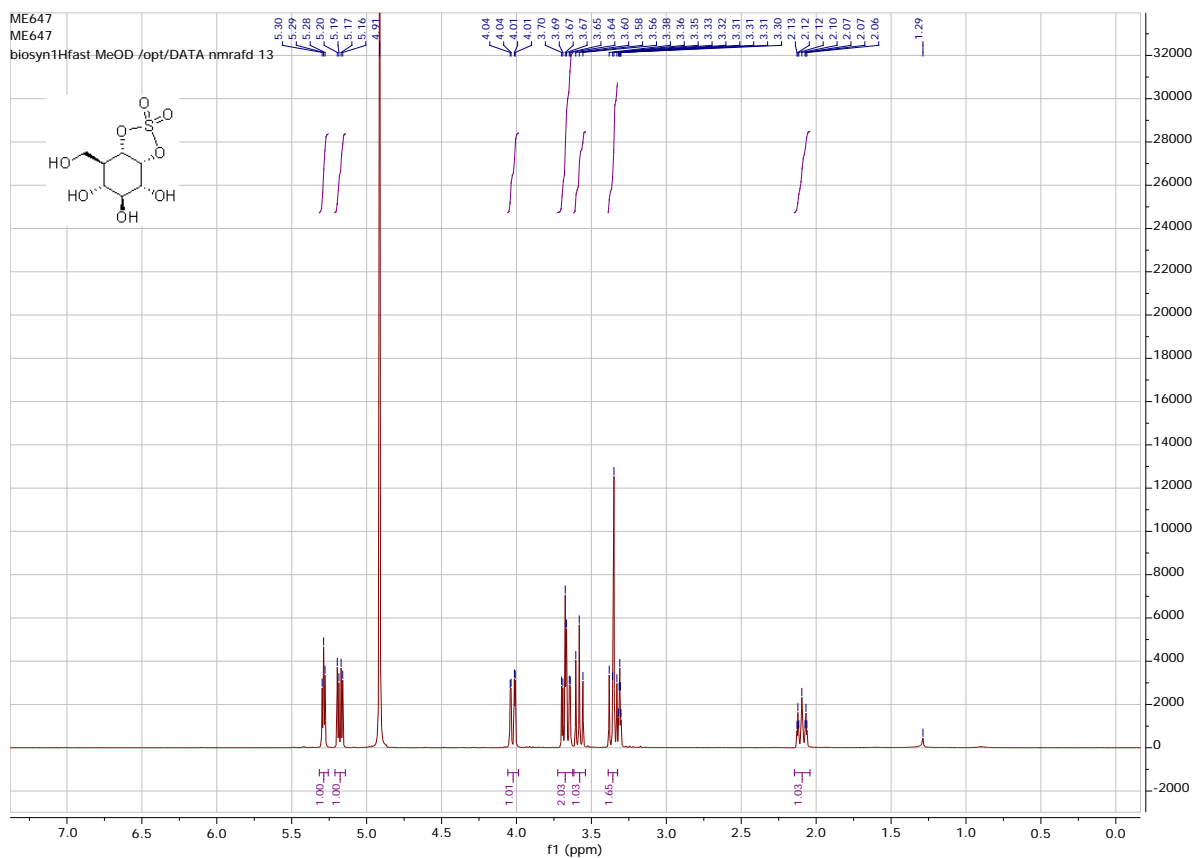
¹H-NMR and ¹³C-NMR spectra of **11** in CDCl₃



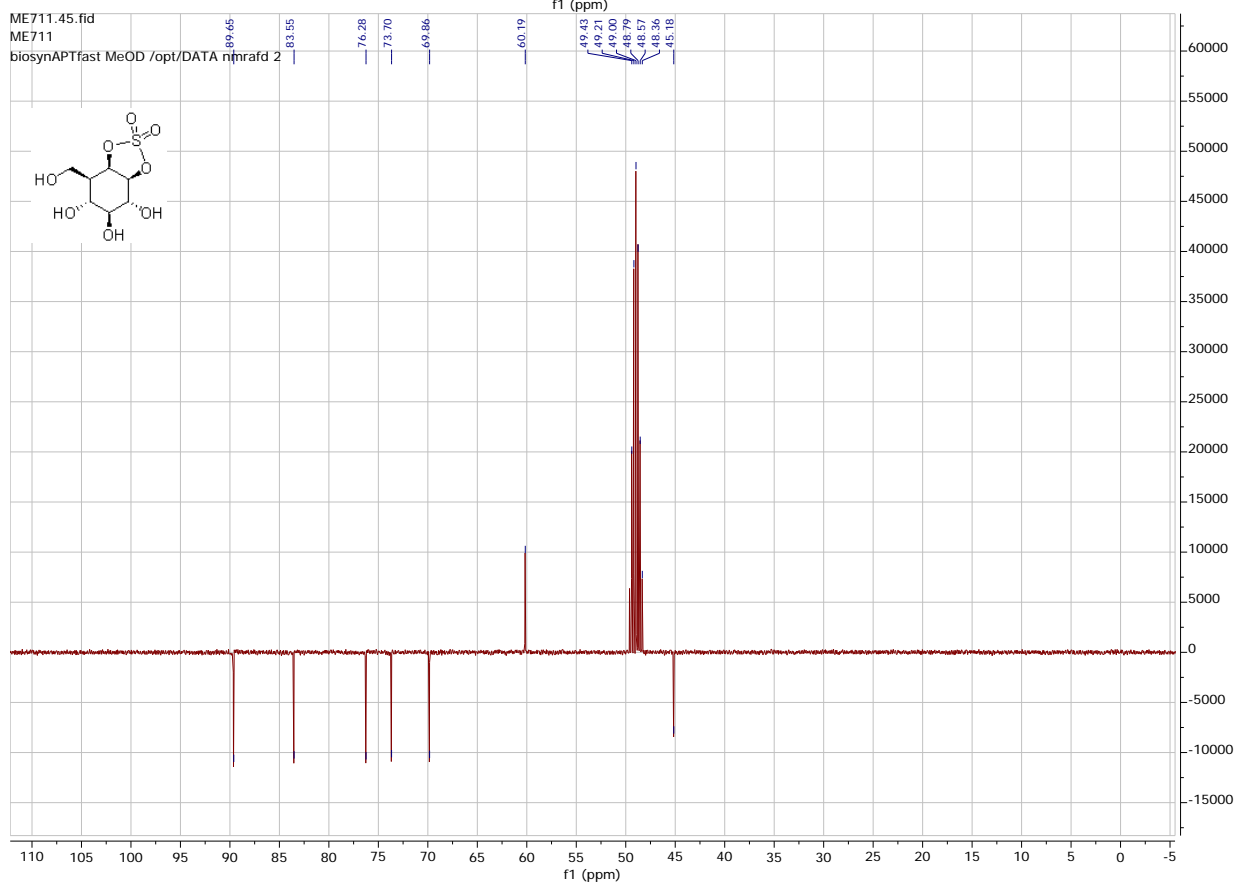
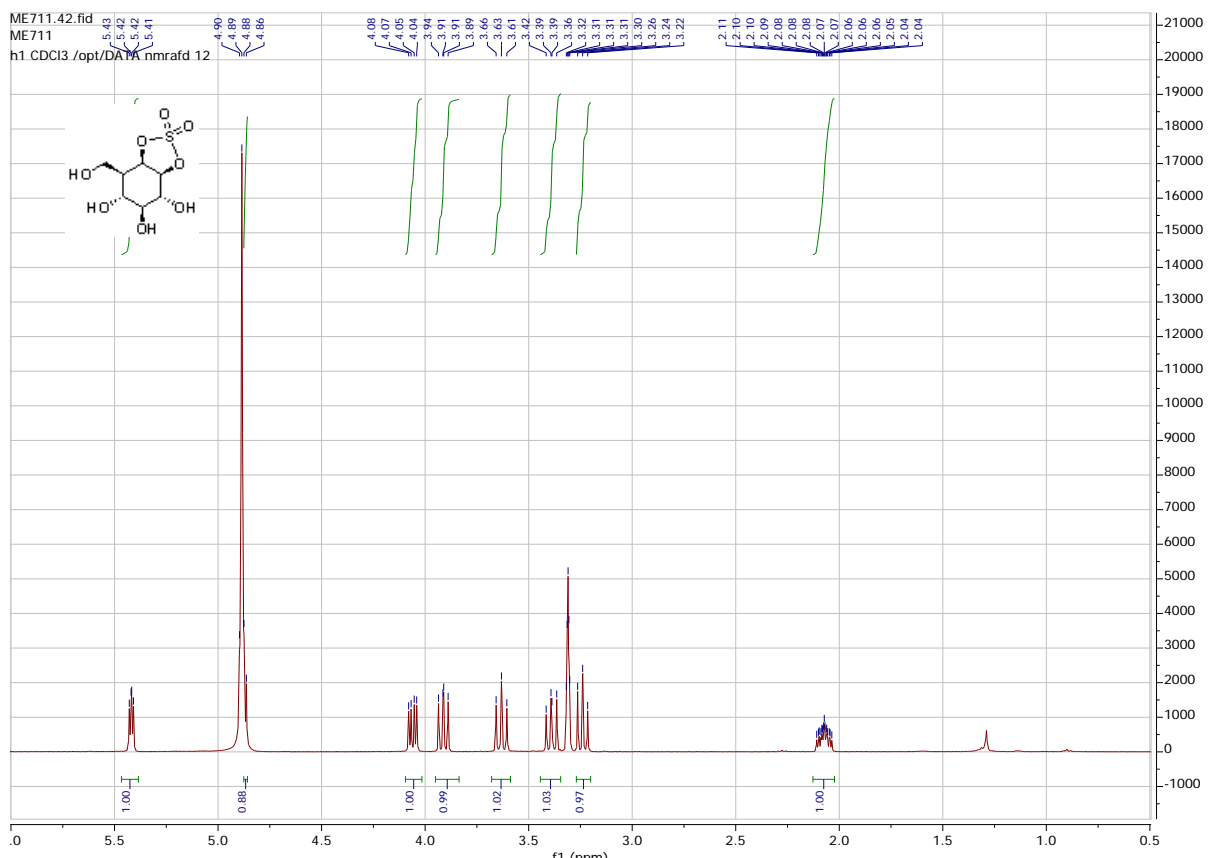
$^1\text{H-NMR}$ and $^{13}\text{C-NMR}$ spectra of **12** in CDCl_3



^1H -NMR and ^{13}C -NMR spectra of **5** in CD_3OD



^1H -NMR and ^{13}C -NMR spectra of **6** in CD_3OD



References

1. Li, K. Y.; Jiang, J.; Witte, M. D.; Kallemeijn, W. W.; Donker-Koopman, W. E.; Boot, R. G.; Aerts, J. M.; Codee, J. D.; van der Marel, G. A.; Overkleeft, H. S., Exploring functional cyclophellitol analogues as human retaining beta-glucosidase inhibitors. *Org. Biomol. Chem.* **2014**, *12*, 7786-7791.
2. Witte, M. D.; Kallemeijn, W. W.; Aten, J.; Li, K. Y.; Strijland, A.; Donker-Koopman, W. E.; van den Nieuwendijk, A. M.; Bleijlevens, B.; Kramer, G.; Florea, B. I.; Hooibrink, B.; Hollak, C. E.; Ottenhoff, R.; Boot, R. G.; van der Marel, G. A.; Overkleeft, H. S.; Aerts, J. M., Ultrasensitive in situ visualization of active glucocerebrosidase molecules. *Nat. Chem. Biol.* **2010**, *6*, 907-913.
3. Gloster, T. M.; Meloncelli, P.; Stick, R. V.; Zechel, D.; Vasella, A.; Davies, G. J., Glycosidase inhibition: An assessment of the binding of 18 putative transition-state mimics. *J. Am. Chem. Soc.* **2007**, *129*, 2345-2354.
4. Charoenwattanasatien, R.; Pengthaisong, S.; Breen, I.; Mutoh, R.; Sansenya, S.; Hua, Y.; Tankrathok, A.; Wu, L.; Songsiririthigul, C.; Tanaka, H.; Williams, S. J.; Davies, G. J.; Kurisu, G.; Cairns, J. R., Bacterial Beta-Glucosidase Reveals the Structural and Functional Basis of Genetic Defects in Human Glucocerebrosidase 2 (GBA2). *ACS Chem. Biol.* **2016**, *11*, 1891-1900.
5. 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.
6. Car, R.; Parrinello, M., Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Lett.* **1985**, *55*, 2471-2474.
7. Troullier, N.; Martins, J. L., Efficient pseudopotentials for plane-wave calculations. *Phys. Rev. B* **1991**, *43*, 1993-2006.
8. Perdew, J. P.; Burke, K.; Ernzerhof, M., Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865-3868.
9. Biarnés, X.; Ardèvol, A.; Planas, A.; Rovira, C.; Laio, A.; Parrinello, M., The conformational free energy landscape of beta-D-glucopyranose. Implications for substrate preactivation in beta-glucoside hydrolases. *J. Am. Chem. Soc.* **2007**, *129*, 10686-10693.
10. Petricevic, M.; Sobala, L. F.; Fernandes, P. Z.; Raich, L.; Thompson, A. J.; Bernardo-Seisdedos, G.; Millet, O.; Zhu, S.; Sollogoub, M.; Jimenez-Barbero, J.; Rovira, C.; Davies, G. J.; Williams, S. J., Contribution of Shape and Charge to the Inhibition of a Family GH99 endo-alpha-1,2-Mannanase. *J. Am. Chem. Soc.* **2017**, *139*, 1089-1097.
11. Jin, Y.; Petricevic, M.; John, A.; Raich, L.; Jenkins, H.; Souza, L. P. D.; Cuskin, F.; Gilbert, H. J.; Rovira, C.; Goddard-Borger, E. D.; Williams, S. J.; Davies, G. J., A b-mannanase with a lysozyme fold and a novel molecular catalytic mechanism *ACS Central Sci.* **2016**, *2*, 896-903.
12. Ardevol, A.; Rovira, C., Reaction Mechanisms in Carbohydrate-Active Enzymes: Glycoside Hydrolases and Glycosyltransferases. Insights from ab Initio Quantum Mechanics/Molecular Mechanics Dynamic Simulations. *J. Am. Chem. Soc.* **2015**, *137*, 7528-7547.
13. Laio, A.; Parrinello, M., Escaping free-energy minima. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 12562-12566.
14. Barducci, A.; Bonomi, M.; Parrinello, M., Metadynamics. *WIREs Comput. Mol. Sci.* **2011**, *1*, 826-843.
15. Cremer, D.; Pople, J. A., General definition of ring puckering coordinates *J. Am. Chem. Soc.* **1975**, *97*, 1354-1358.
16. Zheng, L.; Baumann, U.; Reymond, J. L., An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic Acids Res.* **2004**, *32*, e115.
17. Winter, G., xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* **2010**, *43*, 186-190.
18. Kabsch, W., XDS. *Acta Crystallogr. D* **2010**, *66*, 125-132.
19. Evans, P. R.; Murshudov, G. N., How good are my data and what is the resolution? *Acta Crystallogr. D* **2013**, *69*, 1204-1214.
20. Vagin, A.; Teplyakov, A., Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 22-25.

21. Emsley, P.; Cowtan, K., Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2126-2132.
22. Murshudov, G. N.; Skubak, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A., REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 355-367.
23. 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 Biol. Crystallogr.* **2012**, *68*, 431-440.
24. Agirre, J.; Iglesias-Fernandez, J.; Rovira, C.; Davies, G. J.; Wilson, K. S.; Cowtan, K. D., Privateer: software for the conformational validation of carbohydrate structures. *Nat. Struct. Mol. Biol.* **2015**, *22*, 833-834.
25. Li, K.-Y.; Jiang, J.; Witte, M. D.; Kallemeijn, W. W.; Elst, H. v. d.; Chung-Sing Wong; Chander, S. D.; Hoogendoorn, S.; Beenakker, T. J. M.; Codée, J. D. C.; Aerts, J. M. F. G.; Marel, G. A. v. d.; Overkleef, H. S., Synthesis of Cyclophellitol, Cyclophellitol Aziridine, and Their Tagged Derivatives. *Eur. J. Org. Chem.* **2014**, 6030–6043.
26. Speciale, G.; Thompson, A. J.; Davies, G. J.; Williams, S. J., Dissecting conformational contributions to glycosidase catalysis and inhibition. *Curr. Opin. Struct. Biol.* **2014**, *28*, 1-13.
27. Chen, H. M.; Withers, S. G., Facile synthesis of 2,4-dinitrophenyl alpha-D-glycopyranosides as chromogenic substrates for alpha-glycosidases. *Chembiochem* **2007**, *8*, 719-722.
28. Sharma, S. K.; Corrales, G.; Penades, S., Single-Step Stereoselective Synthesis of Unprotected 2,4-Dinitrophenyl Glycosides. *Tetrahedron Lett.* **1995**, *36*, 5627-5630.