SUPPORTING INFORMATION

Capture-and-release of a sulfoquinovose-binding protein on sulfoquinovose-modified agarose

Thimali Arumapperuma,[a] Alexander J. D. Snow,[b] Mihwa Lee,[a] Mahima Sharma,[b] Yunyang Zhang,[a] James P. Lingford,[c] Ethan D. Goddard-Borger,[c,d]* Gideon J. Davies,[b]* Spencer J. Williams[a]*

[a] School of Chemistry and Bio21 Molecular Science and Biotechnology Institute, University of Melbourne, Parkville, Victoria 3010, Australia

[b] York Structural Biology Laboratory, Department of Chemistry, University of York, Heslington YO10 5DD, United Kingdom

[c] Chemical Biology Division, The Walter and Eliza Hall Institute of Medical Research, Parkville, Victoria 3010, Australia

[d] Department of Medical Biology, University of Melbourne, Parkville, Victoria 3010, Australia

Table of Contents

1. Supplementary Figures

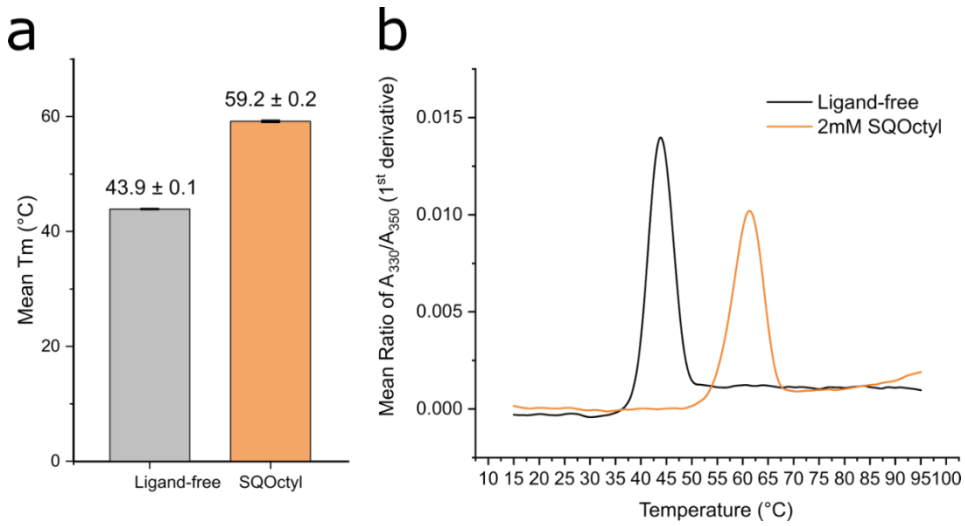


Figure S1. Stabilization of SmoF from *A. tumefaciens* in the presence of octyl-SQ. a) Bar plot showing shift in melting temperature (Tm) of SmoF when ligand-free (grey) was incubated with 2 mM octyl-SQ ligand (coral) using nanoDSF. b) First derivative of the thermal unfolding profile of SmoF under ligand free conditions (black) and in the presence of 2 mM octyl-SQ (coral).

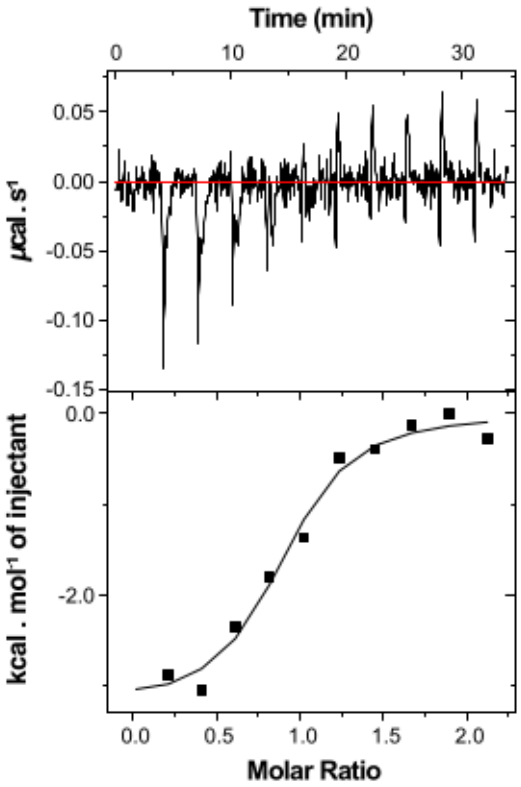


Figure S2. Isothermal titration calorimetry of octyl-SQ titrated into SmoF from *Agrobacterium tumefaciens*. SmoF was at 30 µM.

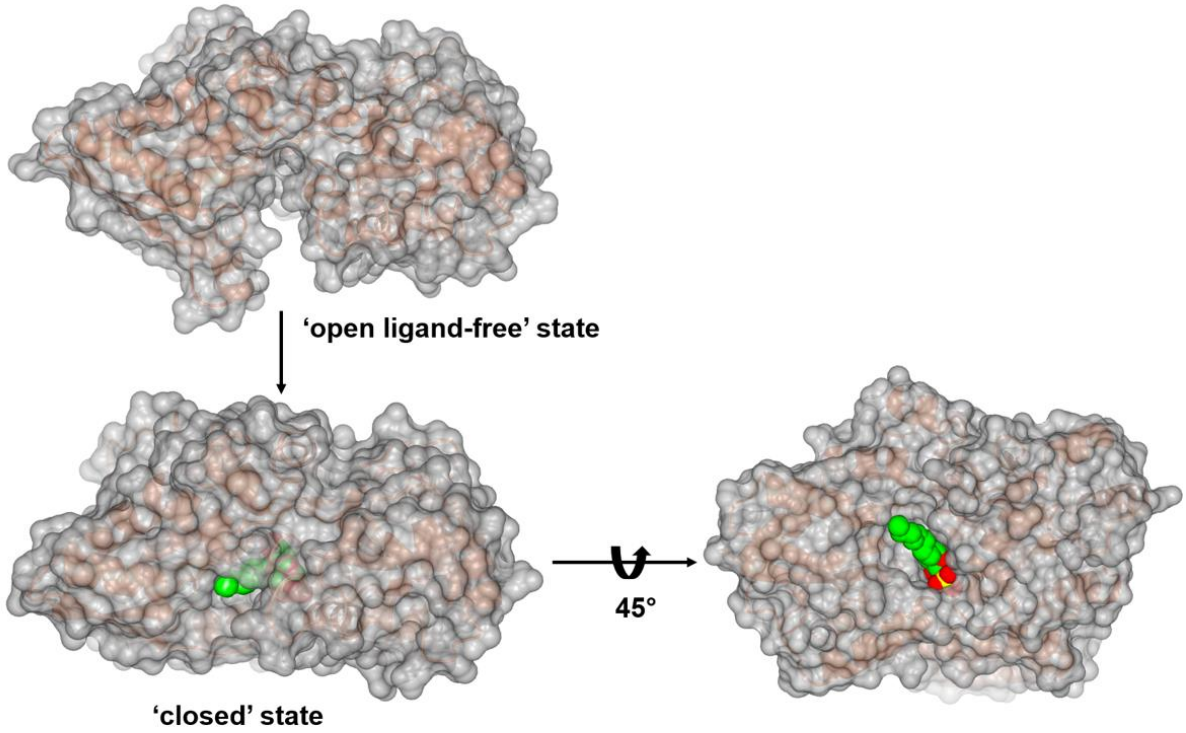


Figure S3. Conformational change of SmoF upon binding octyl-SQ. Comparison of ligand-free 'open' (7NBZ.pdb; top) and octyl-SQ bound 'closed' (this work; bottom left) states of SmoF upon octyl-SQ binding between the two domains. The long alkyl chain of octyl-SQ protrudes out from binding cleft (bottom right; SQBP is rotated 45 deg for clarity). SQBP is shown in ribbon and protein surface representation and octyl-SQ is depicted in standard CPK representation.

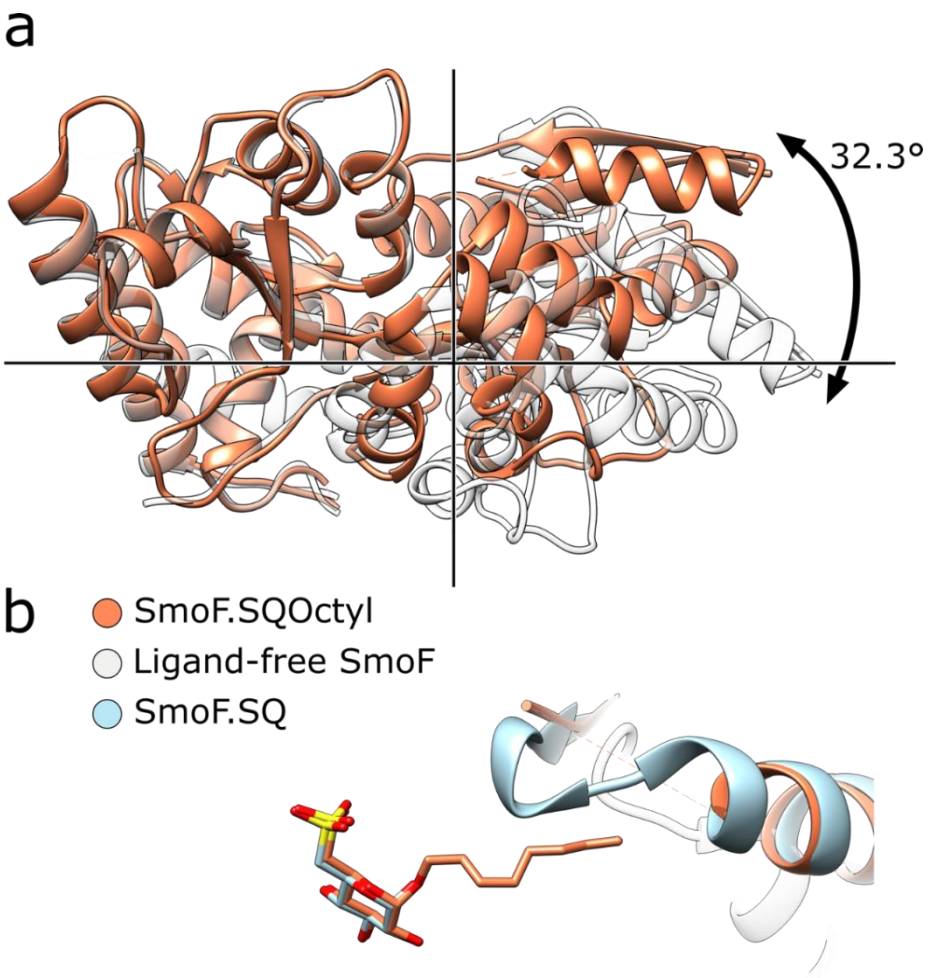


Figure S4. Further details of changes to the SmoF structure upon binding octyl-SQ. a) Domain movement of SmoF upon binding octyl-SQ. An inter-domain rotation of 32° is observed between "open" ligand-free SmoF (grey; pdb 7NBZ) and "closed" SmoF.octyl-SQ (coral). Crossed lines denote the center of rotation and the hinge axes perpendicular to motion. b) Deflection and loss of order in residues 9-19 in SmoF.octyl-SQ compared to SmoF•SQ (blue) and ligand-free SmoF (grey).

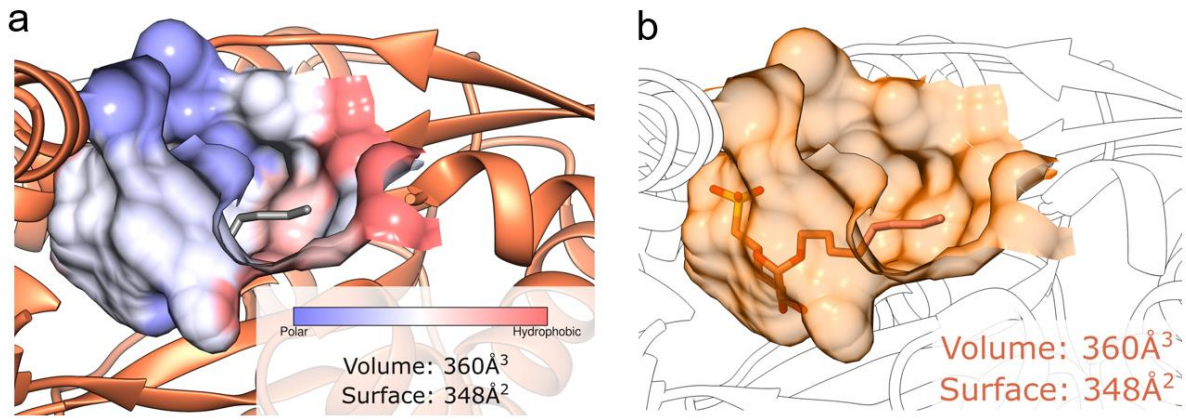


Figure S5. Volume and surface of SmoF ligand binding pocket for octyl-SQ visualized using CASTp server. (**a**) Surface of the binding pocket octyl-SQ **3** is depicted as electrostatic potential with ligand **3** shown in grey (cylinder format) protruding out of the pocket. (**b**) Surface of the binding pocket octyl-SQ **3** is depicted at a transparency of 0.5 with ligand **3** shown as cylinders in coral.

2. Supplementary Table

Table S1. X-ray data collection and refinement statistics. Numbers in brackets refer to data for highest resolution shells. All values are within *I/σI* where applicable.

	SmoF•octyl-SQ (3) complex
Data collection	
Space group	P 21 21 21
Molecules in A.S.U	1
Cell dimensions	
a, *b*, *c* (Å)	53.78, 65.85, 107.0
α, β, ɣ (°)	90.0, 90.0, 90.0
Resolution (Å)	56.1-1.80 (1.84-1.80)
R_{merge}	0.116 (0.470)
R_{pim}	0.065 (0.262)
$I / \sigma I$	10.1 (3.3)
CC1/2	0.965 (0.952)
Completeness (%)	100 (100)
Redundancy	8.0 (7.9)
No. unique reflections	36036 (2098)
Refinement	
Resolution (Å)	1.80
R_{work} / R_{free}	0.23/0.27
No. atoms	
Protein	2850
Ligand/ion	23
Water	243
B-factors (Å^2)	
Protein	21
Ligand/ion	21
Water	28
R.m.s. deviations	
Bond lengths (Å)	0.0145
Bond angles (°)	1.85
Ramachandran Plot Residues	
In most favourable regions (%)	97.04
In allowed regions (%)	2.70
Outliers (%)	0.27
MolProbity score	1.33
PDB code	**8S5B**

3. Chemistry experimental

Deuterochloroform, methanol-d_4 and deuterated water were used as solvents for ^{1}H NMR (400 MHz and 500 MHz) and ^{13}C NMR (100 and 150 MHz) spectroscopy. Multiplicity abbreviations are s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet. ^{1}H-^{1}H COSY and HSQC spectra were used to confirm proton and carbon assignments, respectively. Analytical thin layer chromatography (t.l.c) was performed on aluminium-backed 2 mm thick silica gel 60 F_{254} and were visualized with UV light and/or with sulfuric (5%) or cerium ammonium molybdate stains. High resolution mass spectra (HRMS) were obtained using an ESI-TOF-MS. Dry THF, DMF and CH_2Cl_2 were obtained from a dry solvent apparatus.[1] $[\alpha]_D$ values were measured using polarimeter and are given in deg·mL·g^{-1}·dm^{-1}. Octyl α-D-glucopyranoside and allyl α-D-glucopyranoside were purchased from Biosynth (UK). SQ was purchased from MCAT GmbH (Germany). All other reagents and reactants were from Merck.

Octyl 6-bromo-6-deoxy-α-D-glucopyranoside (2)

N-Bromosuccinimide (2.00 g, 11.2 mmol) was added to a solution of octyl α-D-glucopyranoside **1** (1.00 g, 3.42 mmol) and triphenylphosphine (2.40 g, 9.16 mmol) in DMF (40 mL), whilst the temperature was maintained at 0 °C. The mixture was stirred at rt for 24 h, and then the solvent was evaporated under reduced pressure. Flash chromatography of the residue (100% pet. ether to 100% EtOAc) afforded title compound **2** (0.74 g, 61%) as a light yellow oil, $[\alpha]_D^{20}$ = + 12.3° (c 0.17, MeOH). ^{1}H NMR (CD_3OD, 400 MHz) δ 0.88 (3 H, t, J = 6.7 Hz, CH_3), 1.19-1.42 (10 H, m, octyl-$(CH_2)_5$), 1.60 (2 H, m, octyl-CH_2), 3.26 (1 H, t, $J_{3,4}$ = $J_{4,5}$ 9.3 Hz, H4), 3.36-3.56 (3 H, m, H6a,2,octyl-H1'a), 3.61-3.84 (4 H, m, H3,5,6b,octyl-H1b), 4.78 (1 H, d, $J_{1,2}$ 3.8 Hz, H1); ^{13}C NMR (CD_3OD, 100 MHz) δ 13.2 (CH_3), 22.4, 26.0, 29.0, 29.2, 29.2, 31.6, 33.1 (C6), 67.9 (octyl-C1), 71.3, 72.1, 72.6, 73.5, 98.6 (C1); HRMS (ESI$^+$) calcd for $C_{14}H_{27}O_5BrNa^+$ $[M+Na]^+$ *m/z* 377.0940. Found 377.0833.

Octyl 6-deoxy-6-sulfo-α-D-glucopyranoside, sodium salt (3)

A solution of anhydrous sodium sulfite (0.843 g, 6.67 mmol) and compound **2** (0.25 g, 0.71 mmol) in water (15 mL) was heated under nitrogen at reflux for 8 h. The solution was cooled to room temperature and the solvent was evaporated under reduced pressure. Flash chromatography of the residue (EtOAc to EtOAc/MeOH/H_2O 7:2:1) gave title compound **3** as a white solid (0.25 g, 93%), $[\alpha]_{20}^{D}$ = + 1.5° (c 0.11, MeOH). ^{1}H NMR (CD_3OD, 400 MHz) δ 0.89 (3 H, t, J = 6.6 Hz, CH_3), 1.19-1.42 (10 H, m, octyl-$(CH_2)_5$), 1.62 (2 H, m, octyl-CH_2), 2.95 (1 H, dd, $J_{5,6a}$ = 8.6, $J_{6a,6b}$ = 14.4 Hz, H6a), 3.15 (1 H, t, $J_{3,4}$ = $J_{4,5}$ 9.3 Hz, H4), 3.36-3.56 (3 H, m, H2,6b,octyl-H1a), 3.93 (1 H, m, octyl-H1b), 3.67 (1 H, t, $J_{2,3}$ = $J_{3,4}$ 9.4 Hz, H3), 4.08 (1 H, m, H5), 5.16 (2 H, d, $J_{1,2}$ = 3.8 Hz, H1); ^{13}C NMR (CD_3OD, 100 MHz) δ 13.0 (CH_3), 22.4, 26.0, 29.0, 29.2, 29.4, 31.6, 53.0 (C6), 67.6 (octyl-C1), 68.1, 72.0, 72.4, 73.7, 98.1 (C1); HRMS (ESI$^-$) calcd for $C_{14}H_{28}O_8S^-$ $[M-H]^-$ *m/z* 355.1505. Found 355.1432.

N-Trifluoroacetyl-4,7,10-trioxa-1,13-tridecanediamine (5)

A solution of 4,7,10-trioxa-1,13-tridecanediamine **4** (6.00 g, 27.3 mmol) in anhydrous THF (12 ml) was cooled to -65 °C under N_2. Ethyl trifluoroacetate (3.24 ml, 27.3 mmol) was added dropwise to the solution. The mixture was stirred for 1.5 h at -65 °C, then was allowed to warm to room temperature. The reaction mixture was concentrated in vacuo, and water was added to the residue (20 ml), and the mixture was extracted with CH_2Cl_2 (6 × 25 ml). The combined organic layers were dried (Na_2SO_4) and filtered. The solvent was removed in vacuo to afford a dark yellow oil that was purified by flash chromatography (CH_2Cl_2/MeOH/NH_4OH, 9:1:0.2) affording the title compound **5** (3.53 g, 41%); ^{1}H NMR (400 MHz, $CDCl_3$) δ 1.53 (2 H, p, *J* = 6.0 Hz, H12), 1.65 (2 H, p, *J* = 6.0 Hz, H2), 2.61 (2 H, t, *J* = 6.6 Hz, H13), 3.24 (2 H, t, *J* = 6.3 Hz, H1), 3.33-3.48 (12 H, m, H3,5,6,8,9,11); ^{13}C NMR (100 MHz, $CDCl_3$) δ 28.1, 32.2 (C2,12), 37.6, 38.0 (C1,13), 69.0, 69.2, 69.3, 69.4, 69.79, 69.83, 70.0, 70.16, 70.21 (C3,5,6,8,9,11), 116.4 (q, *J* = 288 Hz, CF_3), 157.1 (q, C=O); HRMS (ESI$^+$) calcd for $C_{12}H_{24}F_3N_2O_3^+$ $[M+H]^+$ *m/z* 317.1680. Found 317.1680.

***N,N'*-Bis-[1-(trifluoroacetamido)-4,7,10-trioxa-tridec-13-yl] 4',4"-dithiodibutyramide (7)**

N-(3-Dimethylaminopropyl)-*N*'-ethylcarbodiimide (EDC) (3.64 g, 18.9 mmol) was added to a mixture of 4,4'-dithiodibutyric acid **6** (1.12 g, 4.7 mmol), compound **5** (3.00 g, 9.49 mmol), 1-hydroxybenzotriazole (HOBt) (1.54 g, 11.4 mmol) and DIPEA (6.61 ml, 38.0 mmol) in DMF (8 ml) at 0 °C, and the resulting mixture was stirred for 20 min at rt. The reaction was quenched with H_2O (30 ml) at 0 °C. The resulting mixture was extracted with ethyl acetate (50 ml), and the combined organic layers were washed with NaCl (50 ml), dried (Na_2SO_4), filtered and concentrated in vacuo. The residue was purified by flash chromatography (EtOAc:MeOH, 9.5:0.5) to afford compound **7** as an off-white solid (4.9 g, 62%), ^{1}H NMR (400 MHz, CD_3OD) δ 1.77 (4 H, p, *J* = 6.4 Hz, H2), 1.83 (4 H, p, *J* = 6.5 Hz, H12), 1.99 (4 H, p, *J* = 7.0 Hz, H3'), 2.31 (4 H, t, *J* = 7.4 Hz, H2'), 2.72 (4 H, t, *J* = 7.1 Hz, H4'), 3.26 (4 H, t, *J* = 6.8 Hz, H1), 3.31 (bs, NH), 3.35 (s, NH), 3.39 (4 H, t, *J* = 6.9 Hz, H13), 3.50-3.66 (12 H, 3m, H3,5,6,8,9,11); ^{13}C NMR (100 MHz, CD_3OD) δ 25.0 (C2'), 28.5, 29.1 (C2,12), 34.2, 34.3 (C3',4'), 37.1, 37.6 (C1,13), 68.3, 68.6, 69.9, 70.0, 70.2 (C3,5,6,8,9,11), 116.3 (q, *J* = 287 Hz, CF_3), 157.3 (q, *J* = 37 Hz, **C**OCF_3), 173.4 (C1', C=O); HRMS (ESI$^+$) calcd for $C_{32}H_{57}F_6N_4O_{10}S_2^+$ $[M+H]^+$ *m/z* 835.3415. Found 835.3400.

5-Aza-4,20-dioxo-21,21,21-trifluoro-9,12,15-trioxa-heneicosane-1-thiol (8)

Tris(2-carboxyethyl)phosphine hydrochloride (0.412 g, 1.64 mmol) was added to disulfide **7** (1.00 g, 1.20 mmol) in acetate buffer at pH 4.0. The reaction mixture was stirred until the disulfide was consumed. The resulting mixture was extracted with deoxygenated ethyl acetate (5 × 10 ml), and the organic layers combined and concentrated under reduced pressure. Crude thiol **8** was used directly in the next reaction.

9-Aza-8,24-dioxo-4-thia-25,25,25-trifluoro-13,16,19-trioxa-pentacosyl α-D-sulfoquinovoside (10)

Allyl α-sulfoquinovoside **9**[2] (200 mg, 0.71 mmol) and the thiol **8** (354 mg, 0.85 mmol) were added to a solution of Vazo44 (68 mg, 0.21 mmol) in acetate buffer (pH 4) (6 ml) under N_2. The reaction mixture was stirred for 8 h at 50 °C. An additional portion of the thiol (296 mg, 0.71 mmol) and Vazo44 (45 mg, 0.14 mmol) were added, and stirred for 6 h under nitrogen atmosphere. The resulting residue was partially purified by flash chromatography to afford an off-white solid. The solid was further purified by cation exchange chromatography (Dowex-50, H^+ form, eluted with deionized water) to obtain compound **10** an off-white oily solid (218 mg, 44%); $[\alpha]_D^{22}$ +33.2 (c 0.885, methanol); ^{1}H NMR (500 MHz, D_2O) δ 1.63 (2 H, p, *J* = 6.6 Hz, H11), 1.63-1.79 (6 H, m, H2,6,21), 2.19 (2 H, t, *J* = 7.3 Hz, H7), 2.43 (2 H, t, *J* = 7.2 Hz, H5), 2.47-2.57 (2 H, m, H3), 2.89 (1 H, dd, *J* = 14.5, 9.8 Hz, H6'b), 3.07-3.11 (3 H, m, H4,10), 3.21 (1 H, dd, *J* = 14.4, 1.6 Hz, H6a), 3.25 (2 H, t, *J* = 6.8 Hz, H22), 3.37-3.57 (16 H, H1b,12,14,15,17,18,20,2',3',4'), 3.76 (1 H, dt, *J* = 9.9, 3.4 Hz, H1a), 3.91 (1 H, td, *J* = 10.2, 1.5 Hz, H5'), 4.70 (1 H, d, *J* = 3.8 Hz, H1'); ^{13}C NMR (150 MHz, CD_3OD) δ 25.1 (C6), 26.5, 27.6, 27.8, 28.2, 28.7, 30.3, 34.7, 36.4, 37.0, 37.7, 52.1 (C6), 66.5, 67.9, 68.2, 68.3, 68.4, 69.3, 69.4, 69.5, 69.6, 71.2, 72.5, 73.0, 97.6 (C1'), 116.0 (q, *J* = 37 Hz, CF_3), 158.5 (q, *J* = 284 Hz, **C**OCF_3), 176.0 (CONH); HRMS (ESI^-) calcd for $C_{25}H_{44}F_3N_2O_{13}S_2^-$ $[M-H]^-$ *m/z* 701.2242. Found 701.2238.

9-Aza-4,4,8,24-tetraoxo-4-thia-25,25,25-trifluoro-13,16,19-trioxa-pentacosyl α-D-sulfoquinovoside (11)

mCPBA (221 mg, 1.28 mmol) was added to a solution of thioether **10** (200 mg, 0.29 mmol) in acetone:water (6:4) and the reaction mixture was stirred for 5 h. The resulting residue was partially purified by flash chromatography (EtOAc:MeOH:H_2O, 7:2:1) to afford an off-white solid. The solid was purified further by cation exchange chromatography (Dowex-50, H^+ form, eluted with deionized water) to obtain compound **11** as an off-white solid (188 mg, 90%); $[\alpha]_D^{22}$ +12.7 (c 0.525, methanol); ^{1}H NMR (500 MHz, D_2O) δ 1.64 (2 H, p, *J* = 6.6 Hz, H15'), 1.71 (2 H, p, *J* = 6.5 Hz, H8'), 1.96 (4 H, m, H5',2'), 2.28 (2 H, t, *J* = 7.4 Hz, H6'), 2.90 (2 H, dd, *J* = 15.5, 10.4 Hz, H6b), 3.11 (5 H, m, H4,4',7'), 3.24 (5 H, m, H6a,3',16'), 3.43 (5 H, m, H2,9',14'), 3.52 (10 H, m, H3,1',10',11',12',13'), 3.80 (1 H, ddd, *J* = 10.0, 6.5, 5.1 Hz, H1'), 3.88 (1 H, m, H5), 4.73 (1 H, d, *J* = 3.8 Hz, H1); ^{13}C NMR (150 MHz, CD_3OD) δ 17.7 (C5'), 21.3 (C2'), 27.7 (C15'), 28.2 (C8'), 33.8 (C6'), 36.4 (C16'), 37.0 (C7'), 49.5 (C3'), 50.9 (C4'), 52.0 (C6), 65.4 (C5), 68.1 (C14'), 68.2 (C9'), 68.4 (C1'), 69.5 (C11',12',13',14'), 71.2 (C3), 72.4 (C4), 73.0 (C2), 97.6 (C1), 115.7 (CF_3), 158.7 ($COCF_3$), 174.6 (CONH); HRMS (ESI^-) calcd for $C_{25}H_{44}F_3N_2O_{15}S_2^-$ $(M-H)^-$ *m/z* 733.2141. Found 733.2102.

9-Aza-4-thia-4,4,8,24-tetraoxo-13,16,19-trioxa-tricosyl α-D-sulfoquinovoside (12; SQ-amine)

NaOMe was added to a solution of compound **11** (150 mg, 0.20 mmol) in methanol until pH turns 8. The reaction mixture was stirred overnight. The reaction mixture was then quenched with 1% acetic acid. It was then concentrated in vacuo. The remainder was mixed with water (10 ml), and the aqueous solution was extracted with EtOAc (2 × 10 ml). The aqueous layer was then concentrated in vacuo to obtain octyl-SQ **12** as an off-white oily solid (118 mg, 91%); $[\alpha]_D^{22}$ +33.8 (c 1.30,

methanol); ^{1}H NMR (500 MHz, D_2O) δ 1.64 (2 H, p, *J* = 6.6 Hz, H15'), 1.80 (2 H, ddd, *J* = 13.4, 6.8, 5.9 Hz, H8'), 1.97 (4 H, m, H2', H5'), 2.28 (2 H, t, *J* = 7.4 Hz, H6'), 2.90 (1 H, dd, *J* = 15.5, 9.9 Hz, H6b), 2.96 (2 H, t, *J* = 7.1 Hz, H16'), 3.11 (5 H, m, H4,4',7'), 3.23 (3 H, m, H6a,3'), 3.42 (3 H, m, H2,14'), 3.53 (12 H, m, H3,1',9',10',11',12',13'), 3.80 (1 H, ddd, *J* = 10.3, 6.0, 5.1 Hz, H1'), 3.87 (1 H, m, H5), 4.73 (1 H, d, *J* = 3.8 Hz, H1); ^{13}C NMR (150 MHz, CD_3OD) δ 17.8 (C5'), 21.4 (C2'), 26.5 (C15'), 28.2 (C8'), 33.8 (C6'), 36.4 (C7'), 37.7 (C16'), 49.5 (C3'), 50.9 (C4'), 52.0 (C6), 65.5 (C5), 68.1 (C1'), 68.3 (C9', C14'), 69.4 (C10', C11', C12', C13'), 71.2 (C3), 72.4 (C4), 72.9 (C2), 97.6 (C1), 174.6 (CONH); HRMS (ESI^-) calcd for *m/z* $C_{23}H_{45}N_2O_{14}S_2^-$ $[M-H]^-$ 637.2318. Found 637.2310.

***N*-Butyryl-4,7,10-trioxa-1,13-tridecanediamine (13)**

Butyric anhydride (677 ml, 4.10 mmol) was added dropwise to a solution of 4,7,10-trioxa-1,13-tridecanediamine **4** (500 mg, 2.30 mmol) and Et_3N (385 ml, 2.80 mmol) in anhydrous THF (5 ml) under nitrogen atmosphere. The mixture was stirred for 4 h at rt. The reaction mixture was concentrated in vacuo and the residue was purified by flash chromatography (0.2% Et_3N in $CHCl_3$:MeOH:water) to afford compound **13** (349 mg, 53%); ^{1}H NMR (500 MHz, CD_3OD) δ 0.95 (3 H, t, *J* = 7.4 Hz, CH_3), 1.63 (2 H, hex, *J* = 7.4 Hz, $CH_3C\mathbf{H}_2$), 1.76 (2 H, p, *J* = 6.6 Hz, H12), 1.92 (2 H, p, *J* = 7.4 Hz, H2), 2.16 (2 H, t, *J* = 7.4 Hz, CH_2CO), 3.08 (2 H, t, *J* = 6.5 Hz, H13), 3.26 (2 H, t, *J* = 7.0 Hz, H1), 3.52 (2 H, t, *J* = 6.2 Hz, CH_2O), 3.57-3.67 (10 H, m, 5 x CH_2O); ^{13}C NMR (150 MHz, CD_3OD) δ 14.0 (CH_3), 20.4 (CH_2C=O), 28.2, 30.5, 37.6, 39.0. 49.9 (2 C, CH_2N), 697, 70.2, 71.05, 71.09, 71.2, 71.4 (C3,5,6,8,9,11), 178.4 (CH_3CO); HRMS $(ESI)^+$ calcd for $C_{14}H_{31}N_2O_4^+$ $[M+H]^+$ *m/z* 291.2278. Found 291.2280.

4. Biochemistry experimental

Expression and purification of SQBP

For protein crystallization: SmoF was cloned, over-expressed and purified as described previously.[3,4]

For affinity column studies: The following modified approach was used. pET29-SmoF[3,4] was transformed into the *E. coli* strain SHuffle T7 express (NEB). The cells were grown in lysogeny broth with 50 µg/ml kanamycin at 30 °C and induced with 0.5 mM isopropyl 1-thio-β-D-galactopyranoside (IPTG) when they reached an A_{600} of 0.8. After incubation at 16 °C overnight, cells were harvested by centrifugation at 6000 × *g*. The cell pellet was resuspended in buffer A (20 mM Tris-HCl (pH 7.5), 500 mM NaCl, 25 mM imidazole), disrupted using a TS Series benchtop cell-disruptor, and clarified by centrifugation at 40000 × *g*. The soluble fraction was applied to a chelating column (Hi-Trap HP, Cytiva) charged with Ni^{2+}. SQBP was eluted using a gradient of 25–500 mM imidazole in buffer A. Pooled fractions were concentrated and applied to a size-exclusion column (Superdex 75 increase 10/300 GL, Cytiva) equilibrated with buffer B (20 mM Tris HCl (at pH 7.5), 200 mM NaCl). Final protein samples at 1 mg/ml were snap-frozen using liquid nitrogen and stored at 80 °C.

Nano Differential Scanning Fluorimetry (nanoDSF)

NanoDSF studies were performed on a Prometheus NT.48 (NanoTemper). Data recording and initial analysis was performed with PR.ThermControl software. All SmoF samples were at 1 $mg.ml^{-1}$ in 50 mM Tris, 300 mM NaCl pH 7.5. with a 15 µl capillary load per sample. 2 mM octyl-SQ was added to SmoF sample and incubated for 10 minutes at 20°C prior to run. Experiments were performed in duplicates with the temperature ramp from 15 °C to 95 °C, at 1.0 °C/min with 15% excitation power. All samples were centrifuged at 13,000 *g* for 5 min prior to loading.

Isothermal titration calorimetry (ITC)

ITC experiments were performed using a MicroCal iTC200 (GE Healthcare) at 25 °C, with a 750 rpm stirring speed and a reference power of 10 $\mu cal.s^{-1}$. *A. tumefaciens* SmoF protein and ligand (octyl-SQ) were equilibrated into degassed and filter-sterilized buffer (50 mM NaPi, 200 mM NaCl, pH 7.4). Protein concentration was determined by BCA assay (Thermo Fisher) before initiating experiments. 300 µM octyl-SQ was titrated into the ITC cell containing 30 µM SmoF as a series of 15 × 2.94 µl injections with a pre-injection of 1 × 0.4 µl. The delay between injections was set at 180 s, with an initial injection delay of 240 s. Data analysis was performed in MicroCal ITC Origin Analysis software (Malvern).

Preparation of SQ-linker and control-linker conjugated resin

SQ linker and control linker were dissolved in the conjugation buffer (0.1 M $NaHCO_3$ (pH 8.4), 0.5 M NaCl). Cyanogen bromide-activated-Sepharose 4B (Sigma-Aldrich) was prepared as per the manufacturer's instructions by swelling and washing the resin extensively with cold 1 mM HCl followed by equilibrating with the conjugation buffer immediately before conjugation. SQ linker or control linker were added to the resin (5 µmole/ml resin) and incubated overnight at 4 °C with gentle rocking. After removing unreacted linkers, the resin was washed with the conjugation buffer and incubated with 0.2 M glycine (pH 8.0) to block unreacted groups for 2 h at room temperature. The conjugated resin was subjected to five cycles of alternate washing with the conjugation buffer and 0.1 M Na acetate (pH 4.0), 0.5 M NaCl. The resin was washed with deionized water and stored with 20% ethanol at 4 °C until use.

Affinity capture of SQBP using SQ-linker conjugated resin

50 µg of purified SQBP or total soluble fraction from 1–1.5 ml cell culture was incubated with 50 µl SQ-linker or control-linker conjugated resin for 1 h at room temperature with gentle rocking. After removing unbound proteins, the resin was washed with 20 column volumes (CV) of buffer B. Elution with the indicated concentrations of SQ, *S*-2,3-dihydroxypropanesulfonate (*S*-DHPS), *S*-sulfolactate (*S*-SL) and 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid (HEPES) in buffer B was performed and the eluants were analysed by SDS-PAGE. Densitometric analysis of SDS-PAGE images was performed using ImageJ.[5]

5. Structural biology experimental

Protein Crystallization and Optimization

Initial screening was performed using commercially available INDEX (Hampton Research), PACT premier and CSSI/II (Molecular Dimensions) screens in 96-well sitting drop trays. Further optimization was carried out in a 48 well sitting drop or 24 well hanging-drop format to obtain optimal crystals for X-ray diffraction. For co-crystallization experiments a 48 well sitting drop plate was used. 2 mM octyl-SQ was added to 50 $mg.ml^{-1}$ SmoF in 50 mM Tris, 300 mM NaCl pH 7.5. A mother liquor containing 200 mM calcium acetate, 100 mM Bis-TRIS pH 6.5, 25% w/v PEG 2000 MME was used in a 1:1 protein: mother liquor ratio with a final drop volume of 1 μL. Octyl-SQ was incubated with SmoF for 10 minutes, then centrifuged at 13,000 xg for 3 minutes prior to drop formation. All incubation and crystal growth was performed at 20 °C.

Data collection and structure determination

All crystals were tested using a Rigaku MicroMax 007HF X-ray generator with an RAXIS IV++ imaging plate detector and test images were collected at 100 K using a 700 Series Cryostream (Oxford Cryosystems). Diffraction pattern quality assessment and resolution estimate performed using ADXV.[6] X-ray data collection occurred at the Diamond Light Source, using beamline I-03 during collection mx18598-51. Data collection statistics are available in Supplementary Table 1. The data were processed and integrated using XDS and scaled using SCALA included in the Xia2 processing system.[7] [8] [9] [10] In all cases AIMLESS was used for data reduction and quality assessment.[11] Resolution was cut to CC1/2 =0.5. Molecular replacement was carried out using MOLREP[12] with 7OFY.pdb as the initial search model, and early model building was automated using BUCCANEER.[13] The structure was built and refined using iterative cycles using COOT [14] and REFMAC[15] the latter employing local NCS restraints. Following building and refinement of the protein and water molecules, clear residual density was observed in the omit maps octyl-SQ ligand. The coordinate and refinement library files for ligand were prepared using ACEDRG [16] and ligand was modelled at an occupancy of 1. All steps were performed from within the CCP4i2 suite.[17] The coordinate files and structure factors have been deposited in the Protein Data Bank (PDB) with accession number: **8S5B.**

All figures were prepared using CCP4MG, Pymol or UCSF Chimera (Schrödinger, L.L.C. (2015). The PyMol Molecular Graphics System), depending on the desired outcome.[18,19] Analyses of conformational changes and internal cavities were performed using the DynDom3D web server and the CASTp V.3.0 Pymol plugin, respectively.[20] [21]

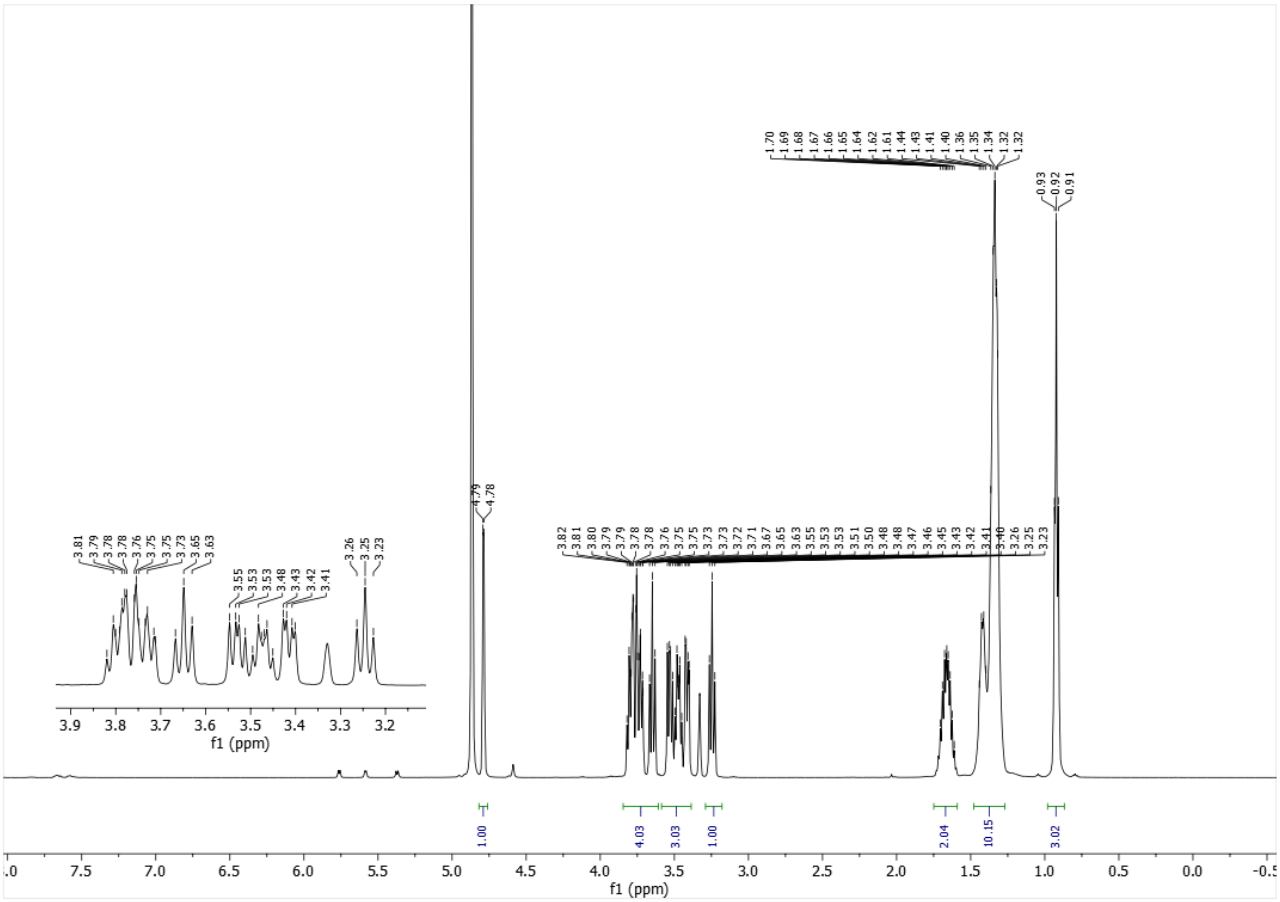
6. References

1. Pangborn, A. B., Giardello, M. A., Grubbs, R. H., Rosen, R. K., and Timmers, F. J. (1996) Safe and Convenient Procedure for Solvent Purification. *Organometallics* **15**, 1518-1520
2. Abayakoon, P., Epa, R., Petricevic, M., Bengt, C., Mui, J. W. Y., van der Peet, P. L., Zhang, Y., Lingford, J. P., White, J. M., Goddard-Borger, E. D., and Williams, S. J. (2019) Comprehensive synthesis of substrates, intermediates and products of the sulfoglycolytic Embden-Meyerhoff-Parnas pathway. *J. Org. Chem.* **84**, 2901-2910
3. Snow, A. J. D., Sharma, M., Lingford, J. P., Zhang, Y., Mui, J. W., Epa, R., Goddard-Borger, E. D., Williams, S. J., and Davies, G. J. (2022) The sulfoquinovosyl glycerol binding protein SmoF binds and accommodates plant sulfolipids. *Curr. Res. Struct. Biol.* **4**, 51-58
4. Sharma, M., Lingford, J. P., Petricevic, M., Snow, A. J. P., Zhang, Y., Jarva, M., , Mui, J. W.-Y., Scott, N. E., Saunders, E. C., Mao, R., Epa, R., da Silva, B. M., Pires, D. E. V., Ascher, D. B., McConville, M. J., Davies, G.J., , and Williams, S. J., Goddard-Borger, E.D. (2022) Oxidative desulfurization pathway for complete catabolism of sulfoquinovose by bacteria. *Proc. Natl. Acad. Sci. USA* **119**, e2116022119
5. Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671-675
6. Porebski, B. T., Ho, B. K., and Buckle, A. M. (2013) Interactive visualization tools for the structural biologist. *J. Appl. Crystallogr.* **46**, 1518-1520
7. Kabsch, W. (2010) Xds. *Acta Crystallogr., Section D: Biol. Crystallogr.* **66**, 125-132
8. Evans, P. (2006) Scaling and assessment of data quality. *Acta Crystallogr. Sect. D* **62**, 72-82
9. Winter, G. (2010) xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* **43**, 186-190
10. Winter, G., Lobley, C. M., and Prince, S. M. (2013) Decision making in xia2. *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1260-1273
11. Evans, P. R., and Murshudov, G. N. (2013) How good are my data and what is the resolution? *Acta Crystallogr. Sect. D* **69**, 1204-1214
12. Alexei Vagin and Alexei, T. (2010) Molecular replacement with MOLREP. *Acta Crystallographica Section D: Biological Crystallography* **66**, 22-25 , pmid = 20057045 , publisher = International Union of Crystallography (IUCr)
13. Vagin, A., and Teplyakov, A. (1997) MOLREP: an Automated Program for Molecular Replacement. *J. Appl. Crystallogr.* **30**, 1022-1025
14. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. Sect. D* **66**, 486-501
15. Murshudov, G. N., Skubak, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D* **67**, 355-367
16. Long, F., Nicholls, R. A., Emsley, P., Grazulis, S., Merkys, A., Vaitkus, A., and Murshudov, G. N. (2017) AceDRG: a stereochemical description generator for ligands. *Acta Crystallogr. Sect. D* **73**, 112-122
17. Potterton, L., Agirre, J., Ballard, C., Cowtan, K., Dodson, E., Evans, P. R., Jenkins, H. T., Keegan, R., Krissinel, E., Stevenson, K., Lebedev, A., McNicholas, S. J., Nicholls, R. A., Noble, M., Pannu, N. S., Roth, C., Sheldrick, G., Skubak, P., Turkenburg, J., Uski, V., von Delft, F., Waterman, D., Wilson, K., Winn, M., and Wojdyr, M. (2018) CCP4i2: the new graphical user interface to the CCP4 program suite. *Acta Crystallogr. D Struct. Biol.* **74**, 68-84
18. McNicholas, S., Potterton, E., Wilson, K. S., and Noble, M. E. M. (2011) Presenting your structures: the CCP4mg molecular-graphics software. *Acta Crystallogr. D* **67**, 386-394
19. Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., Morris, J. H., and Ferrin, T. E. (2021) UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70-82
20. Girdlestone, C., and Hayward, S. (2016) The DynDom3D Webserver for the Analysis of Domain Movements in Multimeric Proteins. *J. Comput. Biol.* **23**, 21-26
21. Tian, W., Chen, C., Lei, X., Zhao, J., and Liang, J. (2018) CASTp 3.0: computed atlas of surface topography of proteins. *Nucleic Acids Res.* **46**, W363-W367

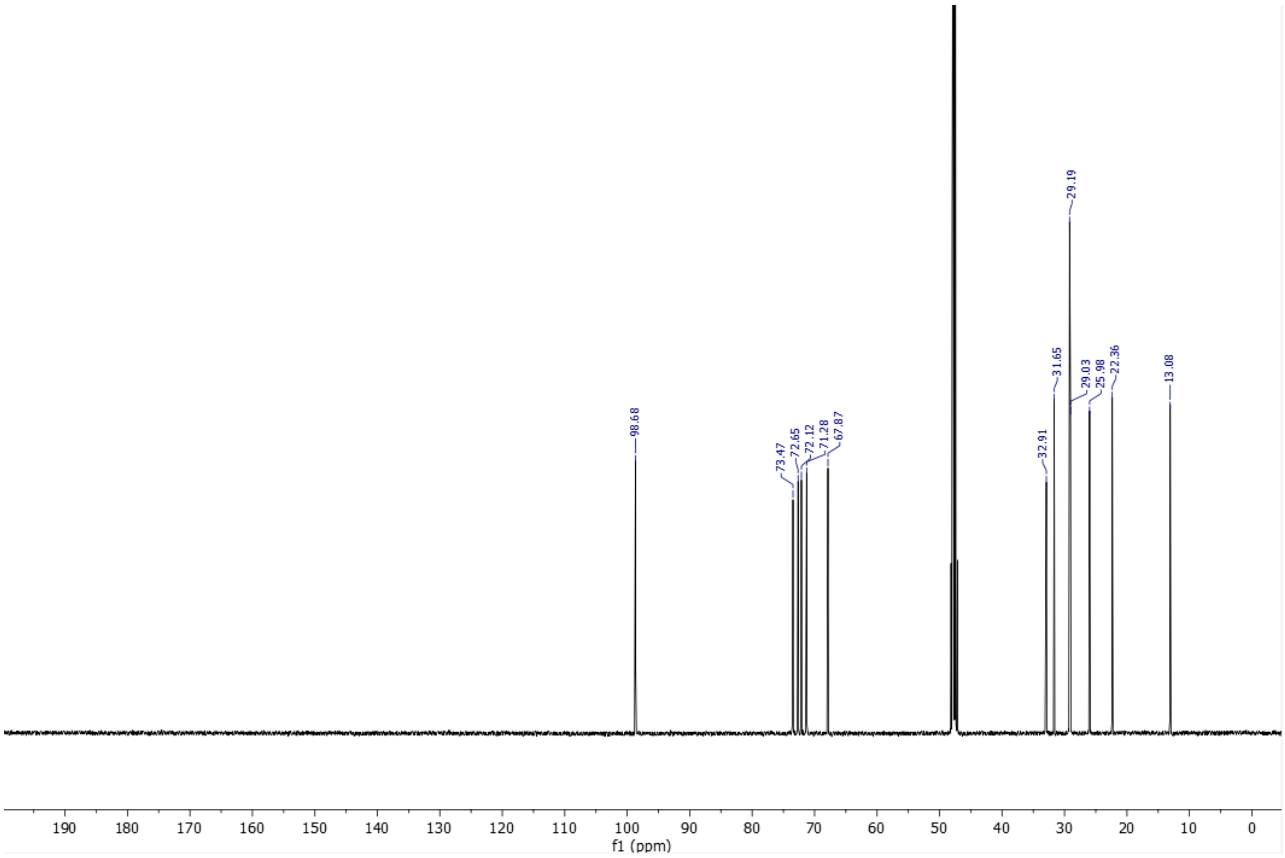
7. NMR spectra

Octyl 6-bromo-6-deoxy-α-D-glucopyranoside (2)

^{1}H NMR

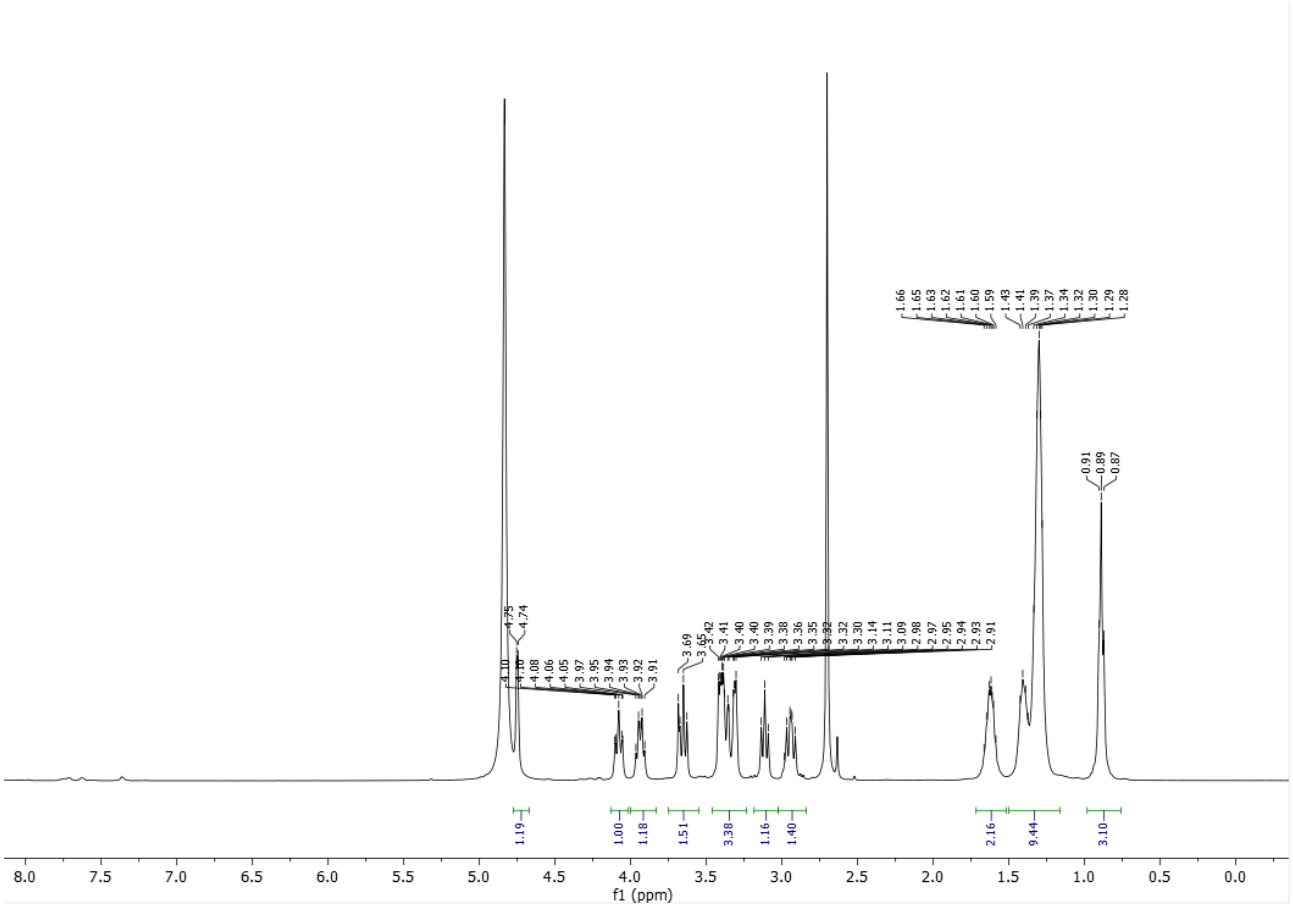


^{13}C NMR

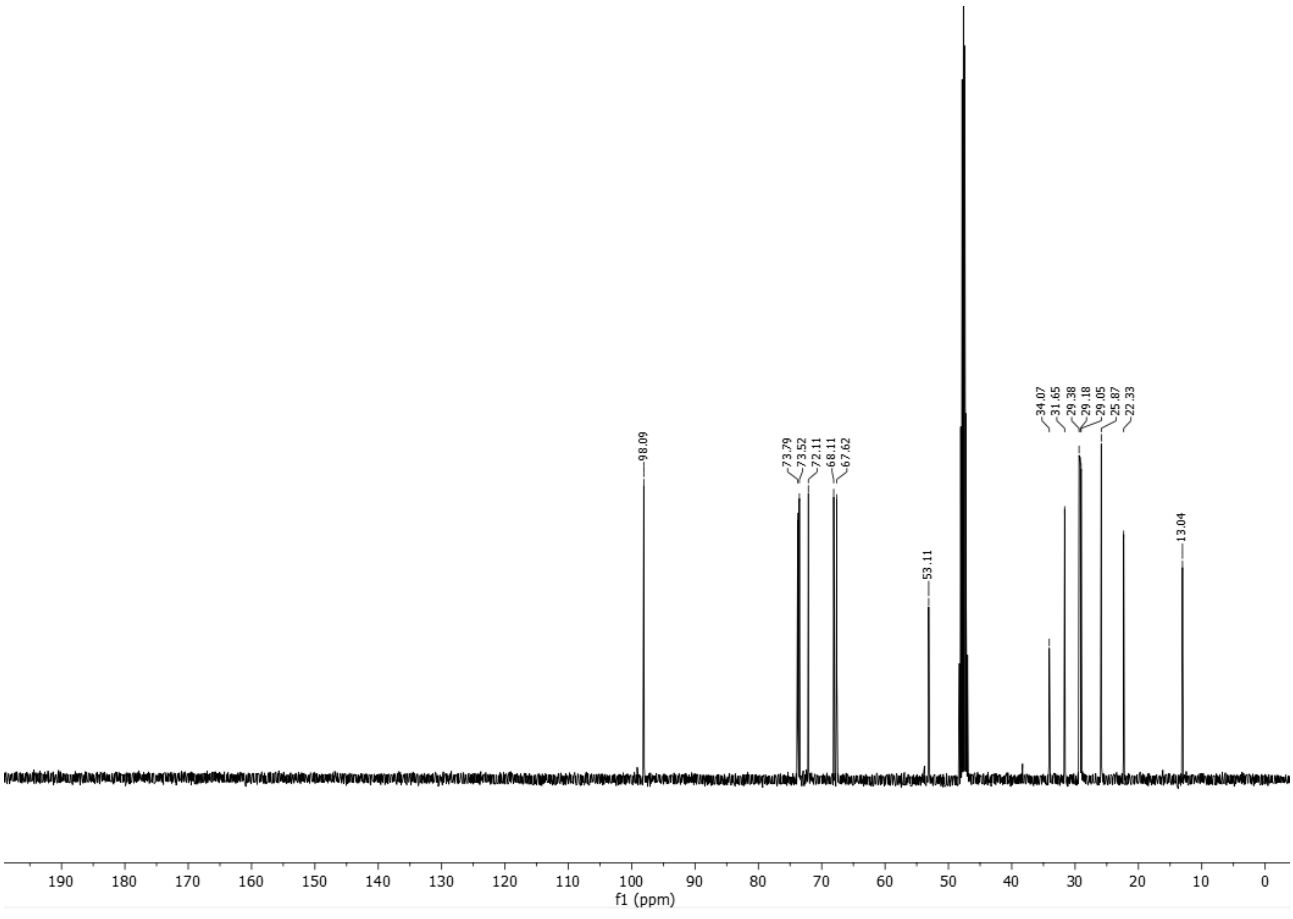


Octyl 6-deoxy-6-sulfo-α-D-glucopyranoside, sodium salt (3)

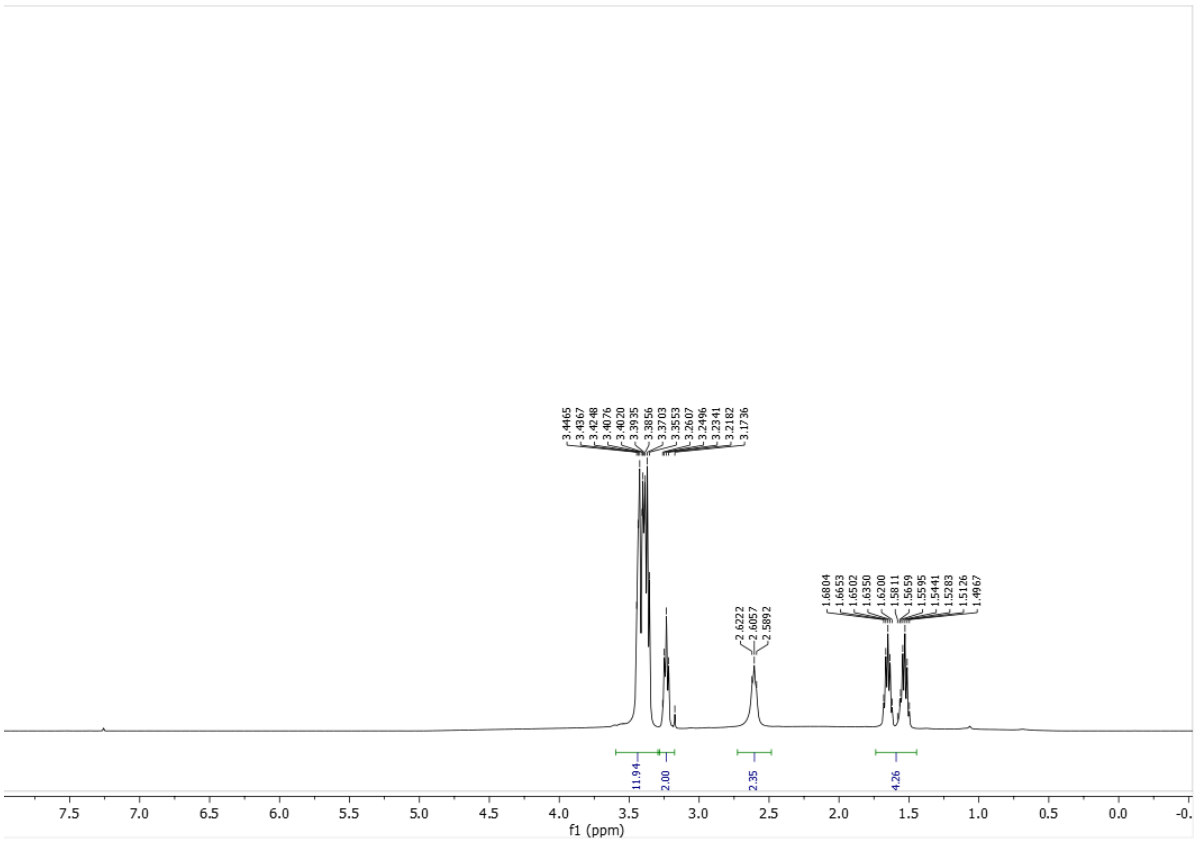
^{1}H NMR



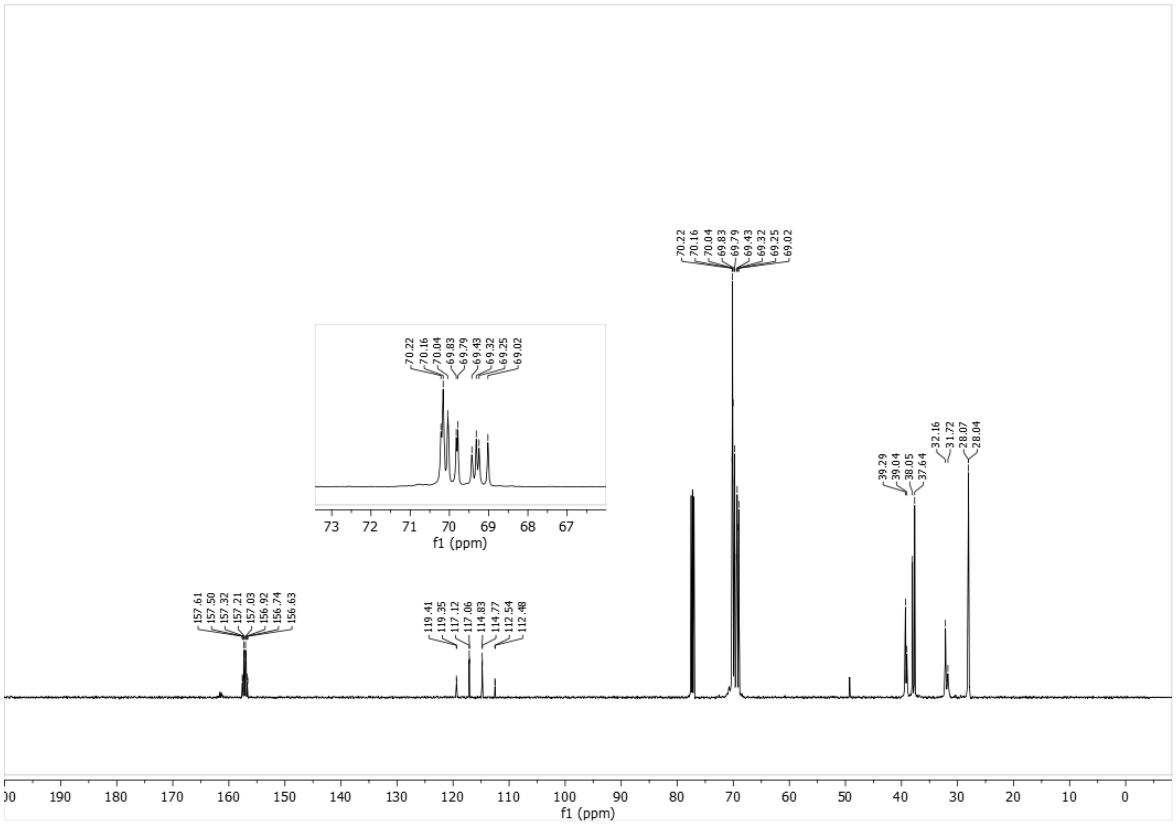
^{13}C NMR



N-Trifluoroacetyl-4,7,10-trioxa-1,13-tridecanediamine (5)

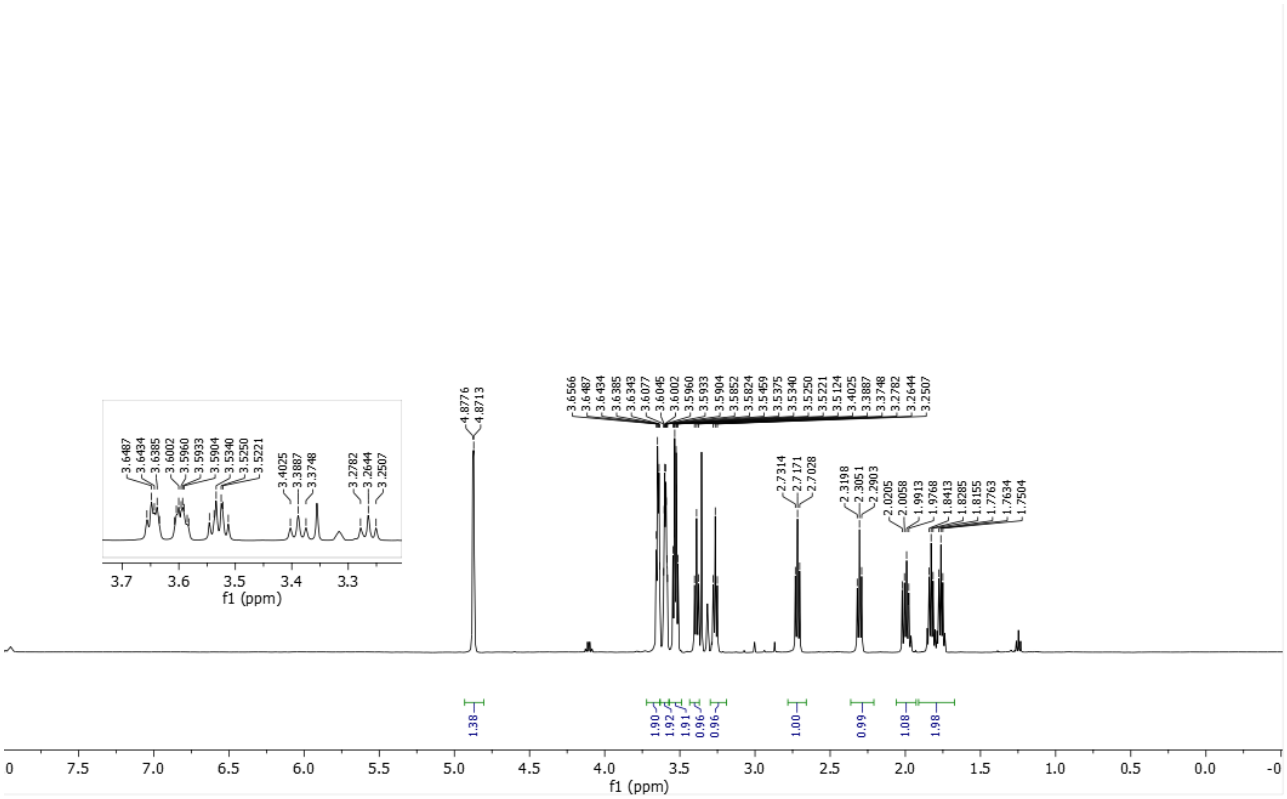


^{13}C NMR

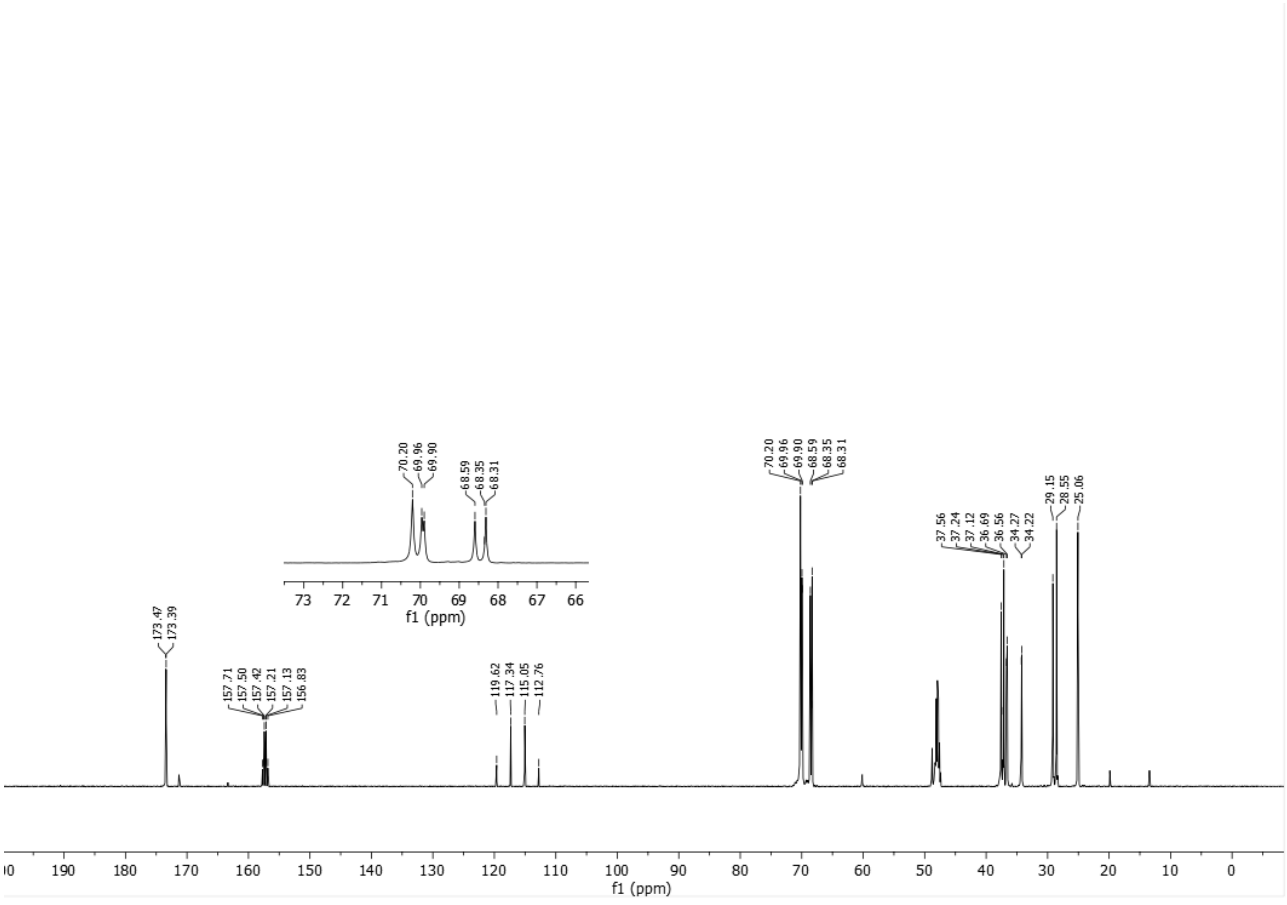


N,N'-Bis-[1-(trifluoroacetamido)-4,7,10-trioxa-tridec-13-yl] 4',4"-dithiodibutyramide (7)

^{1}H NMR

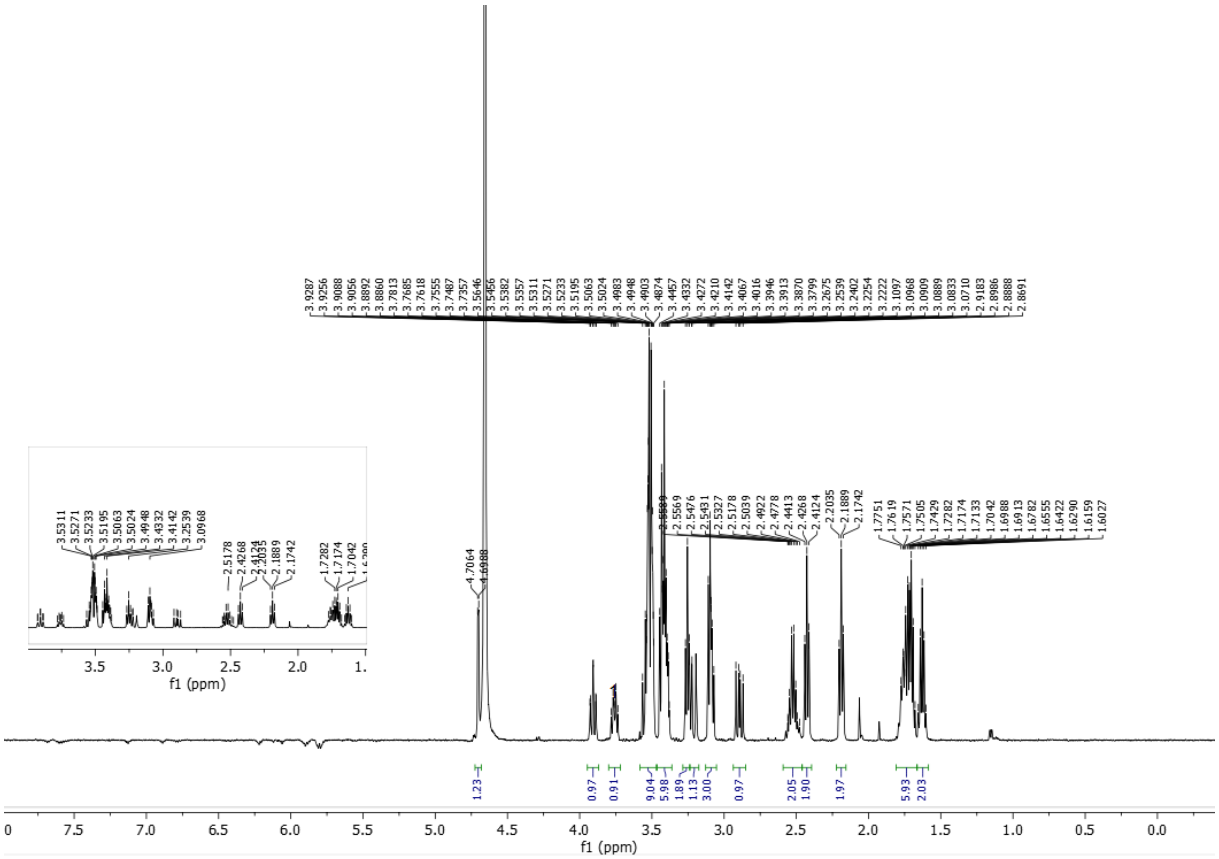


^{13}C NMR

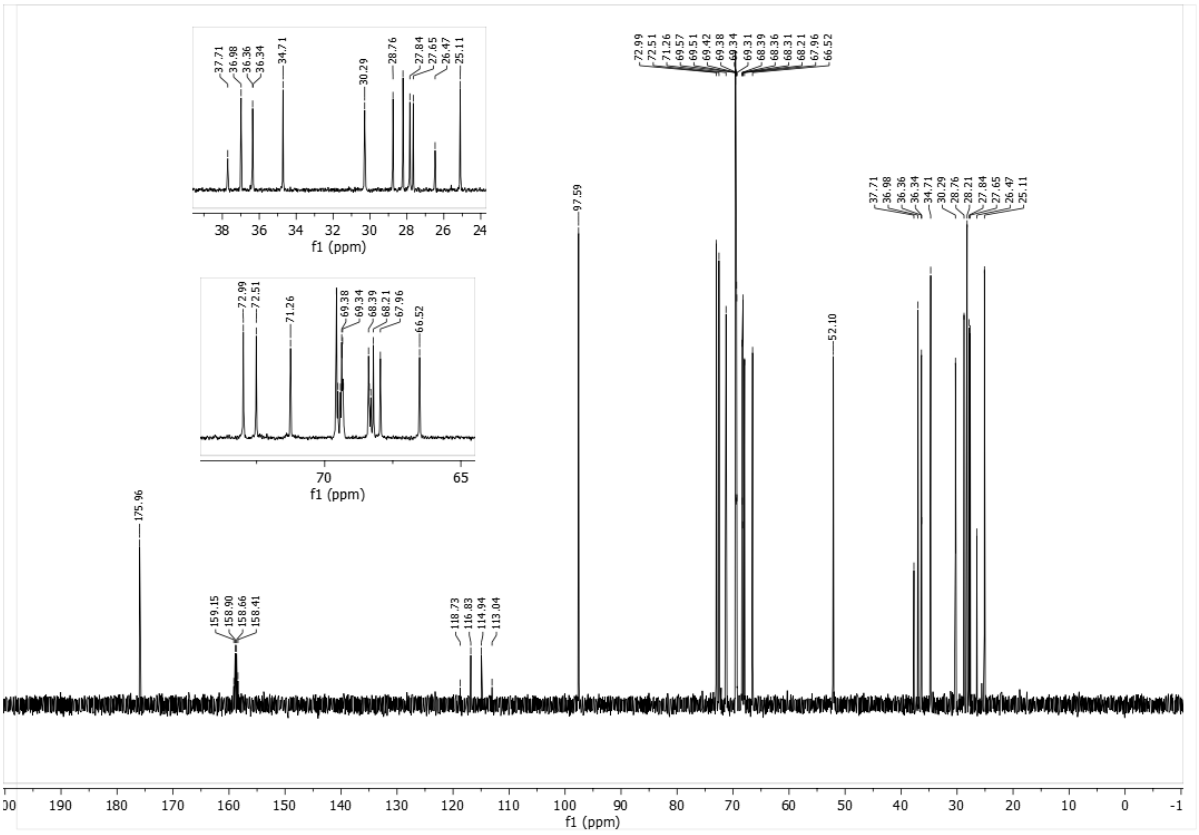


9-Aza-8,24-dioxo-4-thia-25,25,25-trifluoro-13,16,19-trioxa-pentacosyl α-D-sulfoquinovoside (10)

^{1}H NMR

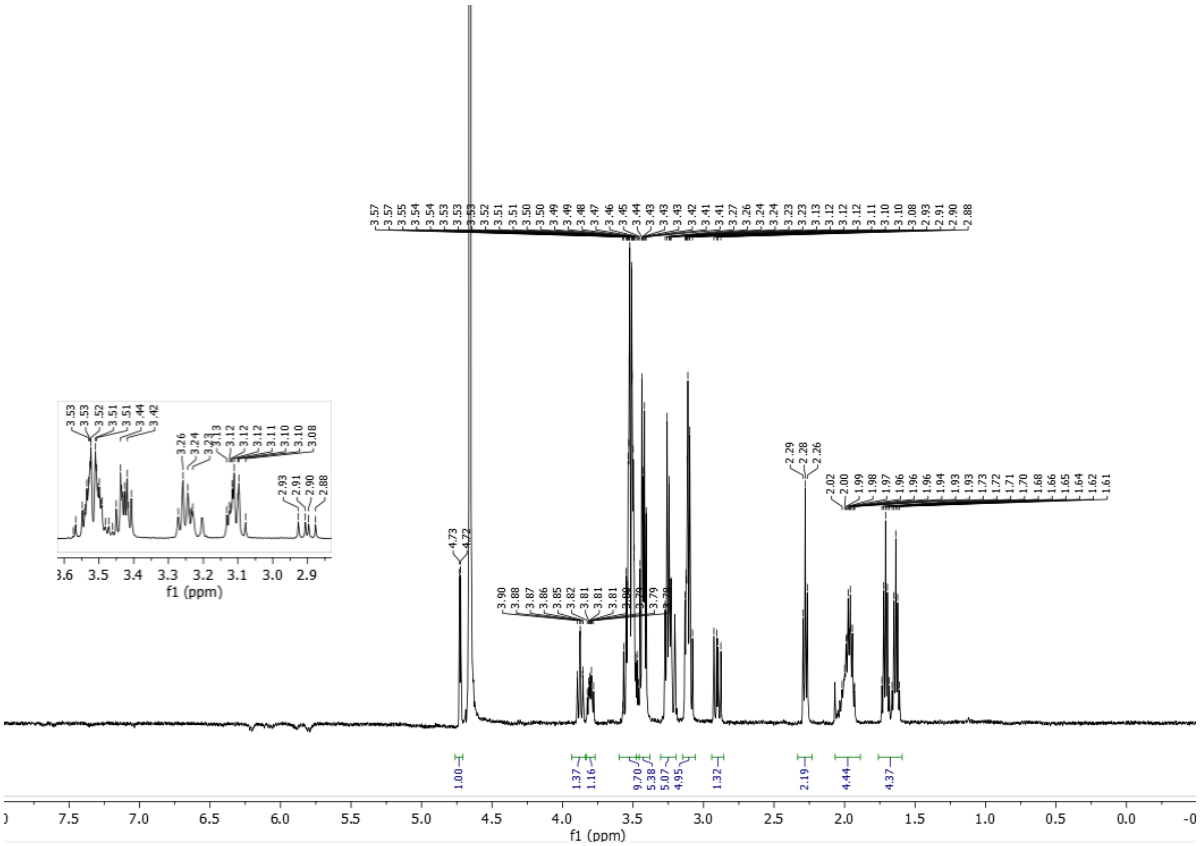


^{13}C NMR

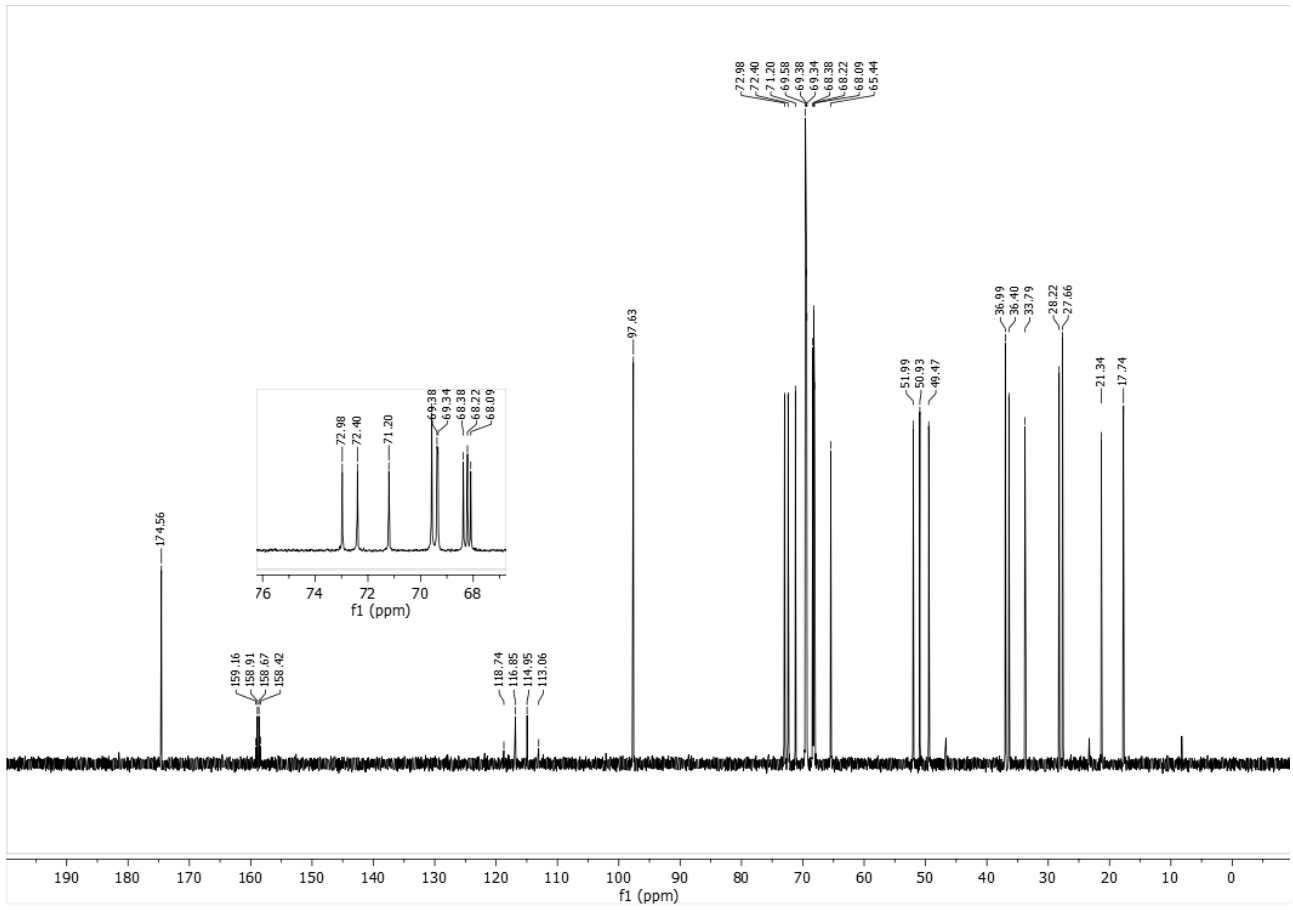


9-Aza-4,4,8,24-tetraoxo-4-thia-25,25,25-trifluoro-13,16,19-trioxa-pentacosyl α-D-sulfoquinovoside (11)

^{1}H NMR

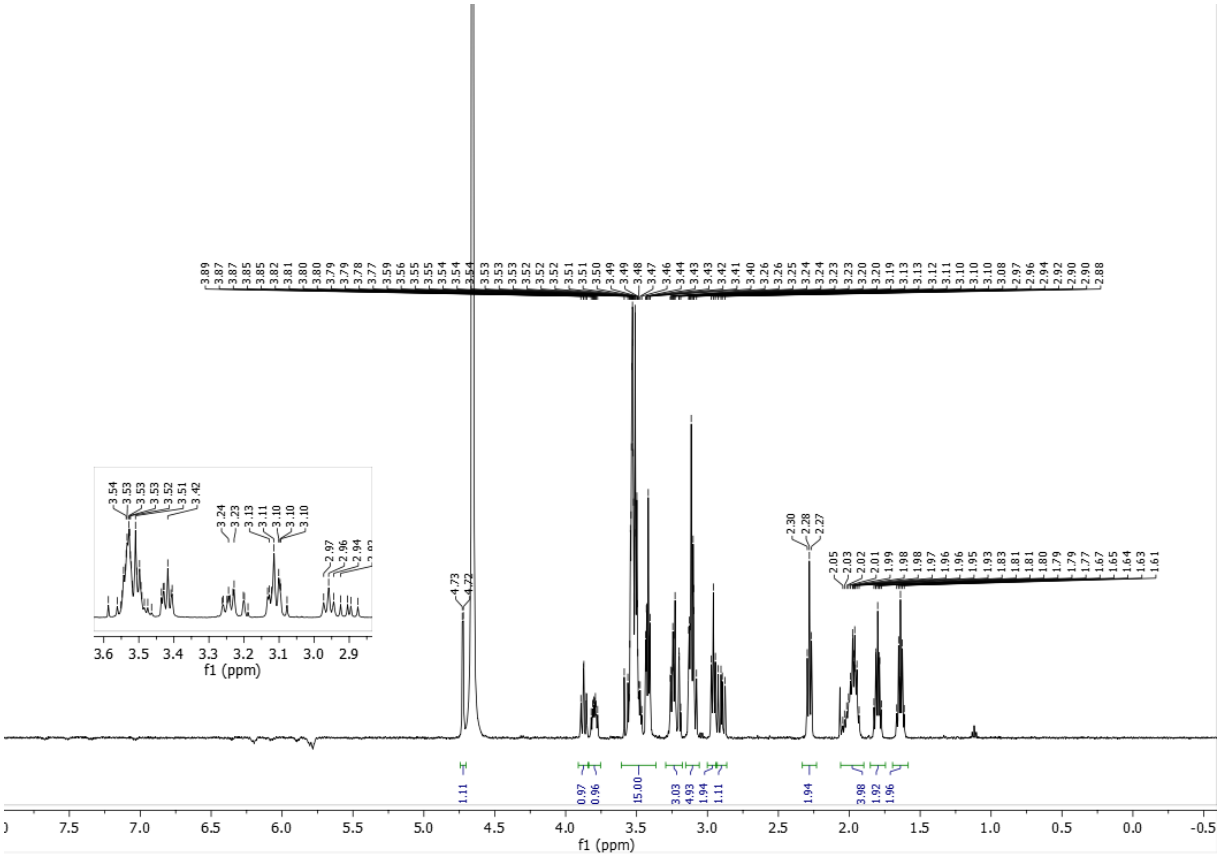


^{13}C NMR

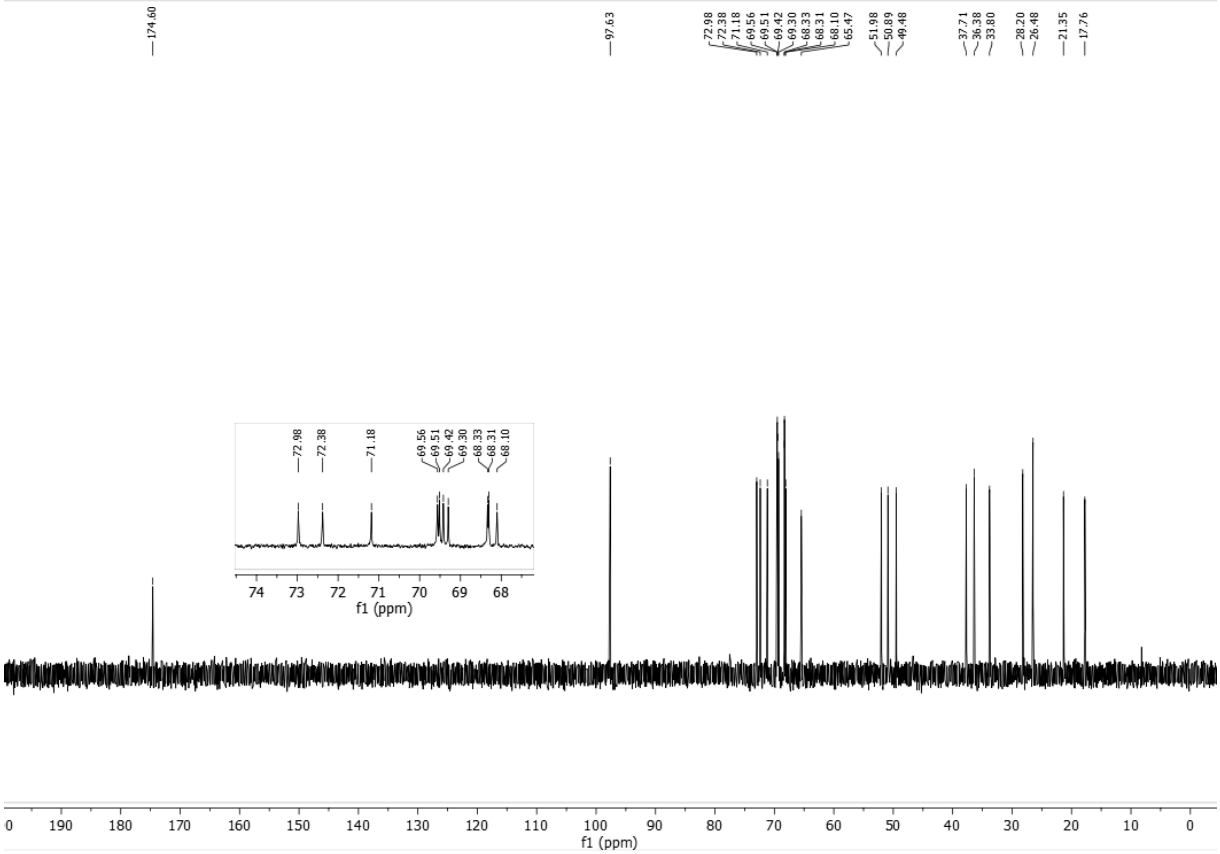


9-Aza-4-thia-4,4,8,24-tetraoxo-13,16,19-trioxa-tricosyl α-D-sulfoquinovoside (12; SQ-amine)

^{1}H NMR

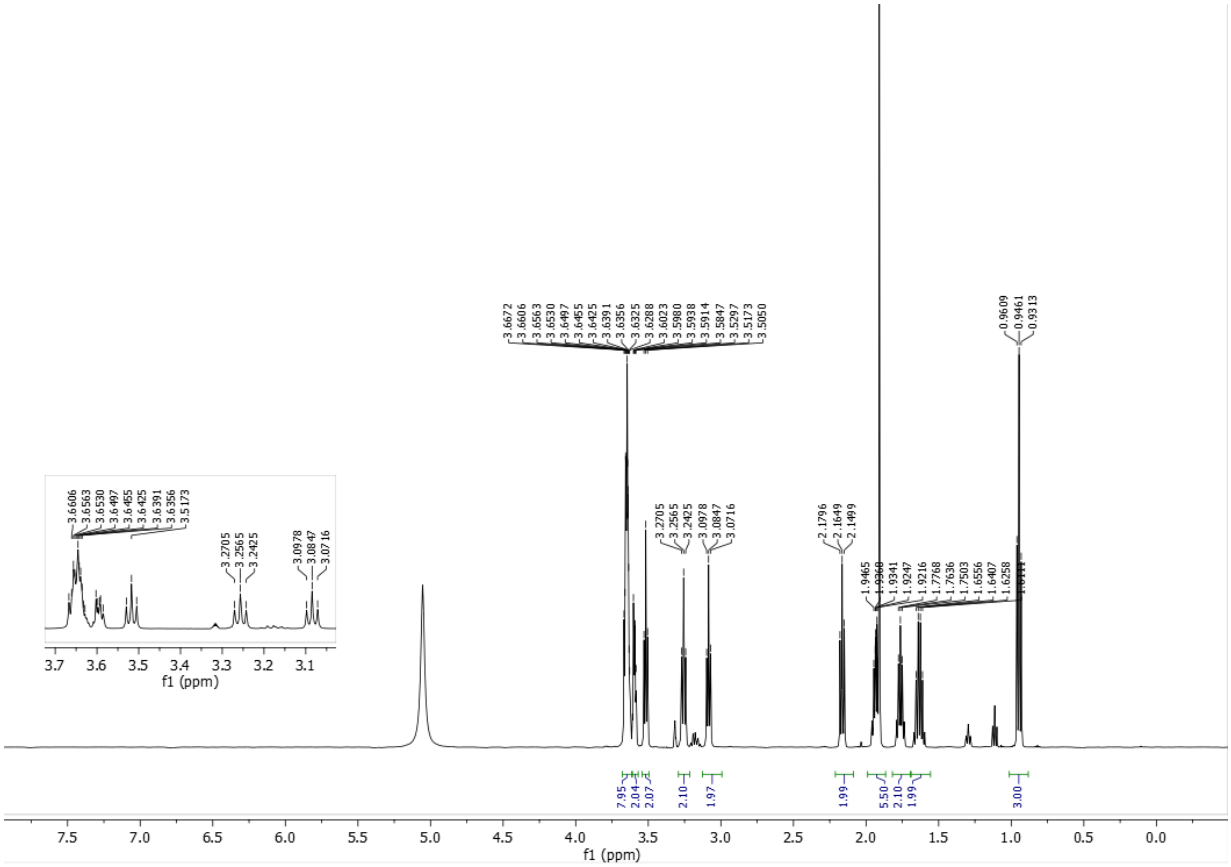


^{13}C NMR



N-Butyryl-4,7,10-trioxa-1,13-tridecanediamine (13)

^{1}H NMR



^{13}C NMR