Towards broad spectrum activity-based glycosidase probes: synthesis and evaluation of deoxygenated cyclophellitol aziridines

Sybrin P. Schröder,[a] Jasper W. van de Sande,[a] Wouter W. Kallemeijn,[b] Chi-Lin Kuo,[b] Marta Artola,[a] Eva J. van Rooden,[a] Jianbing Jiang,[a] Thomas J. M. Beenakker,[a] Bogdan I. Florea,[a] Wendy A. Offen,[c] Gideon J. Davies,[c] Adriaan J. Minnaard,[d] Johannes M. F. G. Aerts,[b] Jeroen D. C. Codée,[a] Gijsbert A. van der Marel[a] and Herman S. Overkleeft[a*]

[a] Department of Bioorganic Synthesis, Leiden Institute of Chemistry, Leiden University, Einsteinweg 55, 2333 CC Leiden, The Netherlands.
[b] Department of Medical Biochemistry, Leiden Institute of Chemistry.
[c] Department of Chemistry, University of York, Heslington, York, YO10 5DD, U.K.
[d] Stratingh Institute for Chemistry, University of Groningen, Nijenborgh 7, 9747 AG Groningen, The Netherlands.
* Corresponding author.

E-mail: h.s.overkleeft@chem.leidenuniv.nl

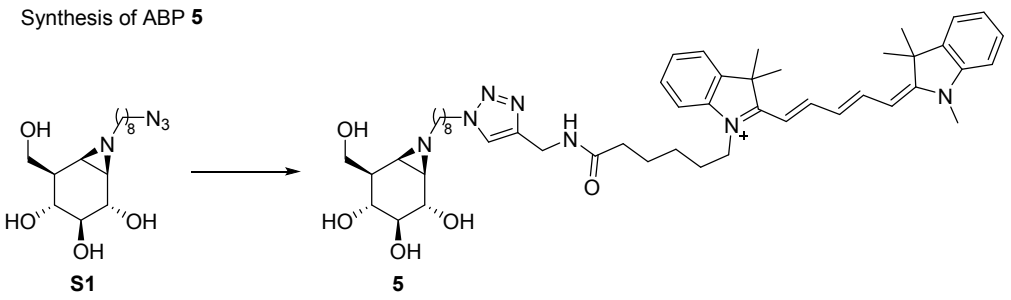
Table of Contents

Experimental procedures and characterization data

General: Chemicals were purchased from Acros, Sigma Aldrich, Biosolve, VWR, Fluka, Merck and Fisher Scientific and used as received unless stated otherwise. Tetrahydrofuran (THF), *N,N*-dimethylformamide (DMF) and toluene were stored over molecular sieves before use. Traces of water from reagents were removed by co-evaporation with toluene in reactions that required anhydrous conditions. All reactions were performed under an argon atmosphere unless stated otherwise. TLC analysis was conducted using Merck aluminum sheets (Silica gel 60 F_{254}) with detection by UV absorption (254 nm), by spraying with a solution of $(NH_4)_6Mo_7O_{24}\cdot 4H_2O$ (25 g/L) and $(NH_4)_4Ce(SO_4)_4\cdot 2H_2O$ (10 g/L) in 10% sulfuric acid or a solution of $KMnO_4$ (20 g/L) and K_2CO_3 (10 g/L) in water, followed by charring at ~150 ˚C. Column chromatography was performed using Screening Device b.v. silica gel (particle size of 40 – 63 µm, pore diameter of 60 Å) with the indicated eluents. ^{1}H NMR and ^{13}C NMR spectra were recorded on a Brüker AV-400 (400 and 101 MHz respectively), Brüker AV-500 (500 and 125 MHz respectively) or Brüker AV-600 (600 and 150 MHz respectively) spectrometer in the given deuterated solvent. Chemical shifts are given in ppm (δ) relative to the residual solvent peak or tetramethylsilane (0 ppm) as internal standard. Coupling constants are given in Hz. High-resolution mass spectrometry (HRMS) analysis was performed with a LTQ Orbitrap mass spectrometer (Thermo Finnigan), equipped with an electronspray ion source in positive mode (source voltage 3.5 kV, sheath gas flow 10 mL/min, capillary temperature 250 °C) with resolution R = 60000 at *m/z* 400 (mass range *m/z* = 150 – 2000) and dioctyl phthalate (*m/z* = 391.28428) as a "lock mass", or with a Synapt G2-Si (Waters) , equipped with an electronspray ion source in positive mode (ESI-TOF), injection via NanoEquity system (Waters), with LeuEnk (*m/z* = 556.2771) as "lock mass". Eluents used: MeCN:H2O (1:1 v/v) supplemented with 0.1% formic acid. The high-resolution mass spectrometers were calibrated prior to measurements with a calibration mixture (Thermo Finnigan).

General procedure for click reactions

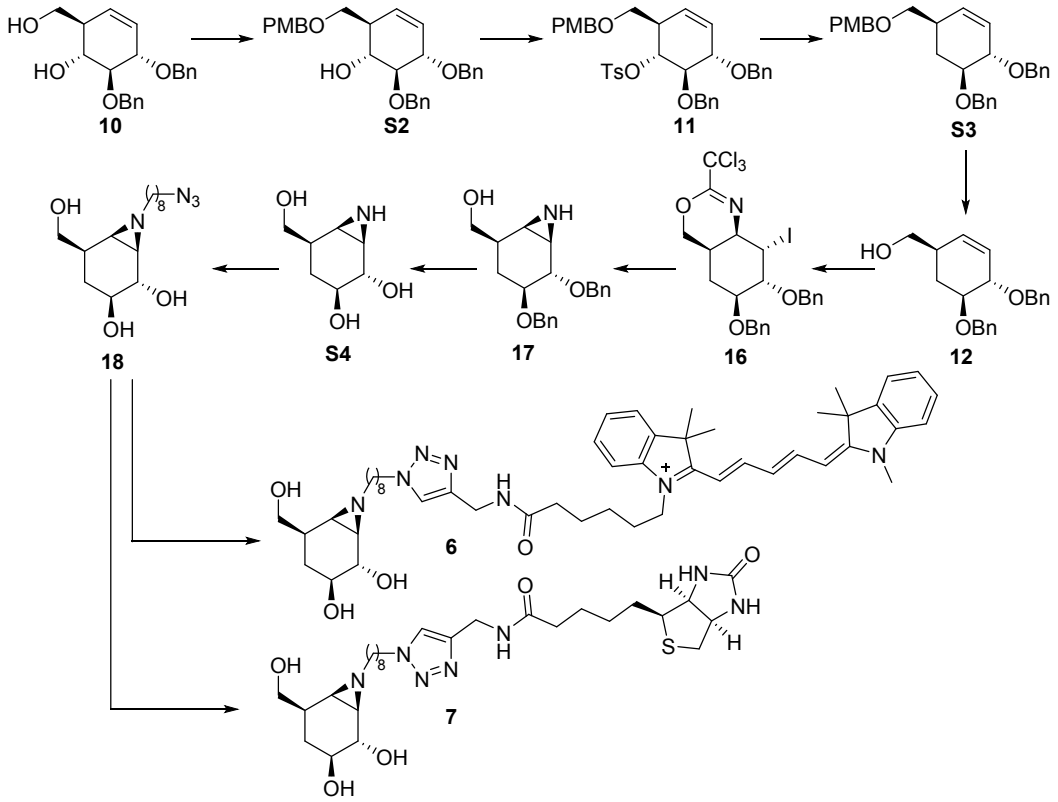
The azido compound was dissolved in DMF (~10 mg/mL), then Cy5-alkyne[1] or biotin-alkyne[2] (1.1 eq), $CuSO_4$ (0.2 eq) and sodium ascorbate (0.4 eq) were added and the mixture was stirred for 16 h at rt. The reaction mixture was concentrated, purified by semi-preparative reversed phase HPLC (linear gradient. Solutions used: A: 50 mM NH_4HCO_3 in H_2O, B: acetonitrile) and lyophilized.



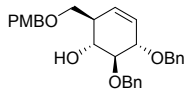
Compound 5

Following the general procedure, reaction of compound **S1**[3] (20 mg, 61 µmol) with Cy5-alkyne afforded the title compound as a blue solid (6.9 mg, 13%). ^{1}H-NMR (600 MHz, CD_3OD): δ 8.28 – 8.22 (m, 2H), 7.84 (s, 1H), 7.50 (d, *J* = 7.5 Hz, 2H), 7.44 – 7.39 (m, 2H), 7.32 – 7.25 (m, 4H), 6.62 (t, *J* = 12.4 Hz, 1H), 6.28 (dd, *J* = 13.7, 2.6 Hz, 2H), 4.45 – 4.32 (m, 4H), 4.09 (t, *J* = 7.5 Hz, 2H), 3.99 (dd, *J* = 10.1, 4.4 Hz, 1H), 3.69 – 3.54 (m, 5H), 3.12 – 3.09 (m, 1H), 3.00 (t, *J* = 9.8 Hz, 1H), 2.35 – 2.29 (m, 1H), 2.25 (t, *J* = 7.3 Hz, 2H), 2.14 – 2.09 (m, 1H), 1.99 – 1.97 (m, 1H), 1.93 (s, 6H), 1.90 – 1.79 (m, 4H), 1.73 – 1.68 (m, 10H), 1.63 (d, *J* = 6.3 Hz, 1H), 1.59 – 1.42 (m, 4H), 1.39 – 1.23 (m, 8H) ppm. ^{13}C-NMR (150 MHz, CD_3OD): δ 175.7, 175.4, 174.6, 155.6, 155.5, 144.2, 143.5, 142.6, 142.5, 129.8, 129.7, 126.6, 126.3, 126.2, 124.2, 123.4, 123.3, 112.0, 111.9, 104.4, 104.2, 79.0, 73.9, 70.1, 63.7, 62.1, 51.3, 50.5, 45.5, 45.5, 44.8, 43.0, 36.5, 35.6, 31.5, 31.3, 30.4, 30.3, 29.9, 28.2, 28.1, 27.9, 27.8, 27.4, 27.3, 26.4 ppm. HRMS (ESI) *m/z*: [M+] calc for $C_{50}H_{70}N_7O_5^+$ 848.54329, found 848.54304.

Synthesis of 4-deoxy ABPs **6**-**7**

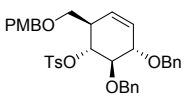


Compound S2



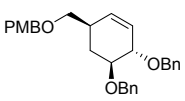
Diol **10**[4] (1.36 g, 4.0 mmol) was dissolved in dry MeCN (20 mL), then PMBCl (0.81 mL, 6.0 mmol), KI (664 mg, 4.0 mmol), K_2CO_3 (608 mg, 4.4 mmol) and 2-aminoethyl diphenylborinate (90 mg, 0.4 mmol) were added and the mixture was stirred at 60 °C for 4 h. Then, the mixture was diluted with EtOAc (200 mL), washed with H_2O (2 x 100 mL) and brine, dried over $MgSO_4$, filtrated and concentrated. Flash purification by silica column chromatography (pentane/EtOAc, 4:1) gave the title compound as a colorless oil (1.75 g, 95%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.39 – 7.26 (m, 10H), 7.24 (d, *J* = 8.6 Hz, 2H), 6.87 (d, *J* = 8.6 Hz, 2H), 5.73 (dt, *J* = 10.2, 2.3 Hz, 1H), 5.60 (d, *J* = 10.2 Hz, 1H), 4.99 (d, *J* = 11.3 Hz, 1H), 4.79 (d, *J* = 11.3 Hz, 1H), 4.67 (q, *J* = 11.5 Hz, 2H), 4.46 (s, 2H), 4.21 – 4.17 (m, 1H), 3.79 (s, 3H), 3.74 – 3.61 (m, 2H), 3.61 – 3.51 (m, 2H), 2.95 (brs, OH), 2.57 – 2.48 (m, 1H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ 159.3, 138.8, 138.4, 130.3, 129.4, 128.6, 128.6, 128.3, 128.1, 128.0, 127.9, 127.8, 126.8, 113.9, 84.0, 80.3, 75.0, 73.1, 71.7, 71.4, 70.9, 55.4, 44.1 ppm. HRMS (ESI): *m/z* = $[M+Na]^+$ calc for $C_{29}H_{32}O_5$ 483.21420, found 483.21368.

Compound 11



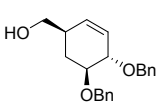
Cyclohexene **S2** (2.54 g, 5.51 mmol) was co-evaporated with toluene and dissolved in pyridine (55 mL). Tosyl chloride (10.5 g, 55.1 mmol) was added in portions over 10 minutes, and then the mixture was stirred at 60 °C for 16 h. The mixture was cooled to 0 °C, quenched with H_2O (20 mL) and diluted with EtOAc (200 mL). The organic phase was washed with aq. 1N HCl (3 x 100 mL), H_2O (100 mL), sat. aq. $NaHCO_3$ (100 mL) and brine, dried over $MgSO_4$, filtered and concentrated. Flash purification by silica column chromatography (pentane/EtOAc, 7:1 → 5:1) gave the title compound as a colorless oil (3.07 g, 90%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.73 (d, *J* = 8.3 Hz, 2H), 7.31 – 7.18 (m, 10H), 7.16 (d, *J* = 5.6 Hz, 2H), 7.07 (d, *J* = 8.0 Hz, 2H), 6.90 – 6.84 (m, 2H), 5.70 (dt, *J* = 10.2, 2.4 Hz, 1H), 5.61 (dt, *J* = 10.2, 1.8 Hz, 1H), 5.00 – 4.92 (m, 1H), 4.65 (d, *J* = 11.4 Hz, 1H), 4.56 – 4.48 (m, 3H), 4.44 (d, *J* = 11.5 Hz, 1H), 4.32 (d, *J* = 11.5 Hz, 1H), 4.23 – 4.17 (m, 1H), 3.80 (s, 3H), 3.73 (dd, *J* = 10.1, 7.5 Hz, 1H), 3.52 (dd, *J* = 9.3, 3.4 Hz, 1H), 3.41 (dd, *J* = 9.3, 5.4 Hz, 1H), 2.71 – 2.63 (m, 1H), 2.29 (s, 3H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ 159.2, 144.3, 138.5, 138.1, 134.7, 130.4, 129.5, 129.4, 128.5, 128.1, 128.1, 128.0, 127.8, 127.8, 127.5, 127.3, 127.0, 113.8, 81.4, 80.4, 80.3, 74.6, 72.9, 72.3, 68.8, 55.4, 43.6, 21.7 ppm. HRMS (ESI): *m/z* = $[M+Na]^+$ calc for $C_{36}H_{38}O_7S$ 637.22305, found 637.22318.

Compound S3



Tosylate **11** (3.07 g, 5.0 mmol) was co-evaporated with toluene (3x) and subsequently dissolved in dry THF (100 mL). The mixture was cooled to 0 °C and $LiAlH_4$ (2.4 M in THF, 15.6 mL, 37.4 mmol) was added and the mixture was refluxed overnight. After cooling to 0 °C, the reaction was quenched by slow addition of EtOAc (100 mL). An aqueous solution of Rochelle's salt (33 wt%, 100 mL) was added and the mixture was stirred vigorously for 1 h. Then, the organic phase was separated and the water phase was extracted with EtOAc (2x 100 mL). The combined organic layers were washed with brine, dried over $MgSO_4$, filtered and concentrated. Flash purification by silica column chromatography (pentane/EtOAc, 8:1) gave the title compound as a colorless oil (1.91 g, 86%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.42 – 7.18 (m, 12H), 6.88 (d, *J* = 7.9 Hz, 2H), 5.69 (s, 2H), 4.78 – 4.64 (m, 4H), 4.44 (s, 2H), 4.12 (d, *J* = 7.2 Hz, 1H), 3.80 (s, 3H), 3.69 (td, *J* = 9.1, 7.2, 3.2 Hz, 1H), 3.38 – 326 (m, 2H), 2.56 (s, 1H), 2.20 (d, *J* = 12.8 Hz, 1H), 1.38 (q, *J* = 11.8 Hz, 1H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ 159.3, 138.9, 130.6, 130.4, 129.4, 128.5, 128.5, 127.9, 127.8, 127.8, 127.6, 113.9, 79.7, 79.4, 73.9, 72.9, 72.2, 71.8, 55.4, 37.1, 30.9 ppm. HRMS (ESI): *m/z* = $[M+H]^+$ calc for $C_{29}H_{32}O_4$ 445.23734, found 445.23729.

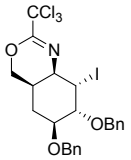
Compound 12



Compound **S3** (1.91 g, 4.29 mmol) was dissolved in a mixture of DCM/HFIP (1:1, 43 mL), and aq. HCl (12 M, 36 µL, 0.43 mmol) was added.[5] The mixture was stirred for 15 minutes and was subsequently quenched by addition of sat. aq. $NaHCO_3$ (10 mL). The mixture was diluted with DCM (200 mL) and washed with brine. The organic phase was dried over $MgSO_4$, filtered and concentrated. Flash purification by silica column chromatography

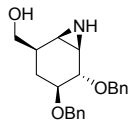
(pentane/EtOAc, 4:1 → 2:1) gave the title compound as a colorless oil (1.48 g, quant.). ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.39 – 7.24 (m, 10H), 5.81 (dt, *J* = 10.1, 2.5 Hz, 1H), 5.74 (d, *J* = 10.2 Hz, 1H), 4.76 – 4.63 (m, 4H), 4.06 (dq, *J* = 6.5, 2.5 Hz, 1H), 3.73 (ddd, *J* = 10.1, 6.4, 3.6 Hz, 1H), 3.62 – 3.52 (m, 2H), 2.51 – 2.42 (m, 1H), 2.20 (dt, *J* = 13.1, 4.6 Hz, 1H), 1.95 (brs, OH), 1.55 – 1.46 (m, 1H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ 138.8, 138.6, 130.7, 128.5, 128.5, 128.1, 127.9, 127.8, 127.8, 127.7, 78.1, 77.7, 72.0, 71.8, 66.6, 38.5, 29.9 ppm. HRMS (ESI): *m/z* = $[M+H]^+$ calc for $C_{21}H_{24}O_3$ 325.17982, found 325.17987.

Compound 16



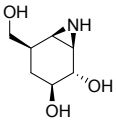
Compound **12** (1.48 g, 4.56 mmol) was dissolved in DCM (23 mL). Trichloroacetonitrile (687 µL, 6.85 mmol) and DBU (68 µL, 0.46 mmol) were added and the mixture was stirred overnight at rt. The mixture was concentrated, and filtered over a small silica plug. Concentration of the eluate afforded an oil (1.9 g) which was directly taken up in $CHCl_3$ (40 mL) and cooled to 0 °C. Then *N*-iodosuccinimide (1.37 g, 6.08 mmol) was added and the mixture was stirred overnight while the ice-bath was allowed to slowly reach rt. The mixture was diluted with $CHCl_3$ (150 mL), washed with aq. 10% $Na_2S_2O_3$ (100 mL) and brine. Flash purification by silica column chromatography (pentane/EtOAc, 20:1) gave the title compound as a colorless oil (2.08 g, 77% over 2 steps). ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.44 – 7.22 (m, 10H), 4.80 (s, 1H), 4.75 (dd, *J* = 11.5, 8.2 Hz, 2H), 4.64 (dd, *J* = 11.5, 4.0 Hz, 2H), 4.38 (dd, *J* = 11.1, 2.9 Hz, 1H), 4.30 (d, *J* = 11.2 Hz, 1H), 3.99 (t, *J* = 3.9 Hz, 1H), 3.82 (td, *J* = 10.5, 4.9 Hz, 1H), 2.91 – 2.63 (m, 2H), 1.97 (dt, *J* = 13.1, 4.4 Hz, 1H), 1.59 – 1.48 (m, 1H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ 152.8, 138.6, 137.9, 128.5, 128.2, 127.9, 127.8, 77.9 (broad, 2 x C), 73.4, 73.4, 72.2, 58.7, 35.7 (broad, assigned by HSQC), 29.6 (broad, assigned by HSQC), 27.0 (broad, assigned by HSQC) ppm. HRMS (ESI) *m/z*: $[M+H]^+$ calc for $C_{23}H_{23}Cl_3INO_3$ 592.97882, found 593.9883.

Compound 17



Compound **16** (35 mg, 59 µmol) was dissolved in MeOH/DCM (1:1, 1.2 mL) and HCl (1.25 M in MeOH, 443 µL, 0.55 mmol) was added. The mixture was stirred 24 h, and was subsequently neutralized by addition of Amberlite IRA-67. After stirring overnight, the reaction mixture was filtered and concentrated. Flash purification by silica column chromatography (DCM/MeOH, 49:1 → 19:1) gave the title compound as a colorless oil (16 mg, 80%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.43 – 7.21 (m, 10H), 4.77 (s, 2H), 4.65 (d, *J* = 4.0 Hz, 2H), 3.76 (dd, *J* = 10.4, 4.8 Hz, 1H), 3.69 (dd, *J* = 10.5, 6.1 Hz, 1H), 3.65 (d, *J* = 8.0 Hz, 1H), 3.43 (ddd, *J* = 11.9, 8.0, 3.7 Hz, 1H), 2.40 (dd, *J* = 5.7, 2.7 Hz, 1H), 2.28 (d, *J* = 6.0 Hz, 1H), 2.16 – 2.05 (m, 1H), 1.65 (dt, *J* = 12.9, 4.2 Hz, 1H), 1.27 (q, *J* = 12.6 Hz, 1H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ 138.9, 138.5, 128.5, 128.4, 127.9, 127.8, 127.7, 127.6, 80.5, 80.1, 72.8, 71.6, 66.0, 36.4, 32.8, 32.3, 24.0 ppm. TLC-MS (ESI) *m/z*: $[M+H]^+$ calc for $C_{21}H_{26}NO_3$ 340.18, found 340.3.

Compound S4

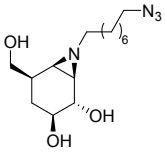


Ammonia (3 mL) was condensed in a flask at -60 °C, and lithium wire (7 mg, 0.94 mmol) was added. The resulting deep-blue solution was stirred for 30 minutes to dissolve all lithium. Aziridine **17** (16 mg, 47 µmol) was taken up in dry THF (1 mL) and added to the reaction mixture. After stirring for 1 h, the mixture was quenched with H_2O. The mixture was slowly warmed to rt and evaporated. The crude was dissolved in H_2O and eluted over a column packed with Amberlite CG-50 (NH_4^+) with 0.5M NH_4OH as eluent, affording the title compound as an oil (8 mg, quant.). ^{1}H NMR (400 MHz, D_2O): δ 3.42 (dd, *J* = 10.8, 7.9 Hz, 1H), 3.37 – 3.30 (m, 2H), 3.21 (ddd, *J* = 12.1, 8.4, 3.5 Hz, 1H), 2.24 (dd, *J* = 5.9, 2.9 Hz, 1H), 2.10 – 1.96 (m, 2H), 1.35 (dt, *J* = 12.7, 4.2 Hz, 1H), 0.68 (q, *J* = 12.4 Hz, 1H) ppm. ^{13}C NMR (101 MHz, D_2O): δ 73.1, 73.0, 64.2, 36.3, 35.0, 31.8, 26.8 ppm. HRMS (ESI) *m/z*: $[M+H]^+$ calc for $C_7H_{14}NO_3$ 160.09682, found 160.09690.

1-Azido-8-iodooctane

8-Chloro-1-octanol (10.5 g, 64 mmol) was dissolved in DMSO (16 mL), NaN_3 (6.2 g, 96 mmol) was added and the mixture was stirred overnight at 80 °C. The mixture was diluted with EtOAc (100 mL) and washed with H_2O (10 x 60 mL). The organic phase was dried over $MgSO_4$, filtrated and concentrated. The crude intermediate was taken up in DCM (200 mL), Et_3N (14.2 mL, 102 mmol) was added and the mixture was cooled to 0 °C. Then, MsCl (7.4 mL, 96 mmol) was added dropwise and the mixture was allowed to reach rt. After 1 h at rt, the mixture was quenched with H_2O (50 mL) and diluted with DCM (100 mL). The organic phase was washed with 1N HCl (3 x 100 mL), sat. aq. $NaHCO_3$ (1 x 100 mL) and brine, dried over $MgSO_4$, filtrated and concentrated. The crude intermediate was taken up in DMF (640 mL), KI (15.9 g, 96 mmol) was added and the mixture was stirred overnight at 70 °C. The solvent was evaporated, then the crude was diluted with H_2O (1 L) and extracted with Et_2O (3 x 100 mL). The combined organic fractions were washed with brine, dried over $MgSO_4$, filtrated, concentrated and flash purification by silica column chromatography (Et_2O/pentane, 1:99) gave the title compound as a pale yellow oil (13.8 g, 77%). ^{1}H-NMR (400 MHz, $CDCl_3$): δ 3.26 (t, *J* = 6.9 Hz, 2H), 3.19 (t, *J* = 7.0 Hz, 2H), 1.88 – 1.73 (m, 2H), 1.66 – 1.50 (m, 2H), 1.46 – 1.26 (m, 8H).

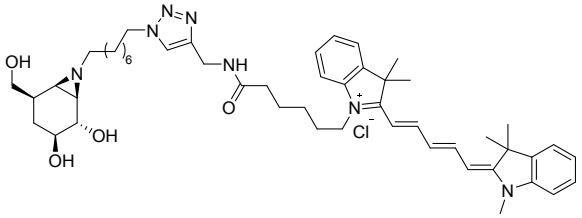
Compound 18



Starting from aziridine **17** (97 mg, 0.29 mmol), the compound was deprotected as described above and purified by elution over Amberlite CG-50 (NH_4^+). After evaporation of the solvent the intermediate was re-dissolved in DMF (1 mL), 1-azido-8-iodooctane (161 mg, 0.57 mmol) and K_2CO_3 (47 mg, 0.34 mmol) were added and the mixture was stirred overnight at 100 °C. The mixture was concentrated at 60 °C, and flash purification by silica column chromatography (DCM/MeOH, 19:1 → 9:1) gave the title compound as a colorless oil (47 mg, 53% over two steps). ^{1}H NMR (400 MHz, MeOD): 3.60 (dd, *J* = 10.2, 7.9 Hz, 1H), 3.52 – 3.44 (m, 2H), 3.28 (t, *J* = 6.8 Hz, 2H), 3.26 – 3.22 (m, 1H), 2.42 (dt, *J* = 11.5, 7.6 Hz, 1H), 2.15 – 2.01 (m, 2H), 1.84 (dd, *J* = 6.1, 3.2 Hz, 1H), 1.63 (d, *J* = 6.2 Hz, 1H), 1.62 – 1.53 (m, 4H), 1.50 (dt, *J* =

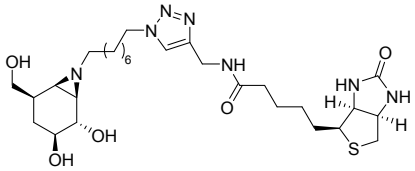
12.7, 4.5 Hz, 1H), 1.36 (m, 8H), 0.97 (q, *J* = 12.1 Hz, 1H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): 74.5, 74.4, 65.8, 61.9, 52.4, 45.7, 42.5, 38.6, 30.5, 30.3, 30.2, 29.9, 29.4, 28.3, 27.8 ppm. HRMS (ESI) *m/z*: $[M+H]^+$ calc for $C_{15}H_{29}N_4O_3$ 313.22342, found 313.22365.

Compound 6



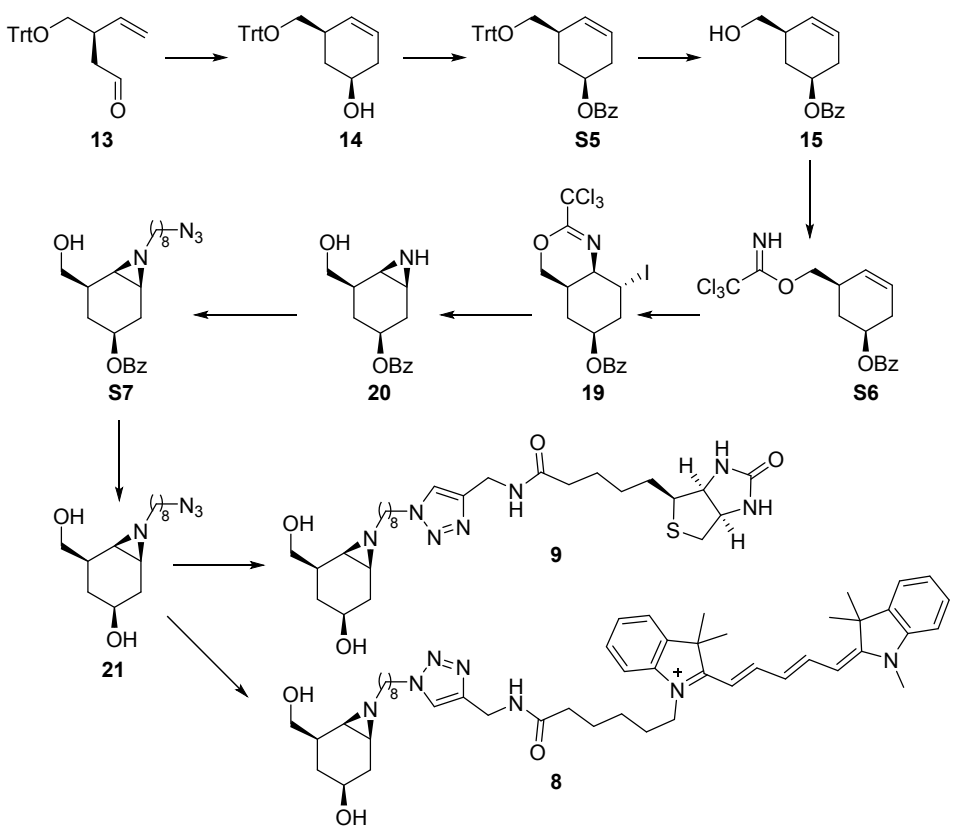
Following the general procedure, reaction of compound **18** (5.7 mg, 18.3 µmol) with Cy5-alkyne[1] afforded the title compound as a blue solid (6.6 mg, 42%). ^{1}H NMR (600 MHz, D_2O): δ 7.81 (s, 1H), 7.81 – 7.73 (m, 2H), 7.33 (dd, *J* = 12.8, 7.3 Hz, 2H), 7.26 (t, *J* = 7.5 Hz, 1H), 7.21 (t, *J* = 9.0 Hz, 2H), 7.14 – 7.07 (m, 2H), 7.05 (t, *J* = 7.2 Hz, 1H), 6.33 (t, *J* = 12.1 Hz, 1H), 6.03 (d, *J* = 13.6 Hz, 1H), 5.97 (d, *J* = 13.4 Hz, 1H), 4.38 (s, 2H), 4.23 (t, *J* = 6.6 Hz, 2H), 3.92 (s, 2H), 3.56 – 3.50 (m, 4H), 3.44 (dd, *J* = 10.5, 6.9 Hz, 1H), 3.34 (ddd, *J* = 12.2, 8.6, 3.5 Hz, 1H), 2.25 (q, *J* = 8.4, 6.5 Hz, 3H), 2.13 (dd, *J* = 6.6, 3.1 Hz, 1H), 1.92 (s, 3H), 1.85 (td, *J* = 11.0, 5.0 Hz, 1H), 1.72 (dd, *J* = 5.8, 3.0 Hz, 1H), 1.70 – 1.63 (m, 4H), 1.63 – 1.55 (m, 4H), 1.40 (d, *J* = 13.9 Hz, 12H), 1.32 (dd, *J* = 14.2, 7.5 Hz, 4H), 1.03 (s, 8H), 0.84 (q, *J* = 12.5 Hz, 1H) ppm. ^{13}C NMR (125 MHz, D_2O): δ 182.4, 176.7, 174.6, 173.5, 154.1, 153.7, 145.7, 143.5, 142.8, 142.0, 129.5, 126.2, 125.9, 125.3, 124.5, 123.3, 123.2, 111.7, 111.6, 104.2, 103.5, 73.7, 73.5, 65.2, 60.8, 51.2, 49.9, 49.8, 45.0, 44.5, 42.0, 37.3, 36.3, 35.2, 31.8, 30.4, 29.5, 29.4, 29.1, 28.9, 27.9, 27.8, 27.6, 27.5, 26.6, 26.4, 26.1 ppm. HRMS (ESI) *m/z*: $[M]^+$ calc for $C_{50}H_{70}N_7O_4^+$ 832.54838, found 832.54865.

Compound 7

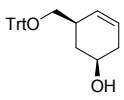


Following the general procedure, reaction of compound **18** (5.7 mg, 18.3 µmol) with biotin-alkyne[2] afforded the title compound as a white solid (5.3 mg, 49%). ^{1}H NMR (600 MHz, D_2O): 7.89 (s, 1H), 4.61 (dd, *J* = 7.9, 4.9 Hz, 1H), 4.45 (d, *J* = 6.8 Hz, 2H), 4.43 – 4.35 (m, 3H), 3.66 (dd, *J* = 10.6, 7.6 Hz, 1H), 3.56 (d, *J* = 8.6 Hz, 1H), 3.51 (dd, *J* = 10.6, 6.7 Hz, 1H), 3.39 (ddd, *J* = 12.3, 8.6, 3.7 Hz, 1H), 3.28 (dt, *J* = 10.0, 5.2 Hz, 1H), 2.99 (dd, *J* = 13.1, 5.0 Hz, 1H), 2.78 (d, *J* = 13.0 Hz, 1H), 2.44 (ddd, *J* = 11.4, 9.7, 6.5 Hz, 1H), 2.30 (t, *J* = 7.1 Hz, 2H), 2.25 – 2.17 (m, 1H), 2.13 (ddd, *J* = 11.7, 9.6, 5.3 Hz, 1H), 1.94 (dd, *J* = 6.2, 3.2 Hz, 1H), 1.92 (d, *J* = 1.1 Hz, 1H), 1.89 (p, *J* = 6.9 Hz, 2H), 1.76 (d, *J* = 6.3 Hz, 1H), 1.68 (dtd, *J* = 18.8, 9.8, 4.9 Hz, 2H), 1.64 – 1.58 (m, 2H), 1.58 – 1.43 (m, 3H), 1.36 – 1.16 (m, 10H), 0.88 (q, *J* = 12.5 Hz, 1H) ppm. ^{13}C NMR (125 MHz, D_2O): 177.6, 166.3, 145.7, 124.8, 73.7, 73.6, 65.3, 63.0, 61.2, 60.7, 56.4, 51.4, 45.0, 41.9, 40.7, 37.3, 36.2, 35.3, 30.2, 29.4, 29.3, 28.9, 28.8, 28.6, 28.6, 27.3, 26.4, 26.1 ppm. HRMS (ESI) *m/z*: $[M+H]^+$ calc for $C_{28}H_{48}N_7O_5S$ 594.34321, found 594.34302.

Synthesis of 2,4-deoxy ABPs **8**-**9**



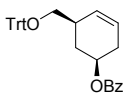
Compound 14



Aldehyde **13**[6] (600 mg, 1.68 mmol) was dissolved in dry THF (20 mL) and cooled to -90 °C. A solution of 1M (+)-Ipc_2B(allyl)borane solution in pentane was added to dry THF (10 mL), cooled to -90°C, and added dropwise via cannula. The mixture was stirred for 1 h, quenched with MeOH (6 mL) and diluted with sodium-phosphate buffer (pH 7, 0.5M, 12 mL). Hydrogen peroxide (13.2 mL, 30%) was added dropwise and the mixture was warmed to rt and stirred for 1 h. Next, the mixture was diluted with sat. aq. $NaHCO_3$ (100 mL) and extracted with EtOAc (3 x 75 mL). The combined organic layers were washed with water and brine, dried over $MgSO_4$, filtrated and concentrated. The crude product was dissolved in DCM (15 mL) and Grubb's II catalyst (71 mg, 5 mol%) was added. The mixture was refluxed overnight under argon. The product was purified by flash chromatography (pentane/EtOAc; 8:1 → 5:1) to afford the title product as a colorless oil (518 mg, 77%, *dr* 86:14). ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.49 – 7.40 (m, 6H), 7.33 – 7.18 (m, 9H), 5.88 – 5.52 (m, 2H), 3.94 (dddd, *J* = 11.1, 9.1, 5.5, 3.5 Hz, 1H), 3.12 – 2.94 (m, 2H), 2.67 – 2.52 (m, 1H), 2.37 (d, *J* = 16.8 Hz, 1H), 2.21 – 2.10 (m, 1H), 2.04 – 1.87 (m, 1H), 1.80 (m, OH), 1.45 – 1.12 (m, 1H). ^{13}C NMR (101 MHz, $CDCl_3$): δ 144.3, 128.8, 128.6, 127.9, 127.0, 125.3, 86.5, 67.6, 67.5, 37.6,

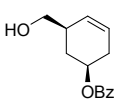
36.4, 35.0. $[\alpha]_D^{20}$ +33.0 (c 0.2, DCM). IR (ATR, cm^{-1}) 3350, 1489, 1448, 1066. HRMS: [M+H^+] calc for $C_{26}H_{26}O_2$ 371.20110 found 371.31553.

Compound S5



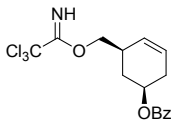
Compound **14** (675 mg, 1.69 mmol) was co-evaporated with toluene, dissolved in dry DCM (17 mL) and cooled to 0°C. Triethyl amine (1.2 mL, 8.5 mmol), benzoyl chloride (0.39 mL, 3.39 mmol) and DMAP (10 mg, 0.09 mmol) were added and the mixture was stirred overnight at rt. The reaction was quenched with sat. aq. $NaHCO_3$ (20 mL), stirred for 30 minutes and then extracted with DCM (3 x 20 mL). The combined organic layers were dried over $MgSO_4$, filtrated and concentrated. The product was purified by flash chromatography (pentane/EtOAc; 50:1) to afford the title product as a single diastereoisomer as a colorless oil (672 mg, 84%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 8.12 – 7.93 (m, 2H), 7.60 – 7.50 (m, 1H), 7.49 – 7.37 (m, 7H), 7.33 – 7.17 (m, 10H), 5.68 (s, 2H), 5.33 – 5.20 (m, 1H), 3.10 (dd, *J* = 8.6, 6.5 Hz, 1H), 3.00 (dd, *J* = 8.5, 6.8 Hz, 1H), 2.74 (s, 1H), 2.62 – 2.48 (m, 1H), 2.32 – 2.23 (m, 1H), 2.24 – 2.14 (m, 1H), 1.55 (q, *J* = 11.6 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$): δ 166.2, 144.3, 132.9, 130.8, 129.7, 128.9, 128.8, 128.4, 127.9, 127.0, 124.7, 86.4, 70.9, 67.4, 37.2, 32.4, 31.5. $[\alpha]_D^{20}$ +36.0 (c 0.2, DCM). IR (ATR, cm^{-1}) 1714, 1448, 1273, 1070. HRMS: [M+Na^+] calc for $C_{33}H_{30}O_3$ 497.20926 found 497.20825.

Compound 15



Compound **S5** (660 mg, 1.39 mmol) was dissolved in a mixture of DCM (7 mL) and MeOH (7 mL). Then, CSA (16 mg, 0.07 mmol) was added and the mixture was stirred for 5 h at rt. The reaction was quenched with Et_3N (0.15 mL) and concentrated. The product was purified by flash chromatography (pentane/EtOAc; 4:1) to afford the title product as a colorless oil (304 mg, 94%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 8.04 (d, *J* = 7.4 Hz, 2H), 7.55 (t, *J* = 7.4 Hz, 1H), 7.43 (t, *J* = 7.7 Hz, 2H), 5.76 (m, 1H), 5.66 (d, *J* = 10.2 Hz, 1H), 5.35 – 5.20 (m, 1H), 3.71 – 3.55 (m, 2H), 2.69 – 2.50 (m, 2H), 2.23 (m, 2H), 2.14 – 1.93 (m, 1H), 1.64 (q, *J* = 11.4 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$): δ 166.2, 133.0, 130.6, 129.6, 128.4, 128.0, 125.9, 70.8, 66.6, 39.0, 31.4, 31.3. $[\alpha]_D^{20}$ +17.3 (c 0.3, DCM). IR (ATR, cm^{-1}) 3400, 1712, 1273, 1114. HRMS: [M+Na^+] calc for $C_{14}H_{16}O_3Na$ 255.0992 found 255.1001.

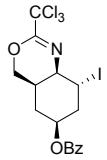
Compound S6



Compound **15** (290 mg, 1.25 mmol) was dissolved in dry DCM (12 mL). Then trichloroacetonitrile (0.25 mL, 2.5 mmol) and DBU (9.3 µL, 0.06 mmol) were added and the mixture was stirred overnight. The mixture was concentrated and the product was purified by flash chromatography (pentane/EtOAc; 20:1) to afford the title product as a colorless oil (429 mg, 91%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 8.30 (s, 1H), 8.05 (d, *J* = 7.5 Hz, 2H), 7.56 (t, *J* = 7.4 Hz, 1H), 7.45 (d, *J* = 7.7 Hz, 2H), 5.78 (ddt, *J* = 10.0, 4.9, 2.5 Hz, 1H), 5.69 (d, *J* = 10.2 Hz, 1H), 5.30 (dddd, *J* = 12.0, 9.2, 5.7, 3.5 Hz, 1H), 4.26 (d, *J* = 6.8 Hz, 2H), 2.95 (m, 1H), 2.67 – 2.50 (m, 1H), 2.35 – 2.16 (m, 2H), 1.67 (q, *J* = 11.3 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$): δ 166.1, 162.9,

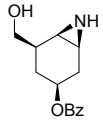
133.0, 130.5, 129.7, 128.4, 126.9, 126.0, 72.5, 70.2, 35.6, 31.5, 31.2. $[\alpha]_D^{20}$ +27.3 (c 0.3, DCM). IR (ATR, cm^{-1}) 1714, 1664, 1271, 1070, 794, 709. HRMS: [M+H$^+$] calc for $C_{16}H_{16}Cl_3NO_3$ 376.02740 found 376.02706.

Compound 19



Compound **S6** (561 mg, 1.49 mmol) was co-evaporated with toluene, dissolved in dry $CHCl_3$ (15 mL) and cooled to 0°C. Then, NIS (570 mg, 2.23 mmol) was added and the mixture was stirred overnight at rt. The reaction was quenched with aq. $Na_2S_2O_3$ (10 mL) and stirred for 10 minutes. The mixture was diluted with sat. aq. $NaHCO_3$ (40 mL) and extracted with CHCl3 (3x 30 mL). The combined organic layers were dried over $MgSO_4$, filtrated and concentrated. The product was purified by flash chromatography (pentane/Et_2O; 10:1) to afford the title product as a white foam (714 mg, 96%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 8.07 – 7.95 (m, 2H), 7.63 – 7.53 (m, 1H), 7.44 (dd, *J* = 8.4, 7.1 Hz, 2H), 5.49 (tt, *J* = 11.0, 4.2 Hz, 1H), 4.85 (q, *J* = 3.3 Hz, 1H), 4.47 (dd, *J* = 11.1, 3.1 Hz, 1H), 4.34 (dd, *J* = 11.2, 1.8 Hz, 1H), 3.92 (t, *J* = 3.6 Hz, 1H), 2.92 – 2.76 (m, 1H), 2.30 (dd, *J* = 14.1, 3.0 Hz, 1H), 2.24 – 2.11 (m, 1H), 1.92 (ddd, *J* = 14.4, 11.0, 3.7 Hz, 1H), 1.55 (q, *J* = 12.8 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$): δ 165.9, 152.9, 133.2, 130.1, 129.7, 128.5, 71.8, 70.4, 57.4, 34.8, 29.6, 28.5, 27.0. $[\alpha]_D^{20}$ +6.0 (c 0.3, DCM). IR (ATR, cm^{-1}) 1714, 1670, 1273, 1111, 821, 711. HRMS: [M+H$^+$] calc for $C_{16}H_{15}Cl_3INO_3$ 501.92406 found 501.92348.

Compound 20



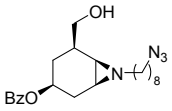
Compound **19** (51 mg, 0.1 mmol) was dissolved in a mixture of DCM (0.5 mL) and MeOH (0.5 mL) and cooled to 0°C. A solution of 1.25M HCl in MeOH (176 µL, 0.22 mmol) was added, and the mixture was stirred 16 h at rt. Then, Amberlite IRA-67 was added until neutral pH and the mixture was stirred for 2 h. The suspension was filtrated and the resin was washed with MeOH (3x). The organic phase was evaporated on an ice-bath. The product was purified by flash chromatography (DCM/MeOH; 16:1) using neutralized SiO_2, and the fractions were evaporated on an ice bath to afford the title product as a colorless oil (23 mg, 93%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 8.01 (m, 2H), 7.55 (m, 1H), 7.48 – 7.38 (m, 3H), 4.97 (dddd, *J* = 12.1, 10.5, 6.7, 3.6 Hz, 1H), 3.85 (dd, *J* = 10.5, 4.4 Hz, 1H), 3.75 (dd, *J* = 10.5, 5.9 Hz, 1H), 2.48 (dtd, *J* = 13.9, 7.0, 1.9 Hz, 1H), 2.40 (dd, *J* = 6.1, 3.1 Hz, 1H), 2.32 (t, *J* = 6.5 Hz, 1H), 2.31 – 2.23 (m, 1H), 1.80 – 1.69 (m, 2H), 1.55 (q, *J* = 12.3 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$): δ 166.1, 133.0, 129.7, 128.5, 128.4, 70.7, 66.2, 36.4, 31.5, 30.1, 27.0, 27.0. $[\alpha]_D^{25}$ +21.6 (c 0.5, DCM). IR (ATR, cm^{-1}) 3300, 1708, 1273, 1112. HRMS: [M+H$^+$] calc for $C_{14}H_{17}NO_3$ 248.12867 found 248.12802. *NOTE: This compound is unstable at elevated temperatures, and must be concentrated on an ice-bath to minimize decomposition.*

8-Azidooctyl trifluoromethanesulfonate

To dry DCM (5.8 mL) was added 8-azidooctan-1-ol (100 mg, 0.58 mmol) and pyridine (57 µL, 0.70 mmol) and the mixture was cooled to -20 °C. Triflic anhydride (118 µL, 0.70 mmol) was

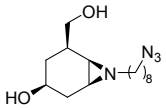
added and the mixture was stirred for 15 minutes. Then the mixture was diluted with DCM, and washed with cold water (3 x 10 mL). The organic layer was dried over $MgSO_4$, filtrated and concentrated at rt. The crude product was used directly in the alkylation of the aziridine. ^{1}H NMR (400 MHz, $CDCl_3$): δ 4.55 (t, *J* = 6.5 Hz, 2H), 3.27 (t, *J* = 6.9 Hz, 2H), 1.90 – 1.75 (m, 2H), 1.65 – 1.55 (m, 2H), 1.49 – 1.30 (m, 8H). ^{13}C NMR (101 MHz, $CDCl_3$): δ 77.8, 51.6, 29.4, 29.0, 28.9, 28.9, 26.7, 25.2 ppm.

Compound S7



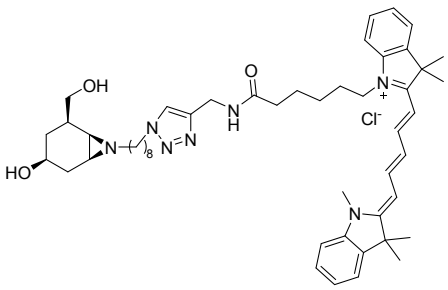
Compound **20** (54 mg, 0.22 mmol) was dissolved in dry THF (2 mL) and cooled to 0°C. DIPEA (42 µL, 0.24 mmol) was added and then 8-azidooctyl trifluoromethanesulfonate (73 mg, 0.24 mmol) in THF (0.5 mL). The mixture was stirred at 0 °C for 2 h and then diluted with water (20 mL) and brine (5 mL). The mixture was extracted with EtOAc (3 x 15 mL) and the combined organic layers were washed with brine, dried over $MgSO_4$, filtrated and concentrated. The product was purified by flash chromatography (pentane/EtOAc; 3:1) to afford the title product as a colorless oil (46 mg, 53%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 8.00 (dd, *J* = 8.3, 1.3 Hz, 2H), 7.62 – 7.48 (m, 1H), 7.42 (t, *J* = 7.7 Hz, 1H), 4.91 (dddd, *J* = 11.9, 10.5, 6.5, 3.5 Hz, 1H), 3.87 (dd, *J* = 10.3, 4.2 Hz, 1H), 3.79 – 3.68 (m, 1H), 3.26 (t, *J* = 6.9 Hz, 2H), 2.47 (m, 2H), 2.39 – 2.30 (m, 1H), 2.23 – 2.10 (m, 2H), 1.78 – 1.65 (m, 3H), 1.64 – 1.51 (m, 8H), 1.34 (m, 7H). ^{13}C NMR (101 MHz, $CDCl_3$): δ 166.0, 132.9, 130.5, 129.5, 128.3, 70.7, 66.2, 61.0, 51.4, 40.0, 36.1, 35.9, 29.8, 29.5, 29.4, 29.1, 28.8, 27.7, 27.3, 26.6. $[\alpha]_D^{25}$ +18.0 (c 0.4, DCM). IR (ATR, cm^{-1}) 3400, 2927, 2094, 1741, 1274, 1112. HRMS: [M+H$^+$] calc for $C_{22}H_{32}N_4O_3$ 401.25527 found 401.25370.

Compound 21



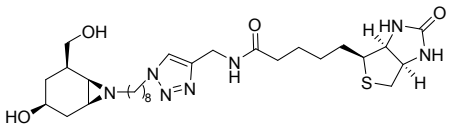
Compound **S7** (46 mg, 0.11 mmol) was dissolved in MeOH (1.1 mL), then 0.1M NaOMe in MeOH (690 µL, 0.07 mmol) was added and the mixture was stirred for 24 h at rt. Then the mixture was diluted with sat. aq. $NaHCO_3$ (10 mL), and extracted with EtOAc (5 x 10 mL). The combined organic layers were dried over $MgSO_4$, filtrated and concentrated. The product was purified by flash chromatography using neutralized SiO_2 (DCM/MeOH; 32:1 → 24:1) to afford the title product as a colorless oil (28 mg, 82%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 3.84 (dd, *J* = 10.4, 7.3 Hz, 1H), 3.79 (m, 1H), 3.74 (dd, *J* = 10.3, 5.8 Hz, 1H), 3.26 (t, *J* = 6.9 Hz, 2H), 2.24 (m, 3H), 1.98 (dd, *J* = 14.2, 5.1 Hz, 1H), 1.85 (dt, *J* = 14.1, 3.7 Hz, 1H), 1.75 (m, 2H), 1.66 – 1.42 (m, 6H), 1.44 – 1.20 (m, 8H). ^{13}C NMR (101 MHz, $CDCl_3$): δ 66.6, 66.6, 60.6, 51.4, 39.7, 37.9, 33.7, 30.9, 30.2, 29.4, 29.3, 29.0, 28.8, 27.2, 26.6. $[\alpha]_D^{25}$ +32.7 (c 0.3, MeOH). IR (ATR, cm^{-1}) 3350, 2927, 2091, 1245. HRMS: [M+H_3O^+] calc for $C_{15}H_{28}N_4O_2$ 315.2391 found 315.2399.

Compound 8



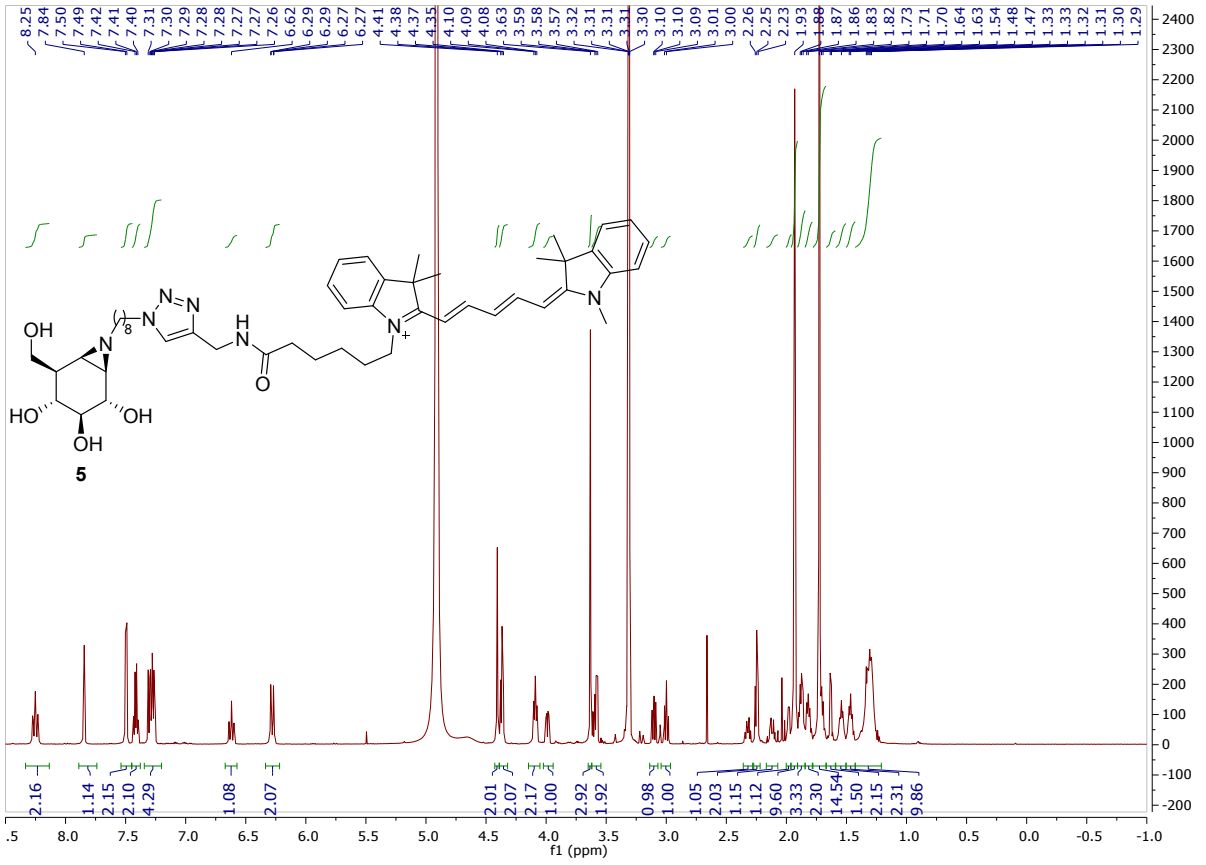
Following the general procedure starting from compound **21** (3.76 mg, 12.7 µmol), the product was obtained as a blue powder (3.2 mg, 30%). ^{1}H NMR (600 MHz, CD_3OD): δ 8.24 (td, *J* = 13.1, 2.9 Hz, 2H), 7.83 (s, 1H), 7.49 (d, *J* = 7.4 Hz, 2H), 7.41 (q, *J* = 7.3 Hz, 2H), 7.28 (m, 4H), 6.62 (t, *J* = 12.4 Hz, 1H), 6.28 (dd, *J* = 14.0, 2.1 Hz, 2H), 4.41 (s, 2H), 4.36 (t, *J* = 7.1 Hz, 2H), 4.09 (t, *J* = 7.5 Hz, 2H), 3.66 – 3.59 (m, 1H), 3.63 (s, 3H), 3.49 (m, 2H), 2.39 – 2.31 (m, 1H), 2.25 (t, *J* = 7.3 Hz, 2H), 2.22 – 2.17 (m, 1H), 2.07 – 2.00 (m, 2H), 1.87 (m, 2H), 1.83 (m, 2H), 1.73 (s, 12H), 1.71 – 1.65 (m, 4H), 1.55 – 1.45 (m, 6H), 1.33 – 1.27 (m, 8H), 0.97 (q, *J* = 11.7 Hz, 1H). ^{13}C NMR (151 MHz, CD_3OD): δ 175.7, 175.4, 174.6, 155.5, 155.5, 146.1, 144.2, 143.5, 142.6, 142.5, 136.0, 131.2, 129.8, 129.7, 128.8, 126.6, 126.3, 126.2, 124.1, 123.4, 123.3, 112.0, 111.8, 104.4, 104.2, 68.3, 66.3, 61.9, 51.3, 50.5, 44.8, 41.5, 38.8, 38.6, 36.5, 35.6, 33.8, 32.4, 31.5, 31.3, 31.1, 30.4, 30.2, 30.0, 28.3, 28.1, 27.9, 27.8, 27.4, 27.3, 26.4. HRMS: [M^+] calc for $C_{50}H_{70}N_7O_3$ 816.55401 found 816.55291.

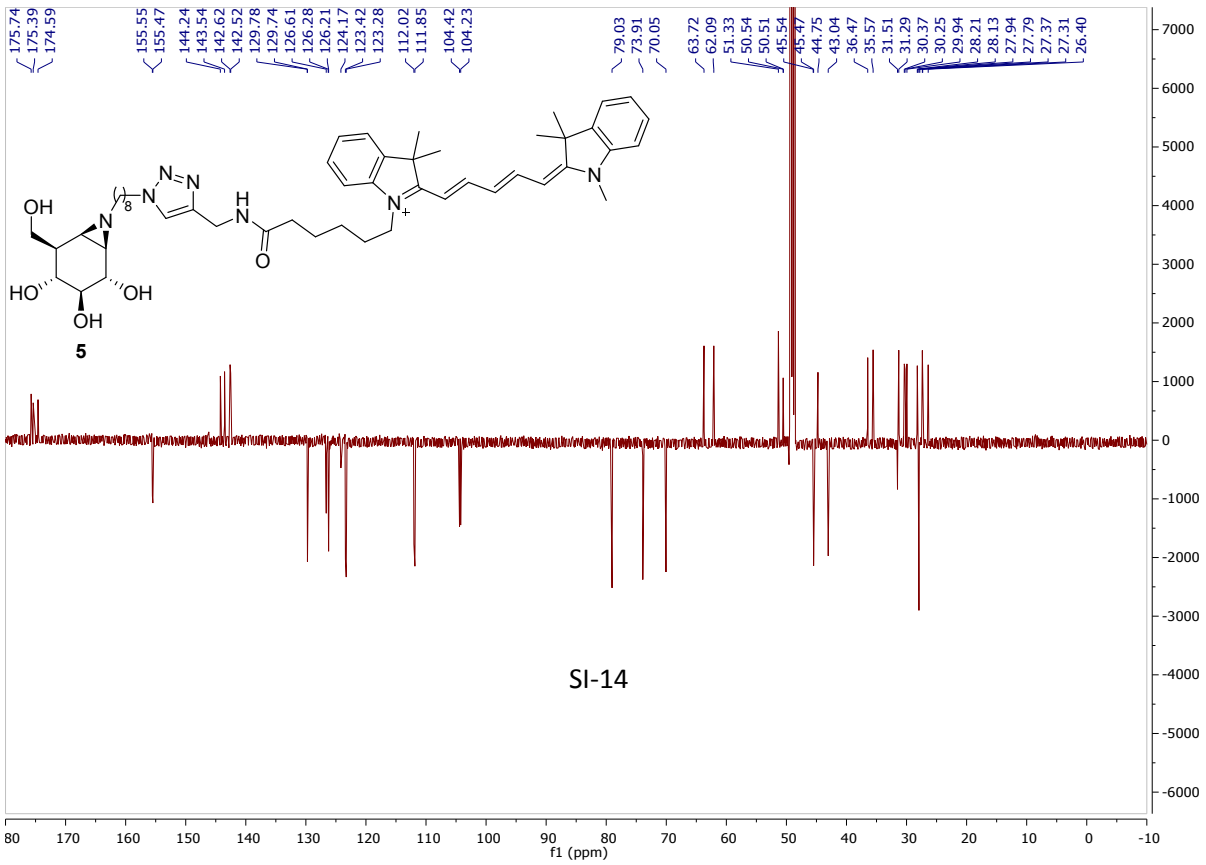
Compound 9

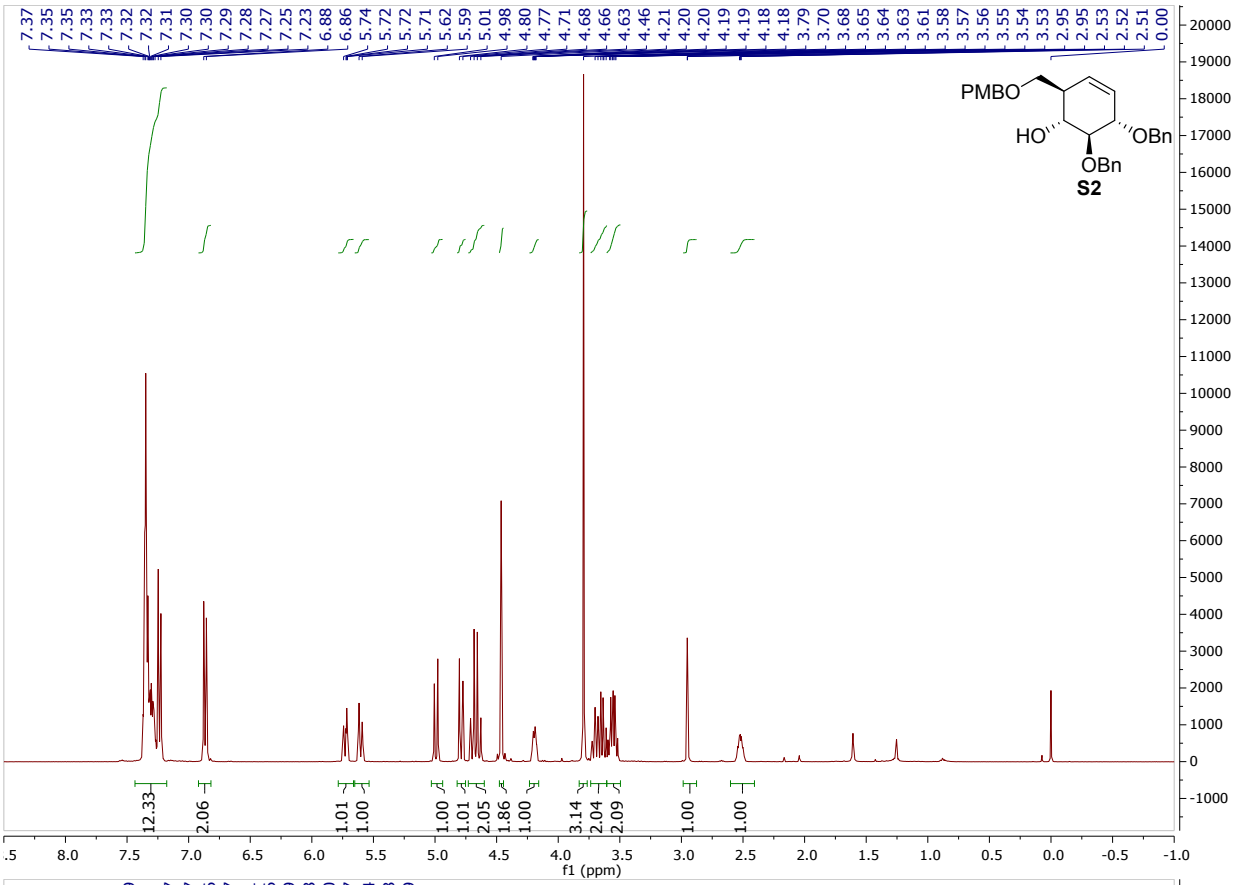


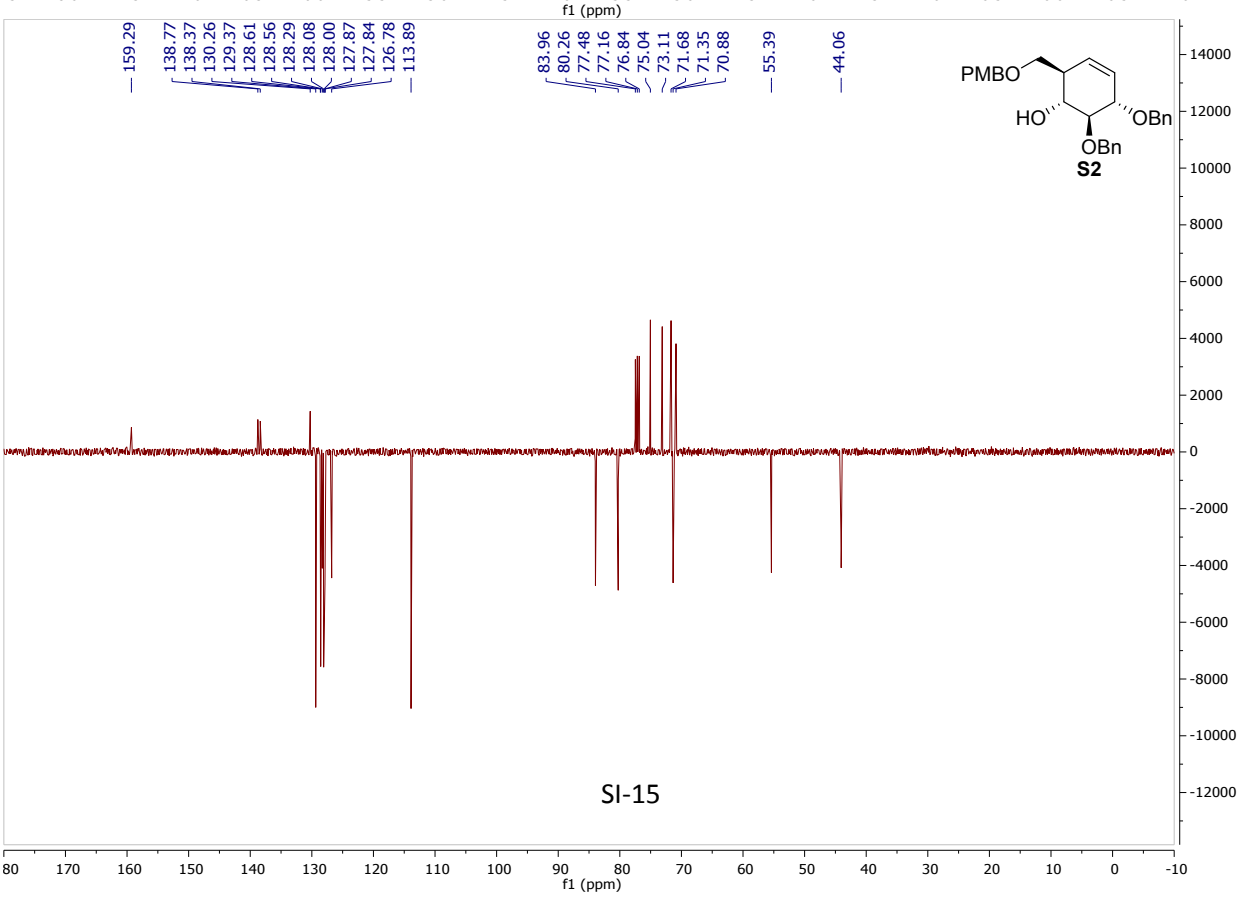
Following the general procedure starting from compound **21** (4.09 mg, 13.8 µmol), the product was obtained as a white powder (3.5 mg, 44%). ^{1}H NMR (600 MHz, CD_3OD): δ 7.84 (s, 1H), 4.49 (dd, *J* = 7.9, 4.9 Hz, 1H), 4.42 (s, 2H), 4.37 (t, *J* = 7.1 Hz, 2H), 4.29 (dd, *J* = 7.9, 4.4 Hz, 1H), 3.63 (dd, *J* = 10.2, 8.0 Hz, 1H), 3.50 (m, 2H), 3.19 (dt, *J* = 9.6, 5.2 Hz, 1H), 2.93 (dd, *J* = 12.7, 5.0 Hz, 1H), 2.71 (d, *J* = 12.7 Hz, 1H), 2.39 (dt, *J* = 11.7, 7.8 Hz, 1H), 2.23 (m, 3H), 2.04 (m, 2H), 1.89 (m, 3H), 1.76 – 1.63 (m, 5H), 1.63 – 1.51 (m, 4H), 1.49 (dd, *J* = 14.1, 10.3 Hz, 1H), 1.42 (m, 2H), 1.39 – 1.25 (m, 8H), 1.18 (t, *J* = 7.0 Hz, 1H), 0.97 (q, *J* = 11.8 Hz, 1H). ^{13}C NMR (151 MHz, CD_3OD): δ 176.0, 166.1, 146.2, 124.1, 68.3, 66.2, 63.3, 61.9, 61.6, 57.0, 51.4, 41.5, 41.1, 38.8, 38.7, 36.5, 35.6, 33.8, 32.4, 31.3, 30.4, 30.2, 30.0, 29.7, 29.5, 28.3, 27.4, 26.7. HRMS: [$M+H^+$] calc for $C_{28}H_{47}N_7O_4S$ 578.3483 found 578.3498.

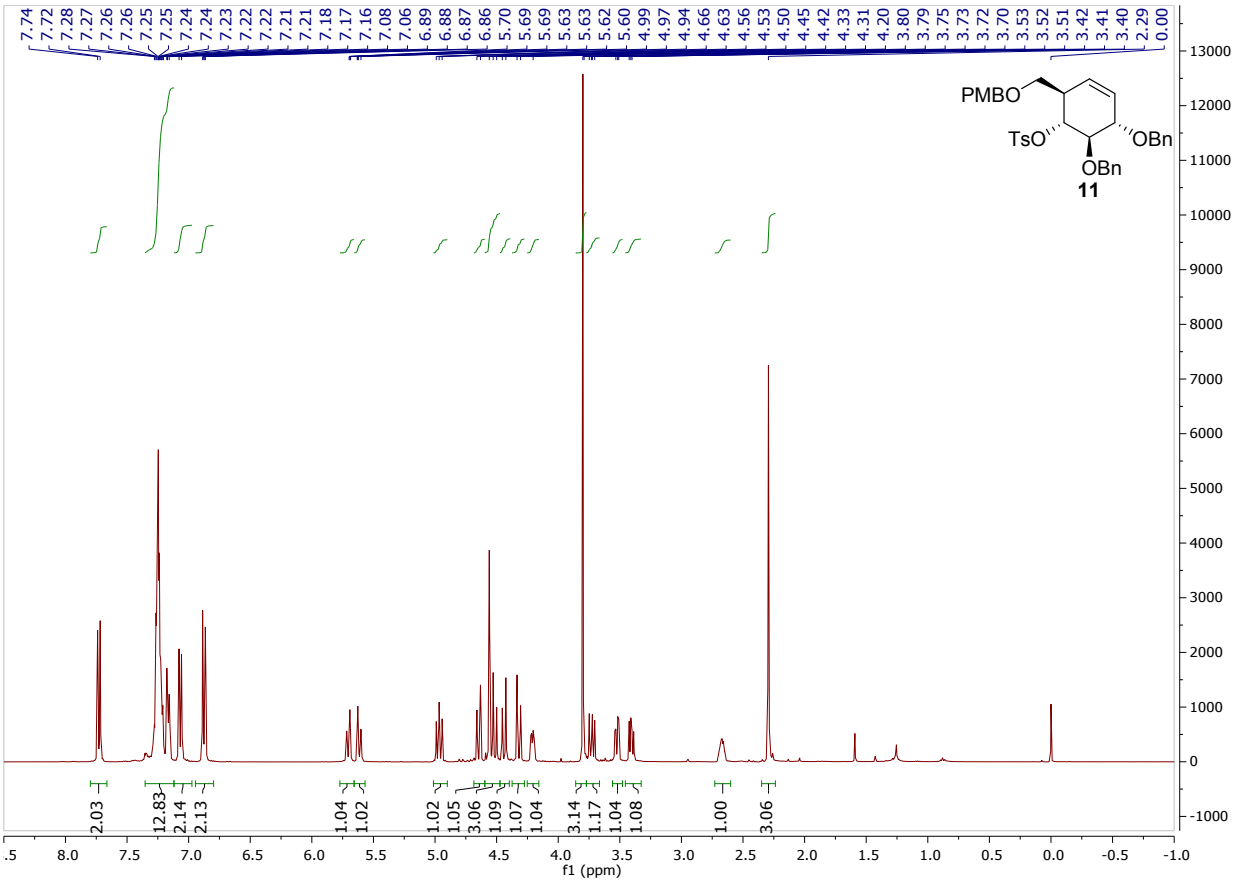
^{1}H and ^{13}C-NMR spectra of all compounds

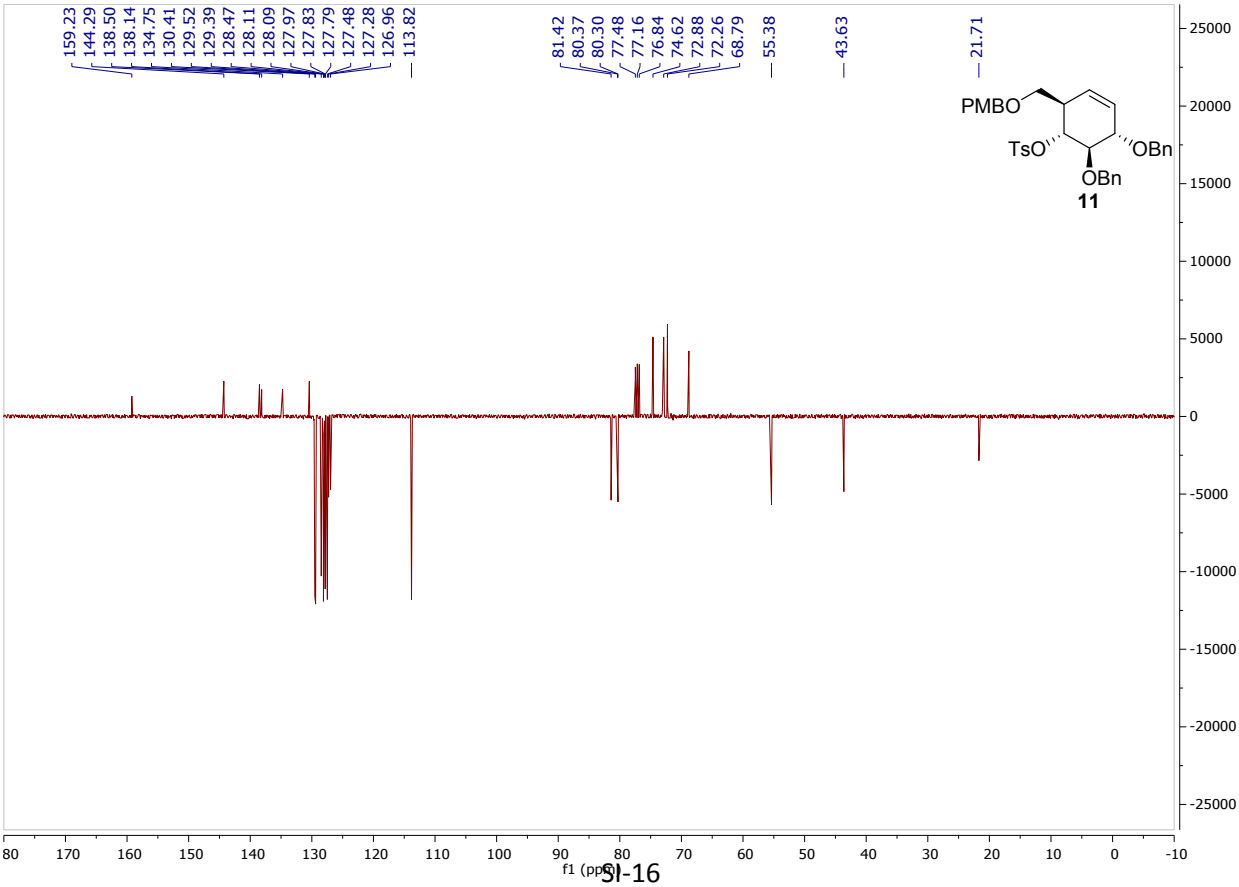


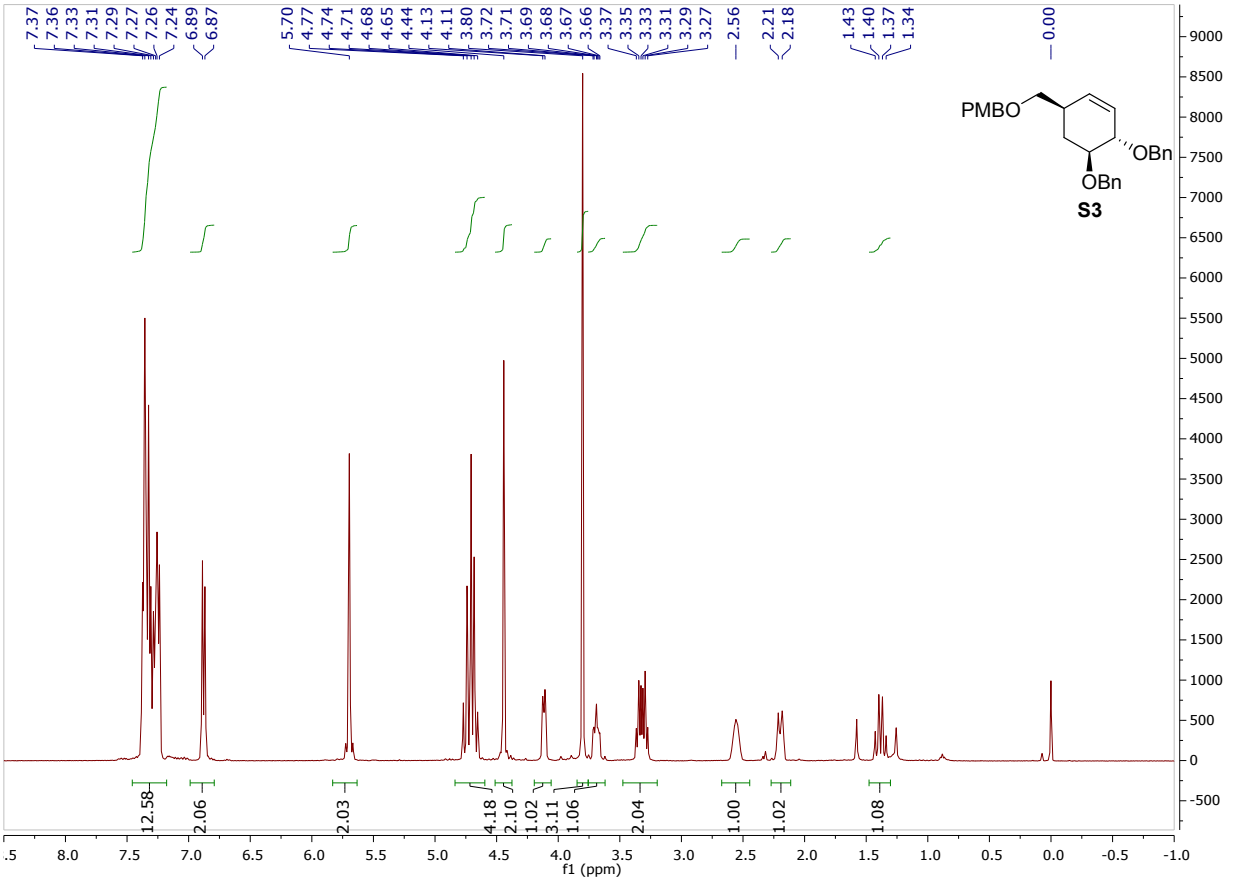


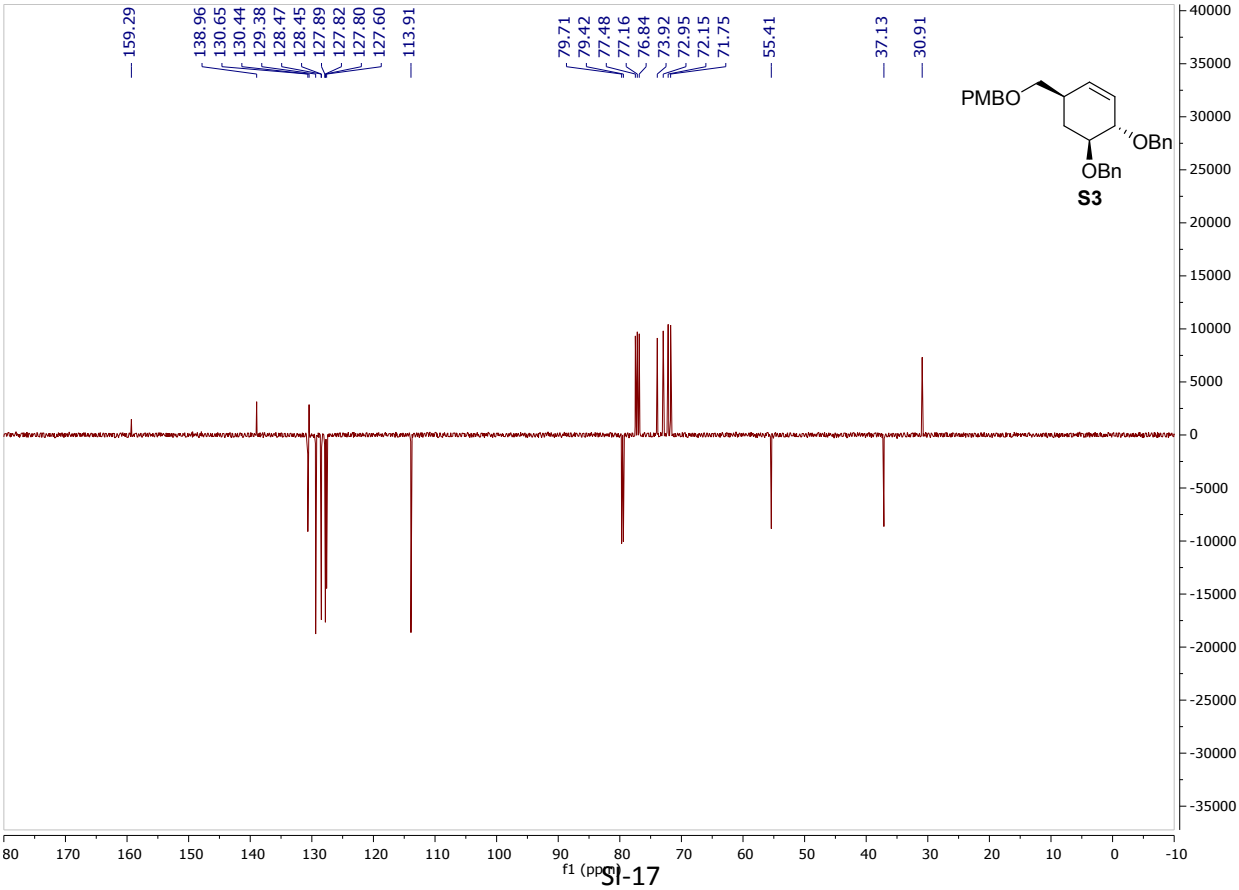

PMBO
HO
OBn
OBn
S2
f1 (ppm)

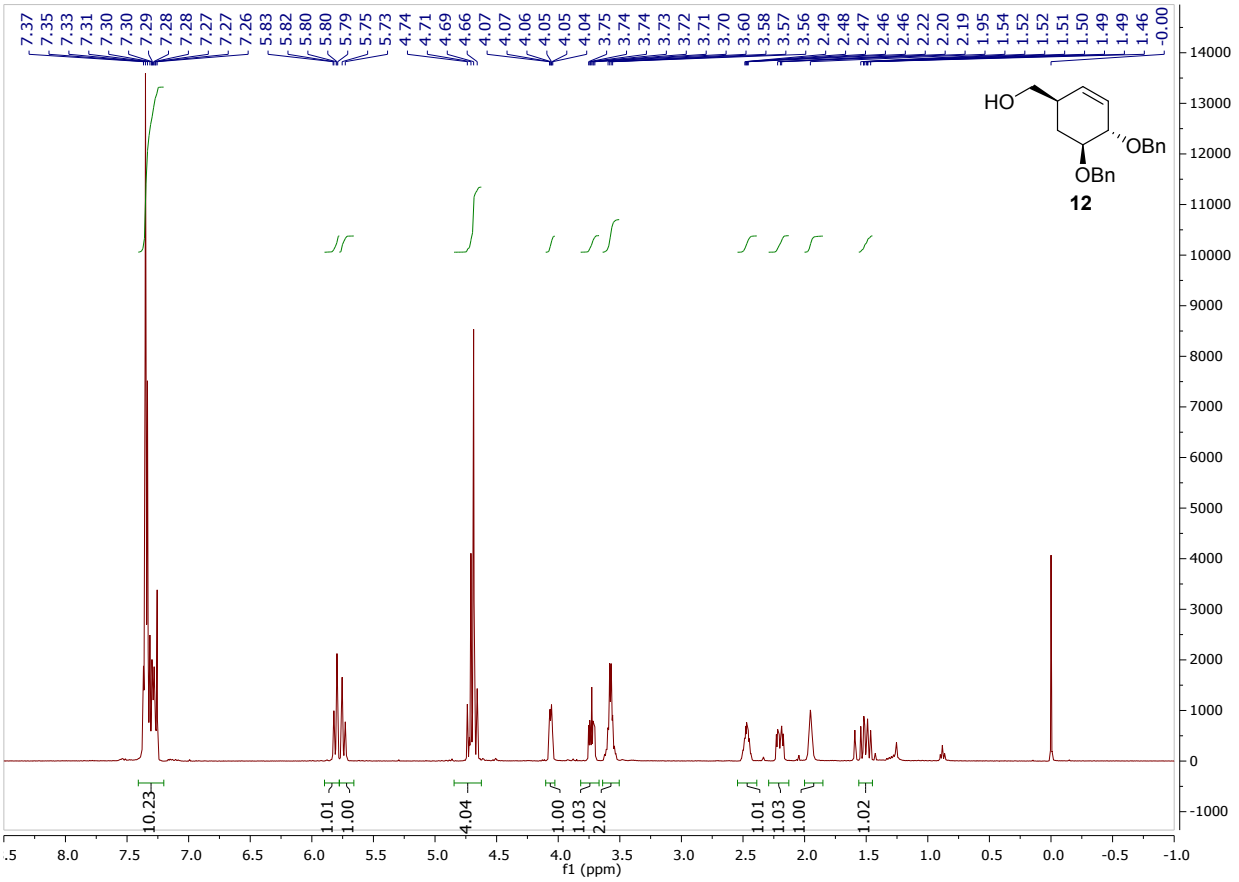

PMBO
HO
OBn
OBn
S2
f1 (ppm)

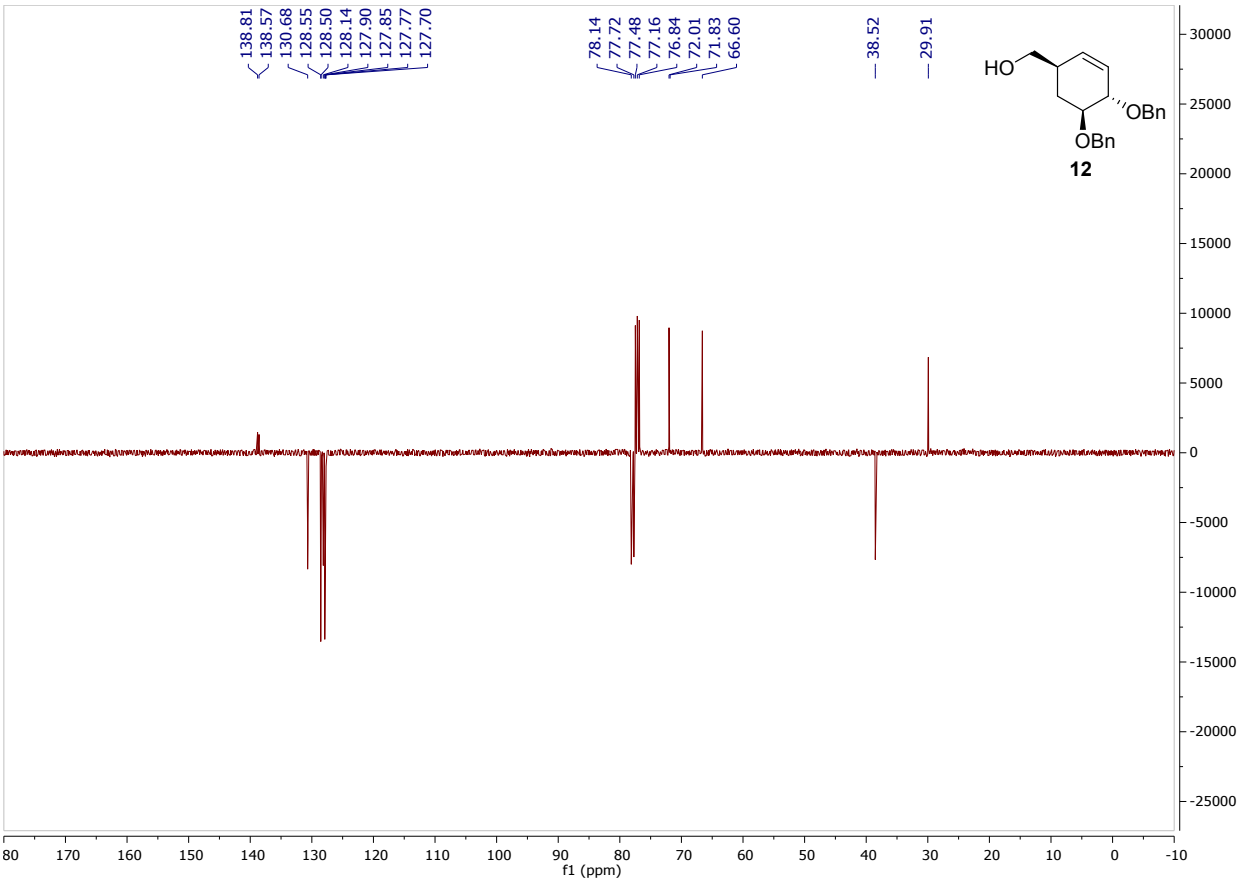

PMBO
TsO
OBn
OBn
11
f1 (ppm)

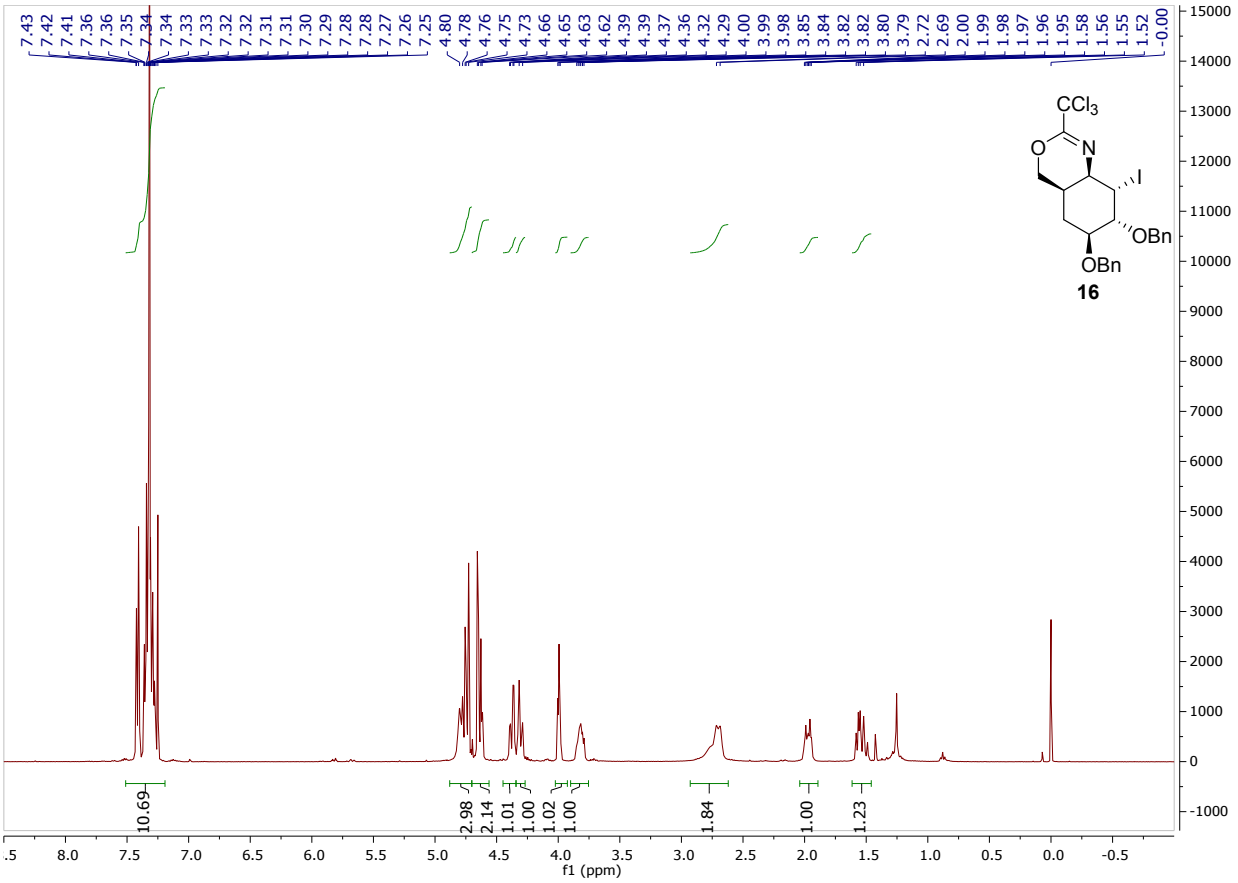

PMBO
TsO
OBn
OBn
11
f1 (ppm)

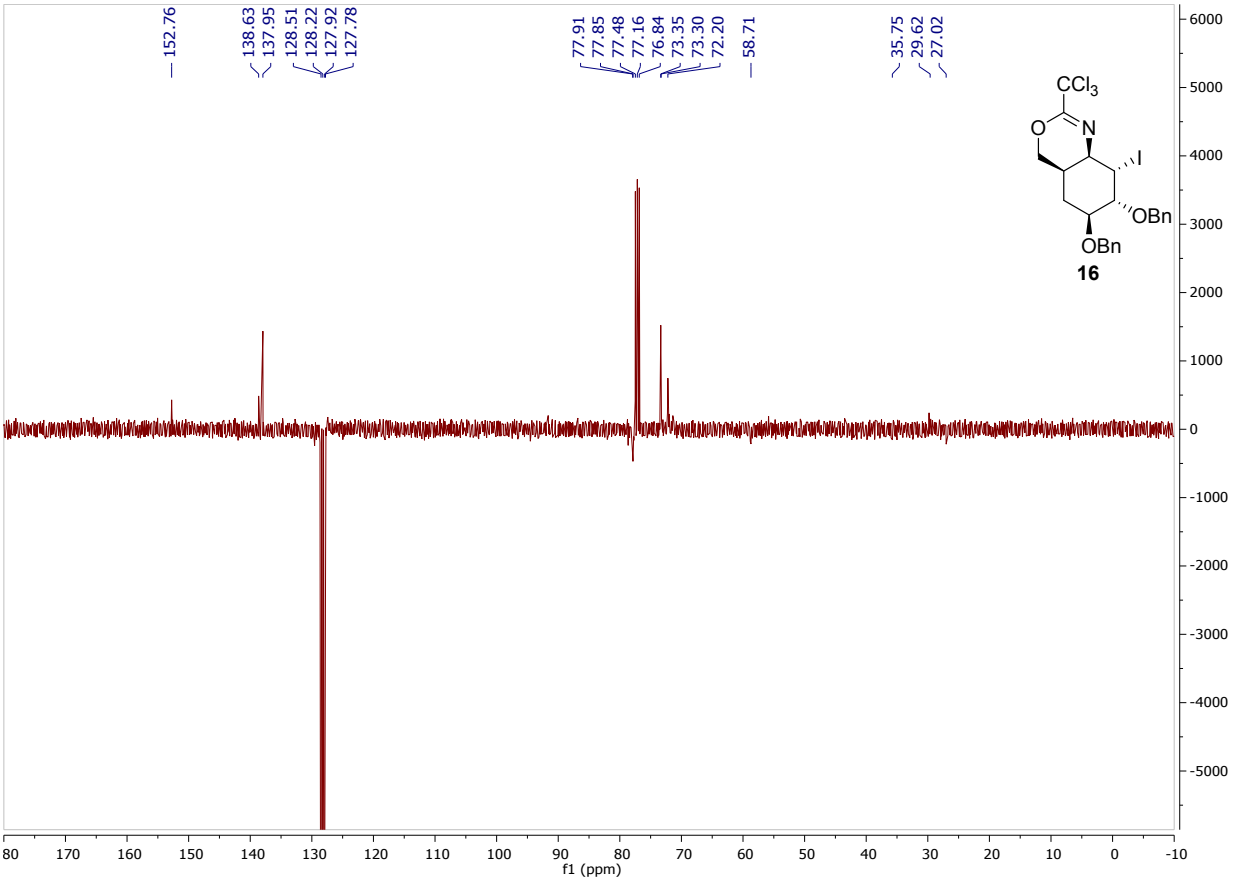

7.37
7.36
7.33
7.31
7.29
7.27
7.26
7.24
6.89
6.87
5.70
4.77
4.74
4.71
4.68
4.65
4.44
4.13
4.11
3.80
3.72
3.71
3.69
3.68
3.67
3.66
3.37
3.35
3.33
3.31
3.29
3.27
2.56
2.21
2.18
1.43
1.40
1.37
1.34
0.00
PMBO
OBn
OBn
S3
12.58
2.06
2.03
4.18
2.10
1.02
3.11
1.06
2.04
1.00
1.02
1.08
f1 (ppm)

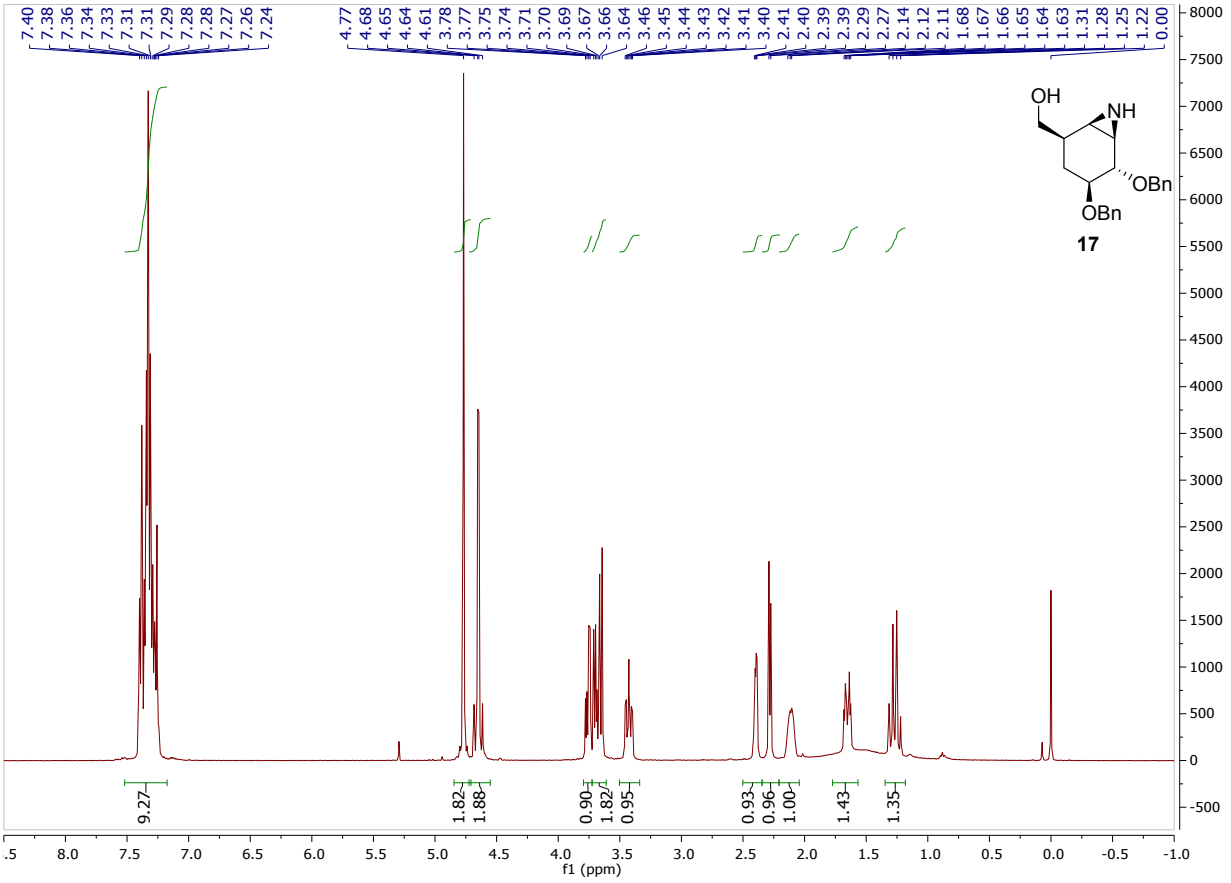

159.29
138.96
130.65
130.44
129.38
128.47
128.45
127.89
127.82
127.80
127.60
113.91
79.71
79.42
77.48
77.16
76.84
73.92
72.95
72.15
71.75
55.41
37.13
30.91
PMBO
OBn
OBn
S3
f1 (ppm)

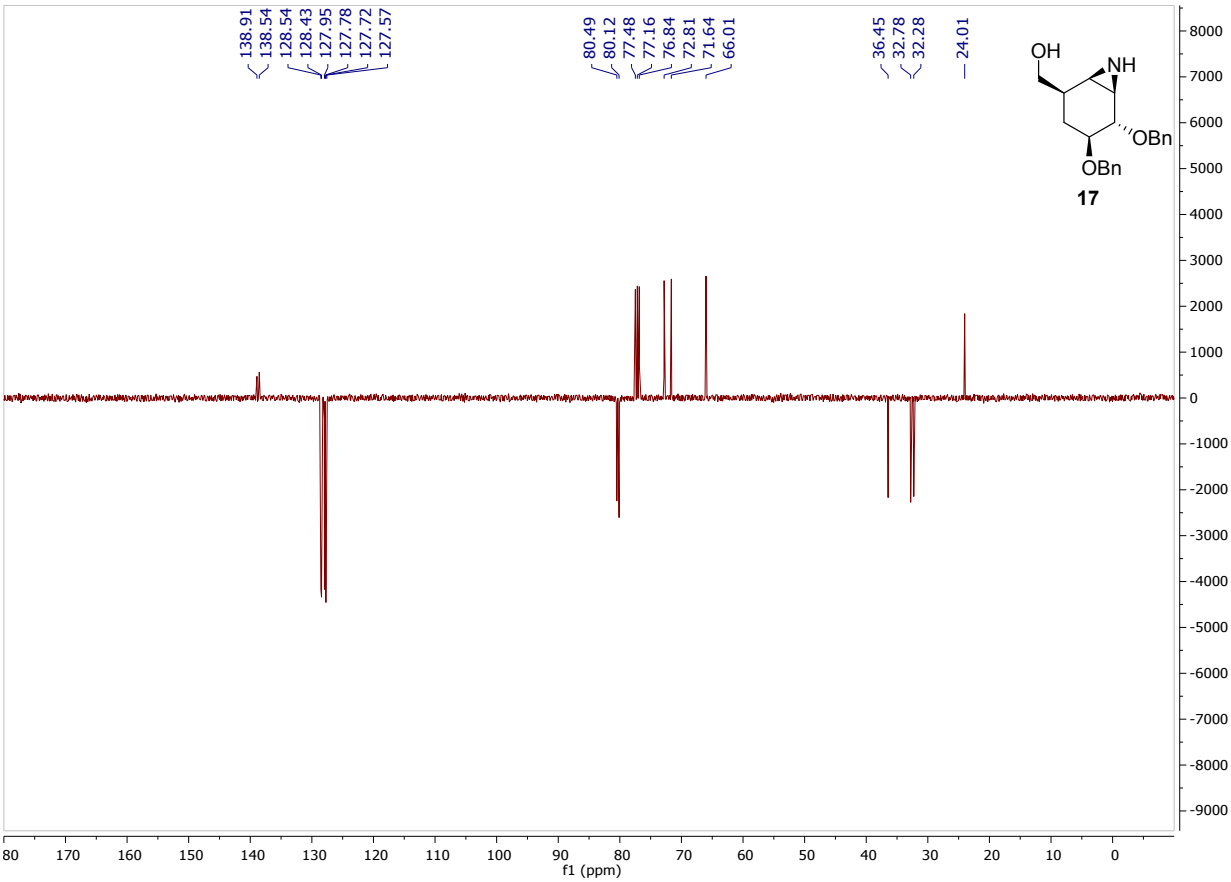

7.37 7.35 7.33 7.31 7.30 7.30 7.29 7.28 7.28 7.27 7.27 7.26 5.83 5.82 5.80 5.80 5.79 5.75 5.73 4.74 4.71 4.69 4.66 4.07 4.07 4.06 4.05 4.05 4.04 3.75 3.74 3.74 3.73 3.72 3.71 3.70 3.60 3.58 3.57 3.56 2.49 2.48 2.47 2.46 2.46 2.22 2.20 2.19 1.95 1.54 1.52 1.52 1.51 1.50 1.49 1.49 1.46 -0.00
HO
OBn
OBn
12
10.23 1.01 1.00 4.04 1.00 1.03 2.02 1.01 1.03 1.00 1.02
f1 (ppm)

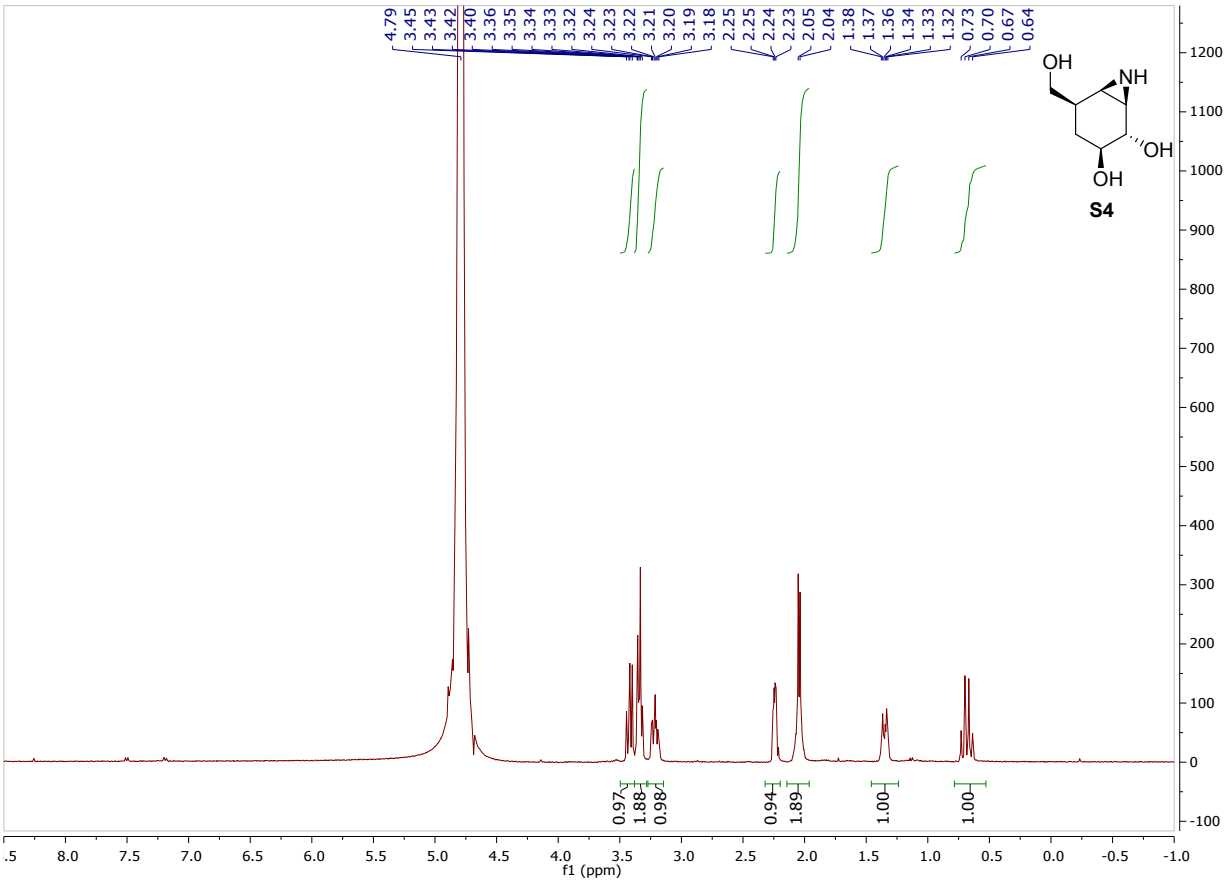

138.81 138.57 130.68 128.55 128.50 128.14 127.90 127.85 127.77 127.70
78.14 77.72 77.48 77.16 76.84 72.01 71.83 66.60
38.52
29.91
HO
OBn
OBn
12
f1 (ppm)

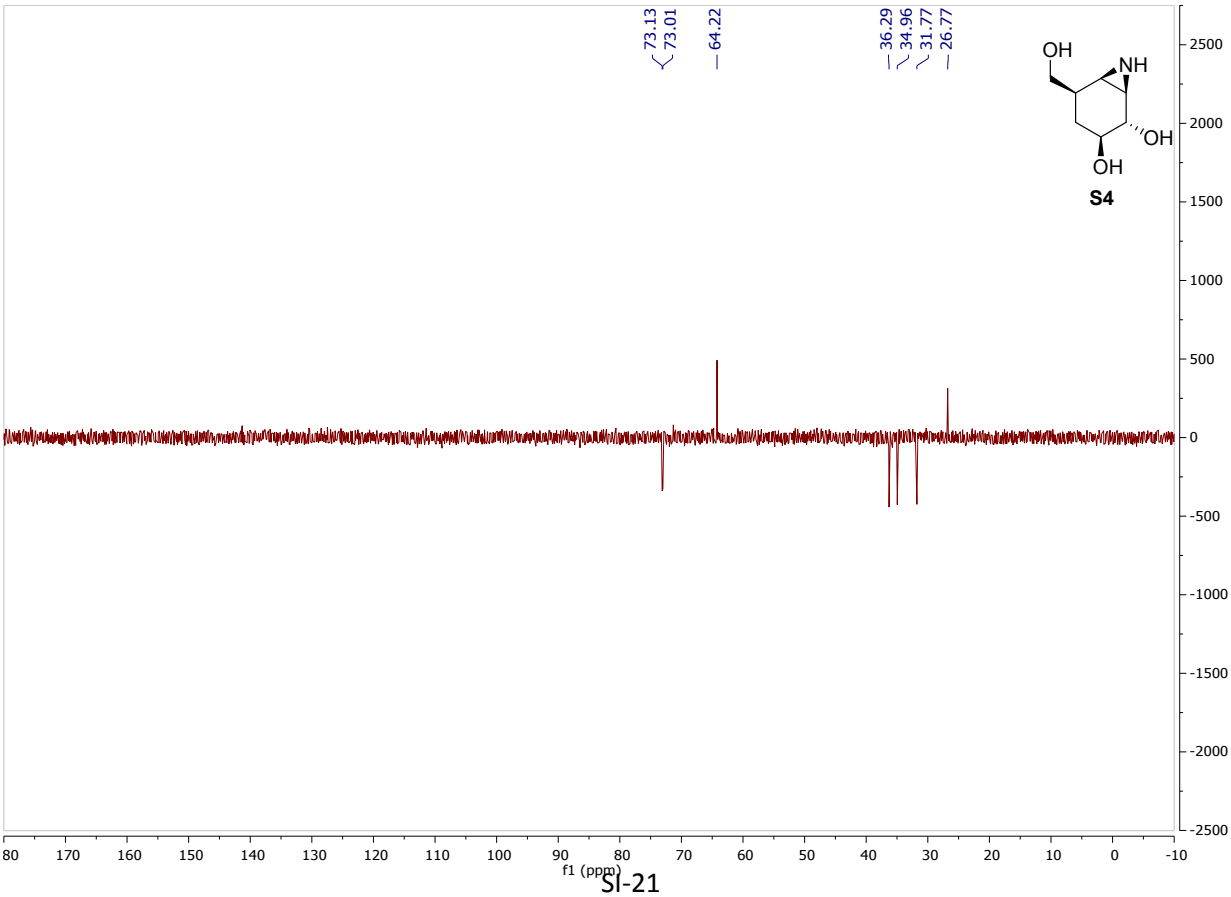

7.43 7.42 7.41 7.36 7.36 7.35 7.34 7.34 7.33 7.33 7.32 7.32 7.31 7.31 7.30 7.29 7.28 7.28 7.27 7.26 7.25
4.80 4.78 4.76 4.75 4.73 4.66 4.65 4.63 4.62 4.39 4.39 4.37 4.36 4.32 4.29 4.00 3.99 3.98 3.85 3.84 3.82 3.82 3.80 3.79 2.72 2.69 2.00 1.99 1.98 1.97 1.96 1.95 1.58 1.56 1.55 1.52 -0.00
10.69
2.98 2.14 1.01 1.00 1.02 1.00
1.84
1.00
1.23
f1 (ppm)
CCl3
O
N
I
OBn
OBn
16

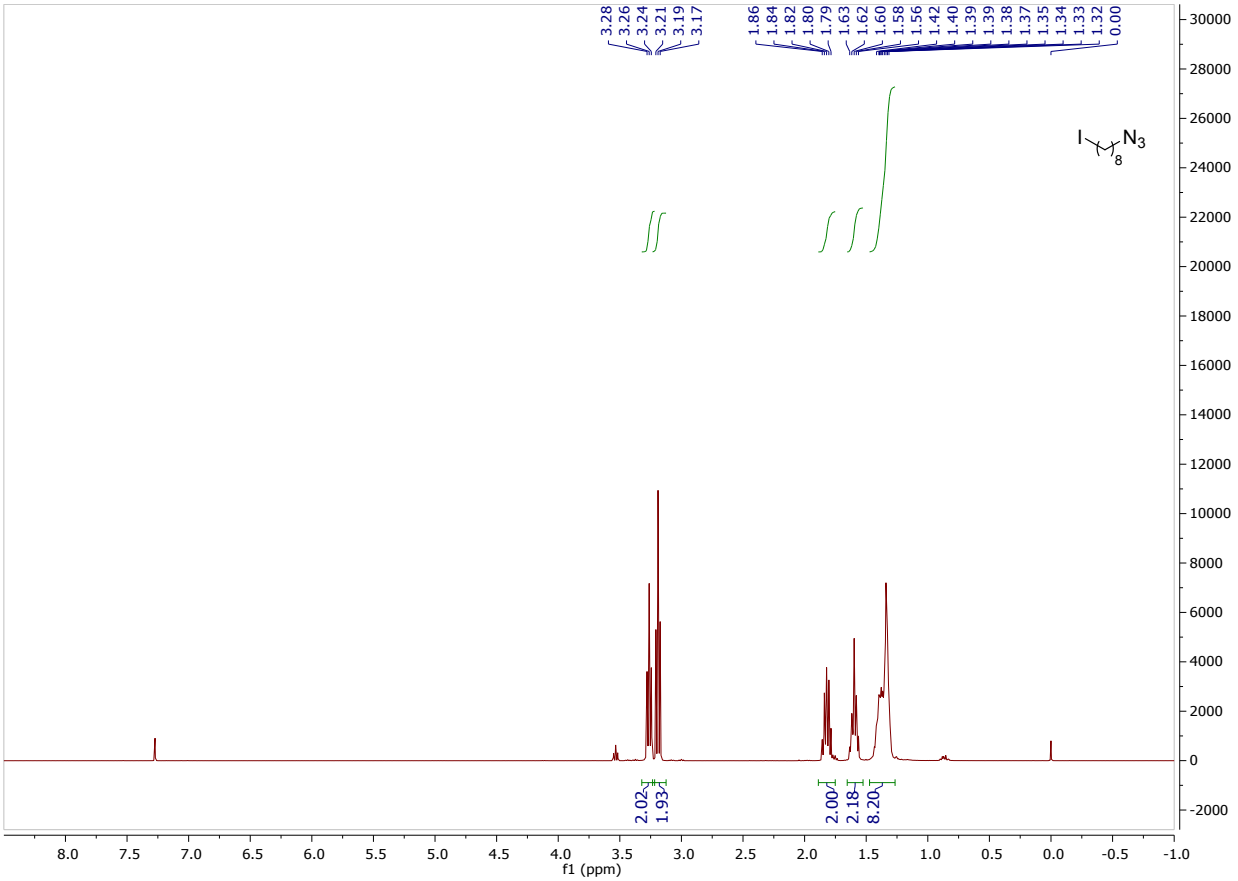

152.76
138.63 137.95 128.51 128.22 127.92 127.78
77.91 77.85 77.48 77.16 76.84 73.35 73.30 72.20
58.71
35.75 29.62 27.02
f1 (ppm)
CCl3
O
N
I
OBn
OBn
16

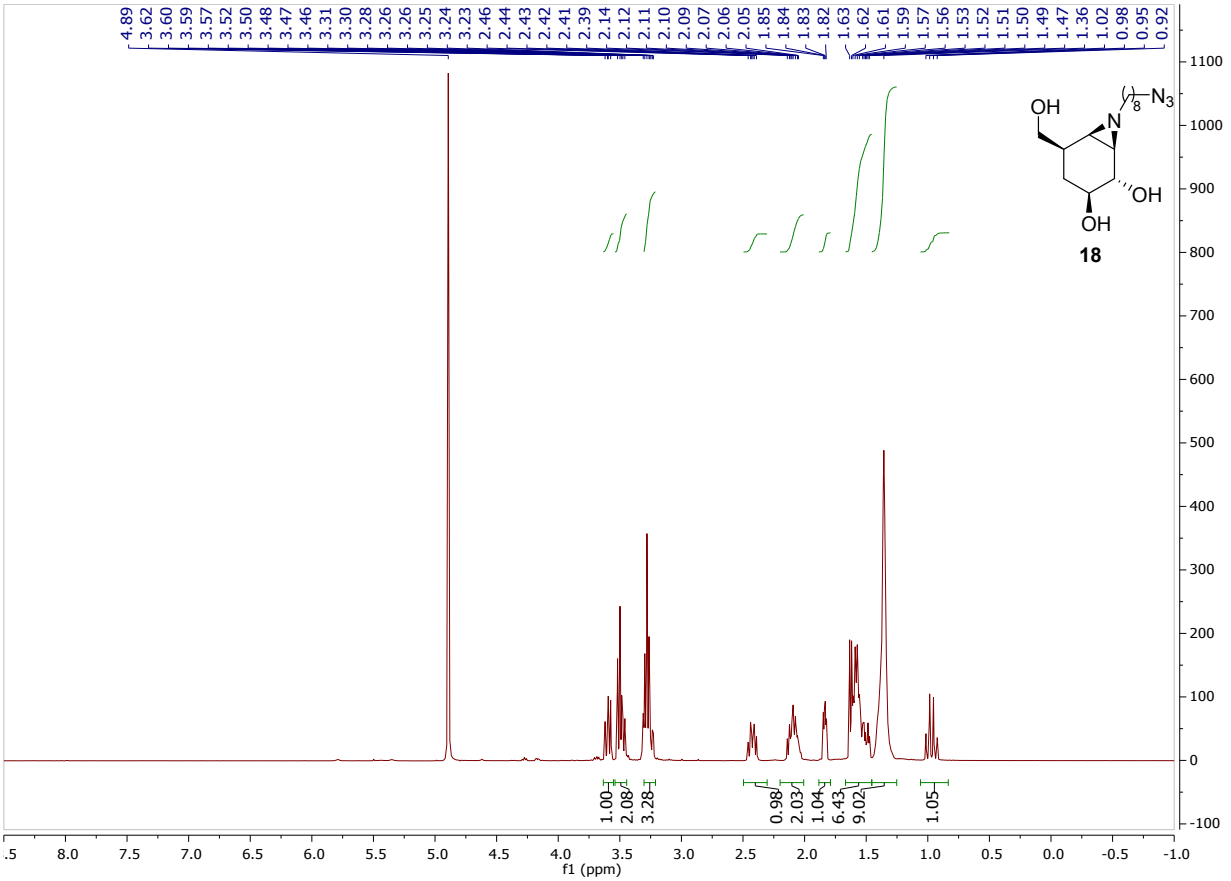

7.40 7.38 7.36 7.34 7.33 7.31 7.31 7.29 7.28 7.28 7.27 7.26 7.24
4.77 4.68 4.65 4.64 4.61 3.78 3.77 3.75 3.74 3.71 3.70 3.69 3.67 3.66 3.64 3.46 3.45 3.44 3.43 3.42 3.41 3.40 2.41 2.40 2.39 2.29 2.27 2.14 2.12 2.11 1.68 1.67 1.66 1.65 1.64 1.63 1.31 1.28 1.25 1.22 0.00
OH
NH
OBn
OBn
17
9.27
1.82
1.88
0.90
1.82
0.95
0.93
0.96
1.00
1.43
1.35
f1 (ppm)

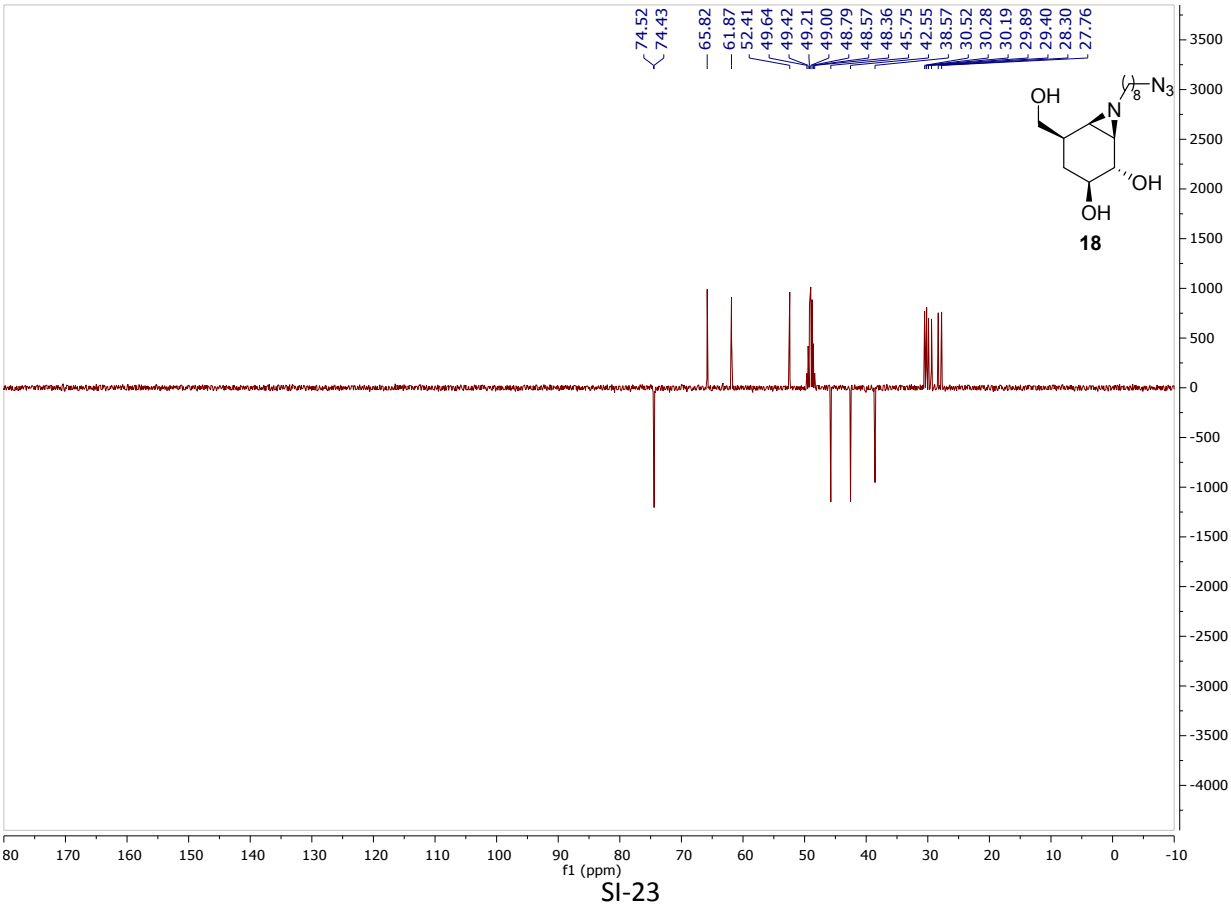

138.91 138.54 128.54 128.43 127.95 127.78 127.72 127.57
80.49 80.12 77.48 77.16 76.84 72.81 71.64 66.01
36.45 32.78 32.28
24.01
OH
NH
OBn
OBn
17
f1 (ppm)

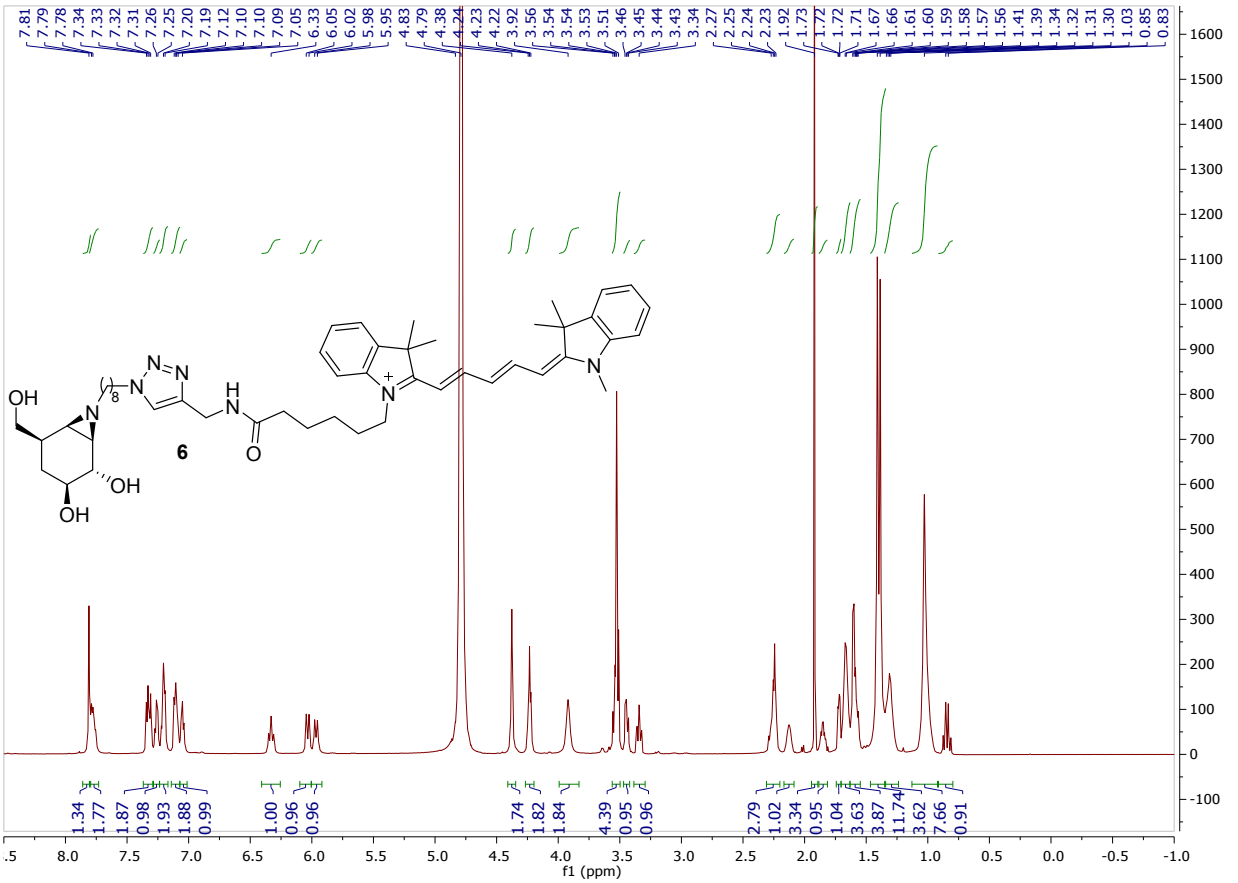

4.79
3.45
3.43
3.42
3.40
3.36
3.35
3.34
3.33
3.32
3.24
3.23
3.22
3.21
3.20
3.19
3.18
2.25
2.25
2.24
2.23
2.05
2.04
1.38
1.37
1.36
1.34
1.33
1.32
0.73
0.70
0.67
0.64
OH
NH
OH
OH
S4
0.97
1.88
0.98
0.94
1.89
1.00
1.00
f1 (ppm)

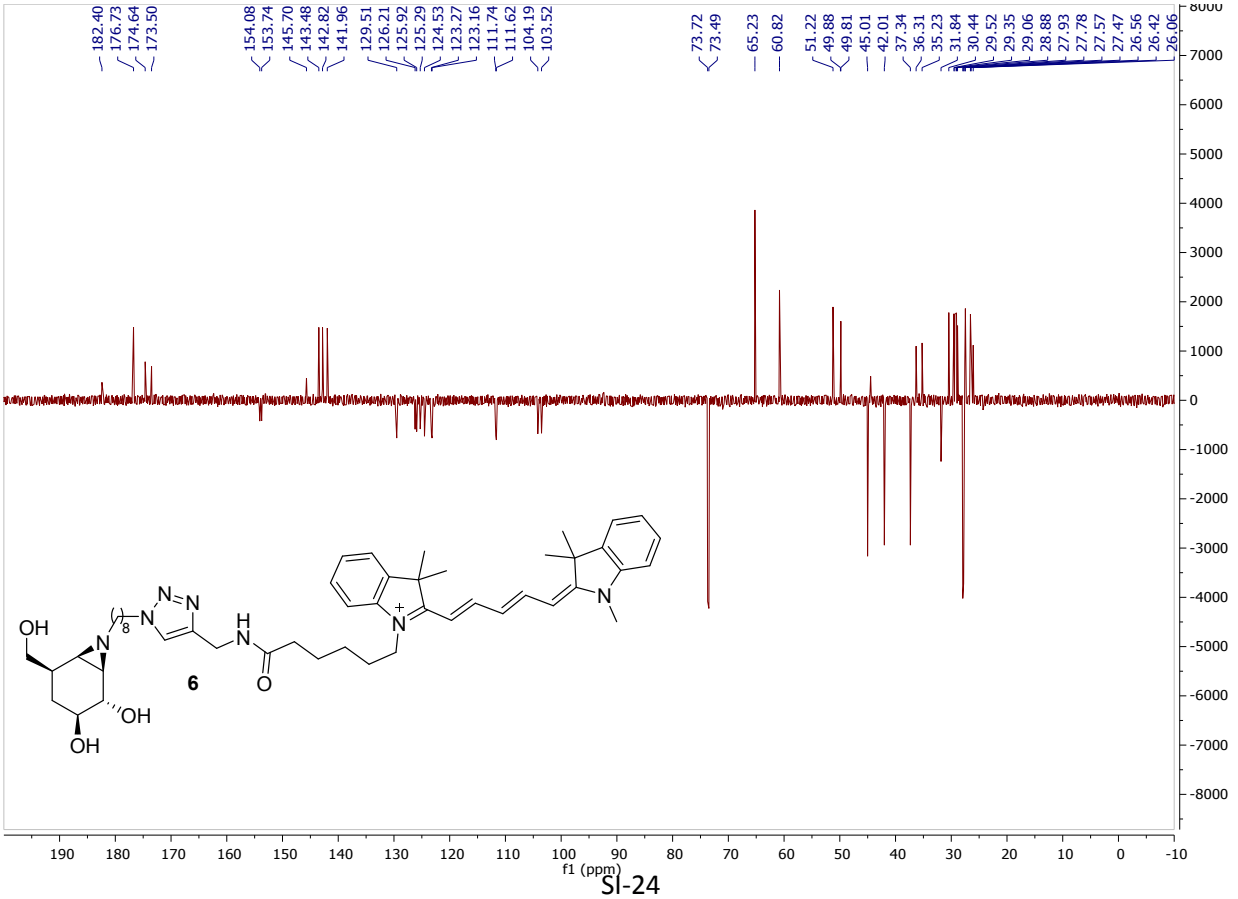

73.13
73.01
64.22
36.29
34.96
31.77
26.77
OH
NH
OH
OH
S4
f1 (ppm)

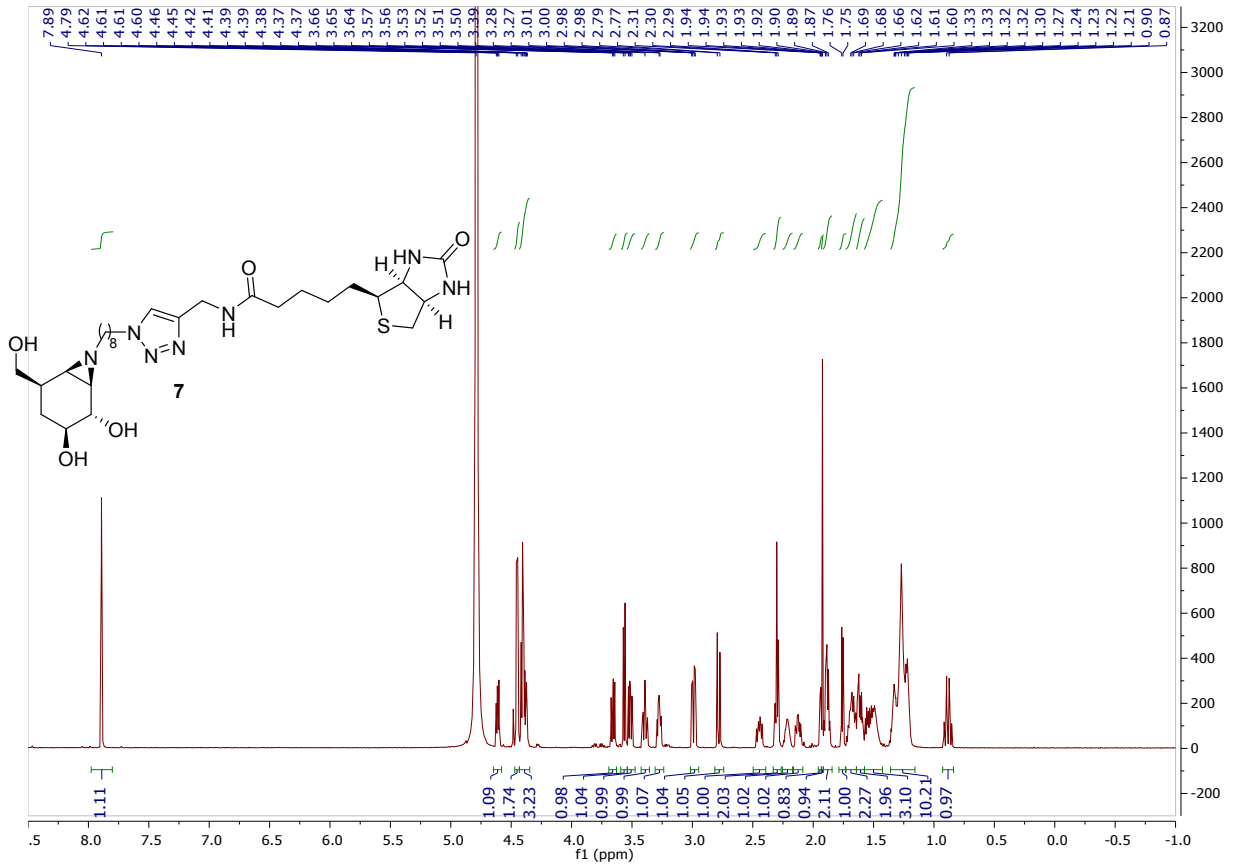

3.28
3.26
3.24
3.21
3.19
3.17
1.86
1.84
1.82
1.80
1.79
1.63
1.62
1.60
1.58
1.56
1.42
1.40
1.39
1.39
1.38
1.37
1.35
1.34
1.33
1.32
0.00
I
N3
8
2.02
1.93
2.00
2.18
8.20
f1 (ppm)

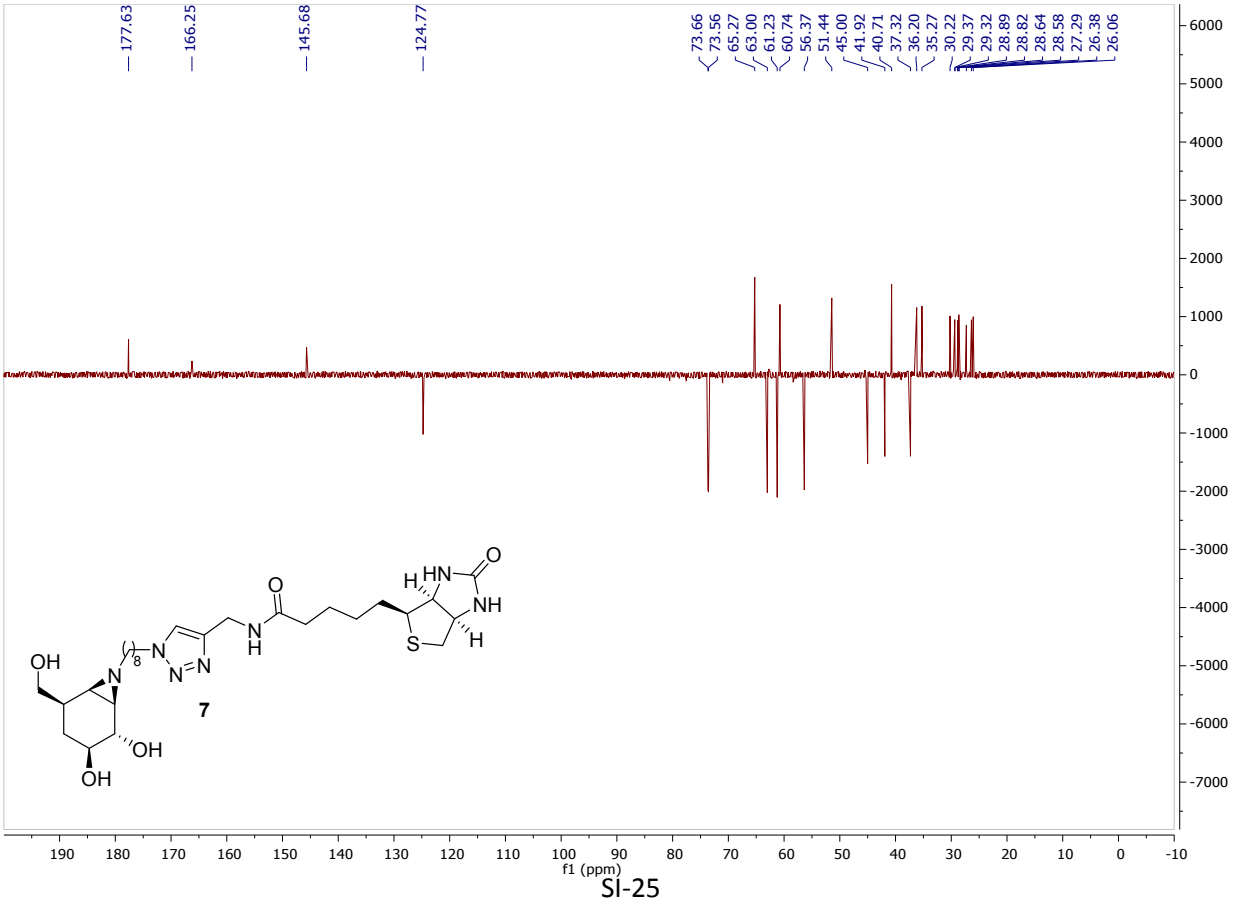

4.89 3.62 3.60 3.59 3.57 3.52 3.50 3.48 3.47 3.46 3.31 3.30 3.28 3.26 3.26 3.25 3.24 3.23 2.46 2.44 2.43 2.42 2.41 2.39 2.14 2.12 2.11 2.10 2.09 2.07 2.06 2.05 1.85 1.84 1.83 1.82 1.63 1.62 1.61 1.59 1.57 1.56 1.53 1.52 1.51 1.50 1.49 1.47 1.36 1.02 0.98 0.95 0.92
OH
N
N3
8
OH
OH
18
1.00 2.08 3.28 0.98 2.03 1.04 6.43 9.02 1.05
f1 (ppm)

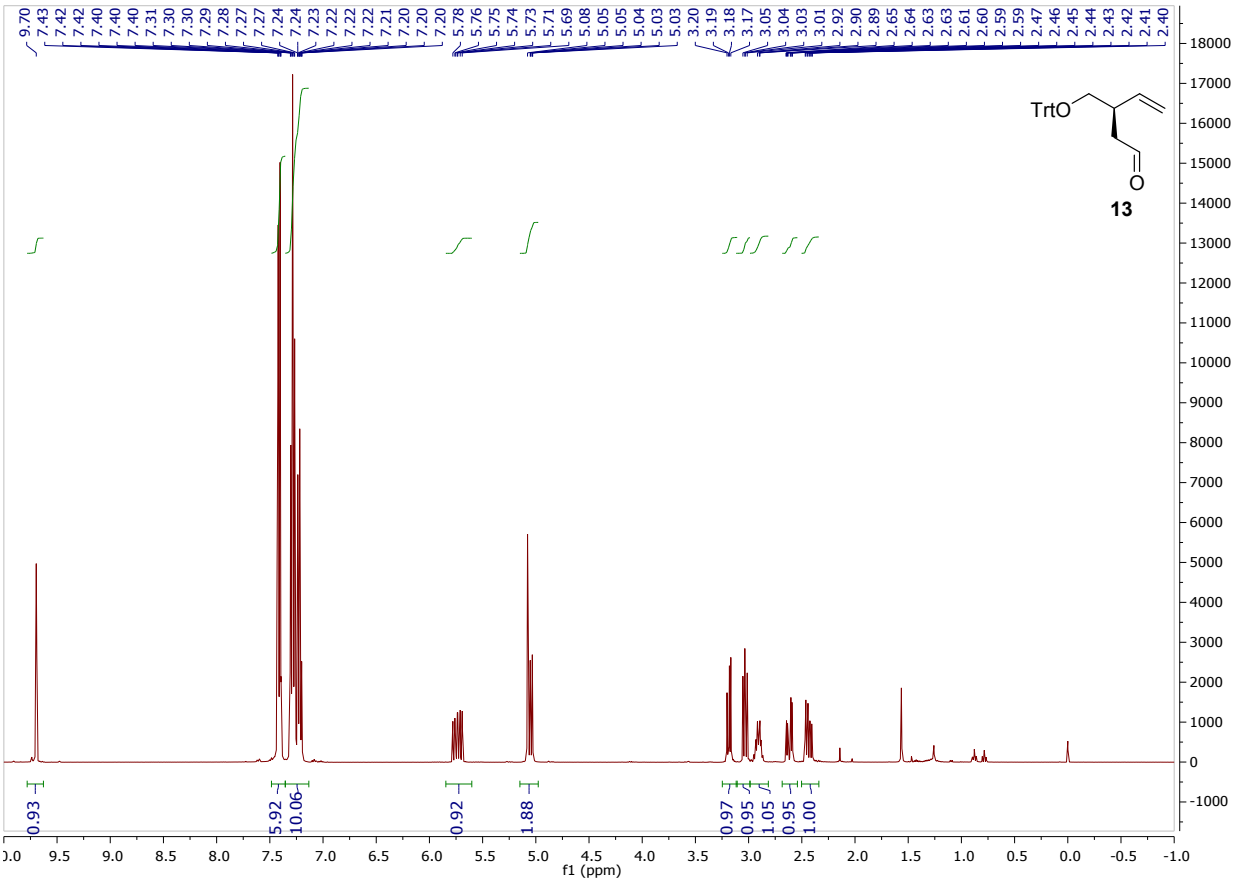

74.52 74.43 65.82 61.87 52.41 49.64 49.42 49.21 49.00 48.79 48.57 48.36 45.75 42.55 38.57 30.52 30.28 30.19 29.89 29.40 28.30 27.76
OH
N
N3
8
OH
OH
18
f1 (ppm)

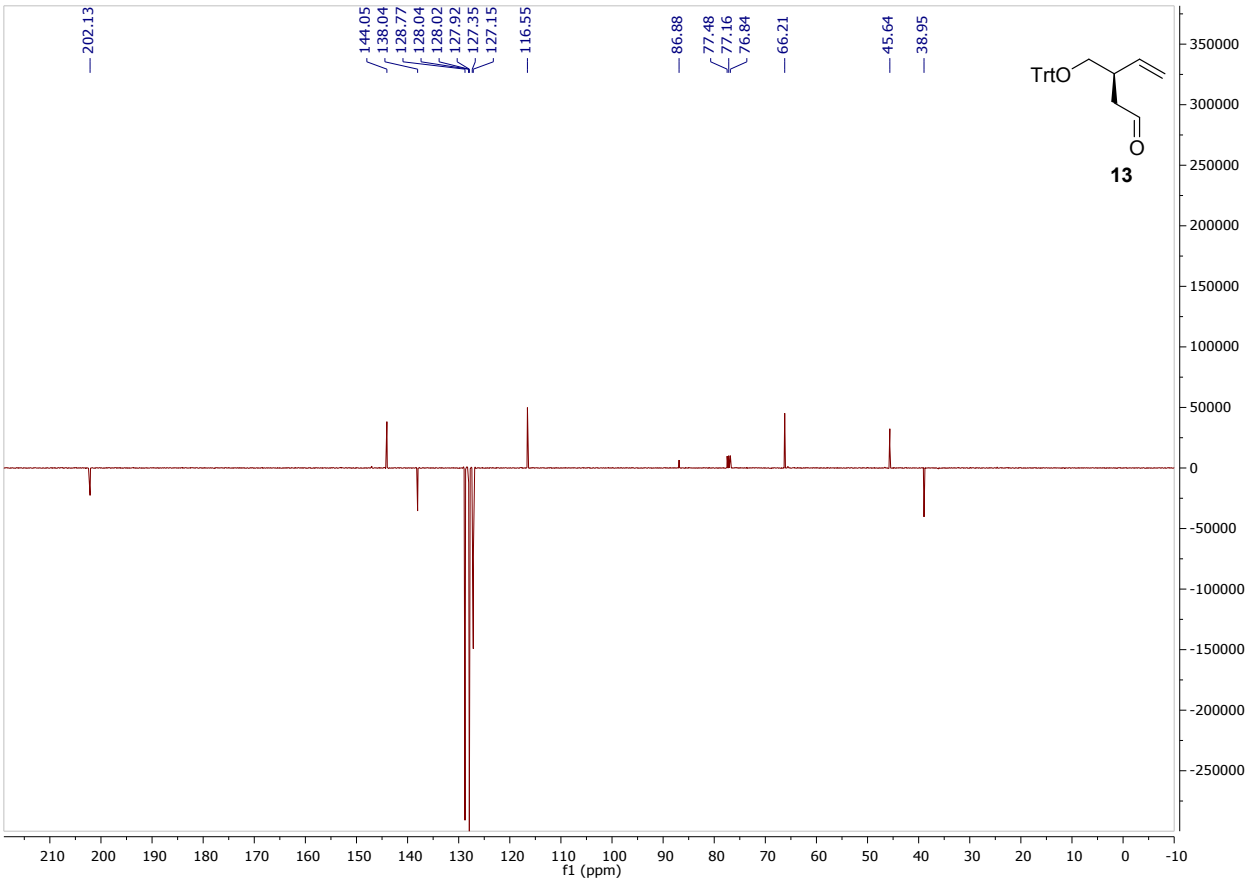

7.81 7.79 7.78 7.34 7.33 7.32 7.31 7.26 7.25 7.20 7.19 7.12 7.10 7.10 7.09 7.05 6.33 6.05 6.02 5.98 5.95 4.83 4.79 4.38 4.24 4.23 4.22 3.92 3.56 3.54 3.54 3.53 3.51 3.46 3.45 3.44 3.43 3.34 2.27 2.25 2.24 2.23 1.92 1.73 1.72 1.72 1.71 1.67 1.66 1.61 1.60 1.59 1.58 1.57 1.56 1.41 1.39 1.34 1.32 1.31 1.30 1.03 0.85 0.83
1.34 1.77 1.87 0.98 1.93 1.88 0.99 1.00 0.96 0.96 1.74 1.82 1.84 4.39 0.95 0.96 2.79 1.02 3.34 0.95 1.04 3.63 3.87 11.74 3.62 7.66 0.91
f1 (ppm)
OH
OH
OH
6

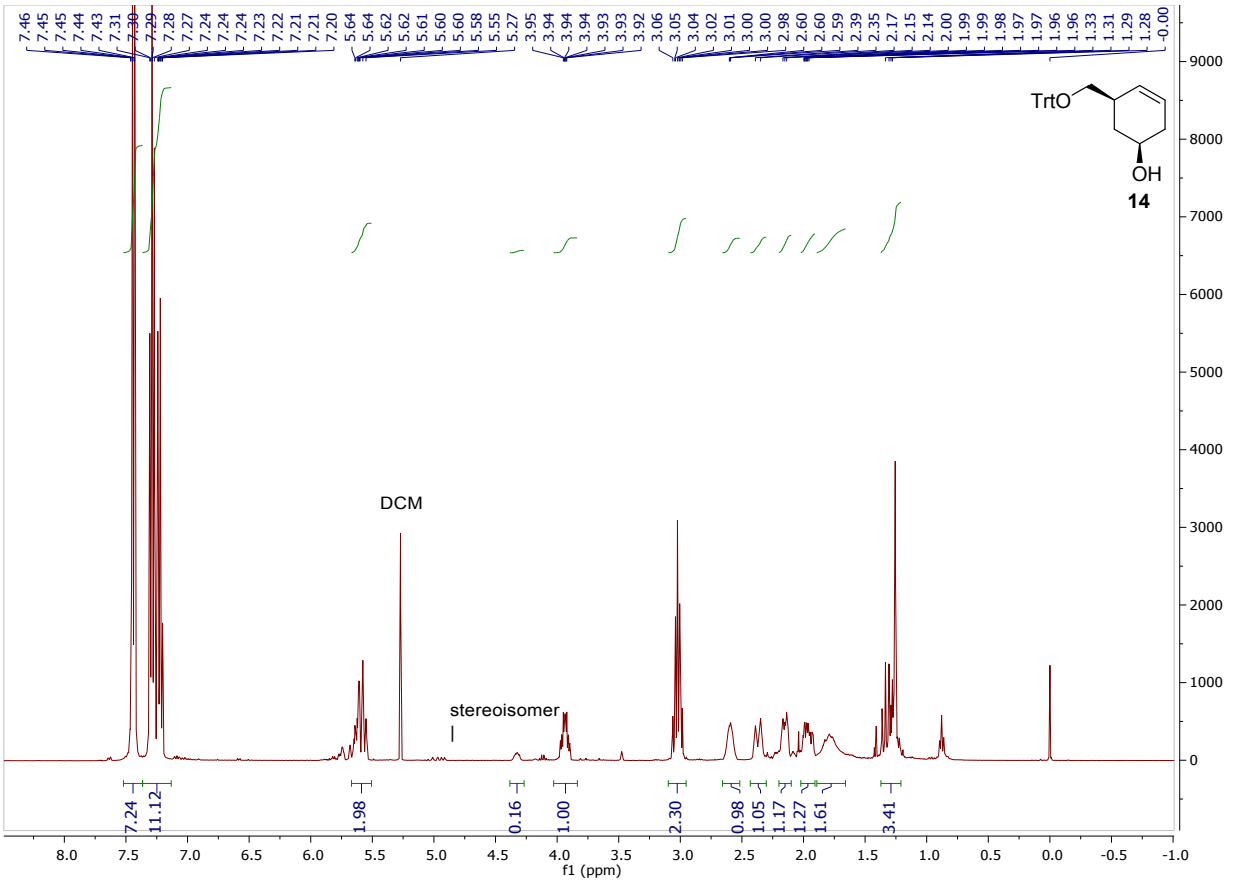

182.40 176.73 174.64 173.50 154.08 153.74 145.70 143.48 142.82 141.96 129.51 126.21 125.92 125.29 124.53 123.27 123.16 111.74 111.62 104.19 103.52 73.72 73.49 65.23 60.82 51.22 49.88 49.81 45.01 42.01 37.34 36.31 35.23 31.84 30.44 29.52 29.35 29.06 28.88 27.93 27.78 27.57 27.47 26.56 26.42 26.06
f1 (ppm)
OH
OH
OH
6

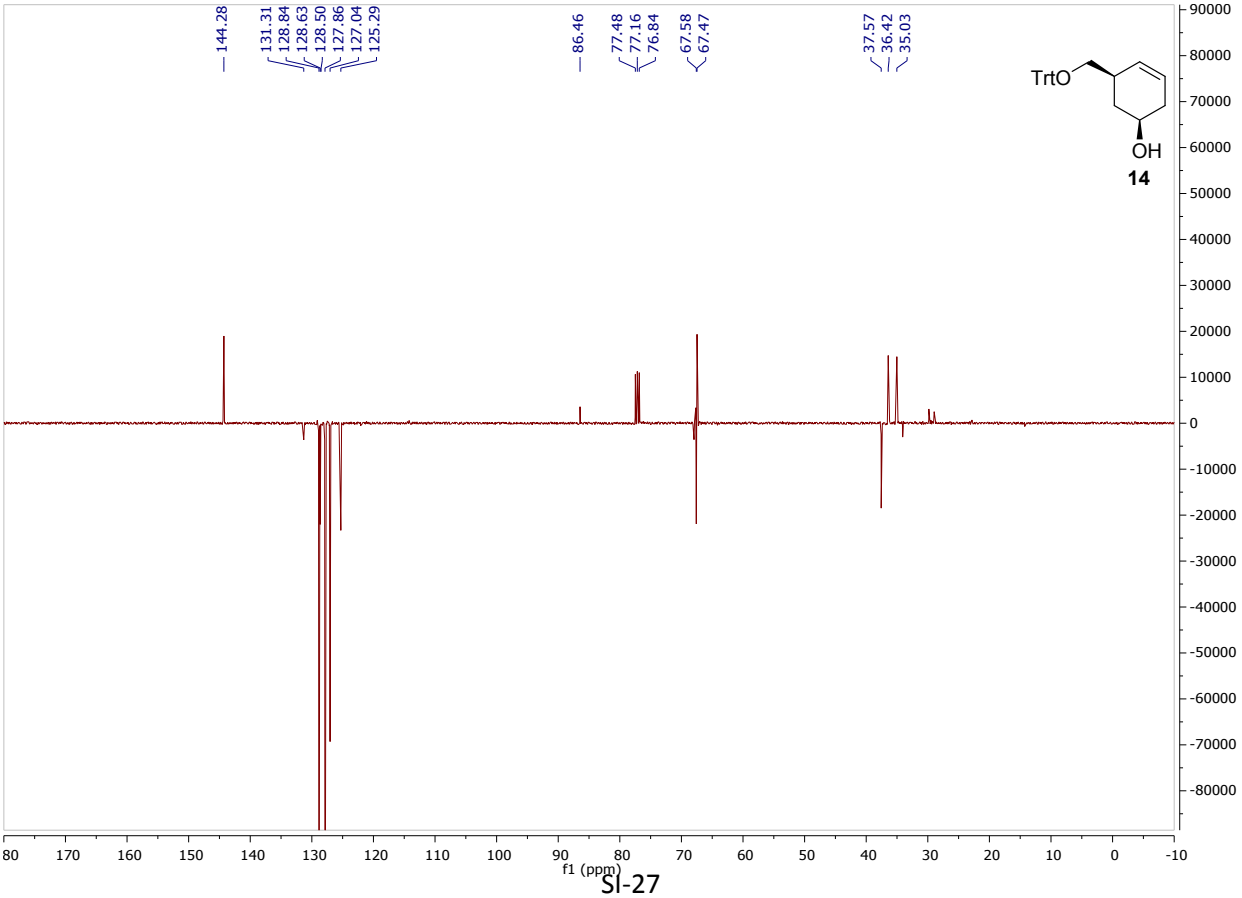

7.89 4.79 4.62 4.61 4.61 4.60 4.46 4.45 4.42 4.41 4.39 4.39 4.38 4.37 4.37 3.66 3.65 3.64 3.57 3.56 3.53 3.52 3.51 3.50 3.39 3.28 3.27 3.01 3.00 2.98 2.98 2.79 2.77 2.31 2.30 2.29 1.94 1.94 1.93 1.93 1.92 1.90 1.89 1.87 1.76 1.75 1.69 1.68 1.66 1.62 1.61 1.60 1.33 1.33 1.32 1.30 1.27 1.24 1.23 1.22 1.21 0.90 0.87
1.11 1.09 1.74 3.23 0.98 1.04 0.99 0.99 1.07 1.04 1.05 1.00 2.03 1.02 1.02 0.83 0.94 2.11 1.00 2.27 1.96 3.10 10.21 0.97
f1 (ppm)
7

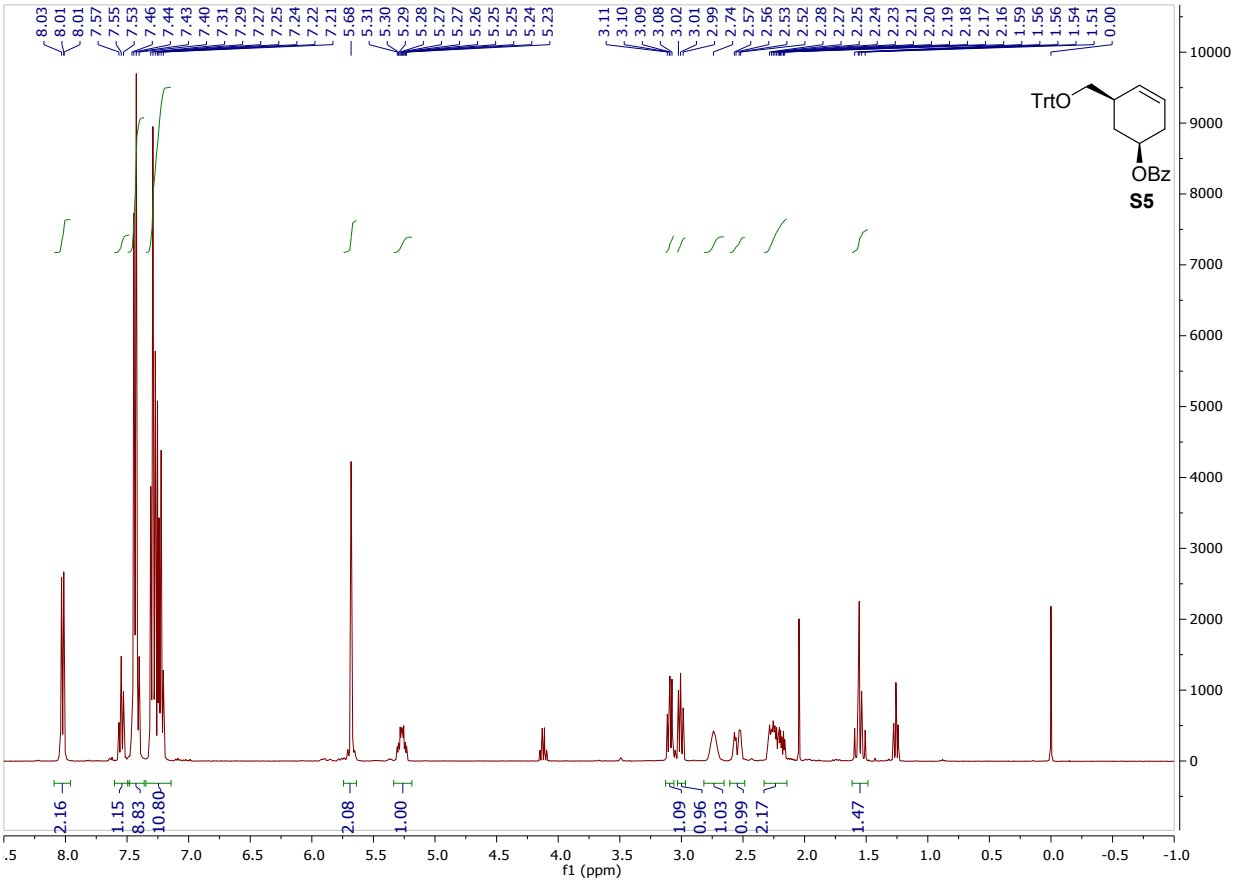

177.63 166.25 145.68 124.77
73.66 73.56 65.27 63.00 61.23 60.74 56.37 51.44 45.00 41.92 40.71 37.32 36.20 35.27 30.22 29.37 29.32 28.89 28.82 28.64 28.58 27.29 26.38 26.06
f1 (ppm)
7

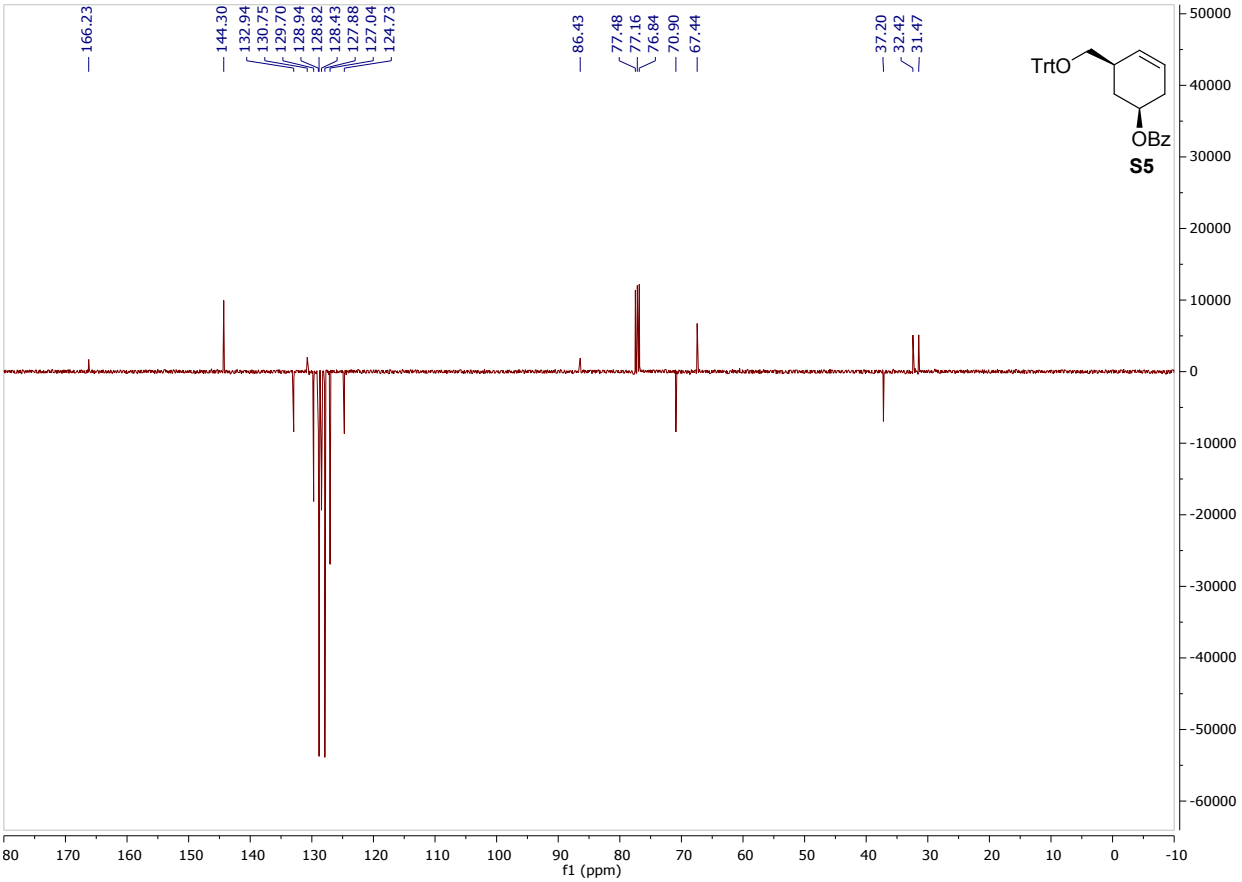

TrtO
O
13
f1 (ppm)

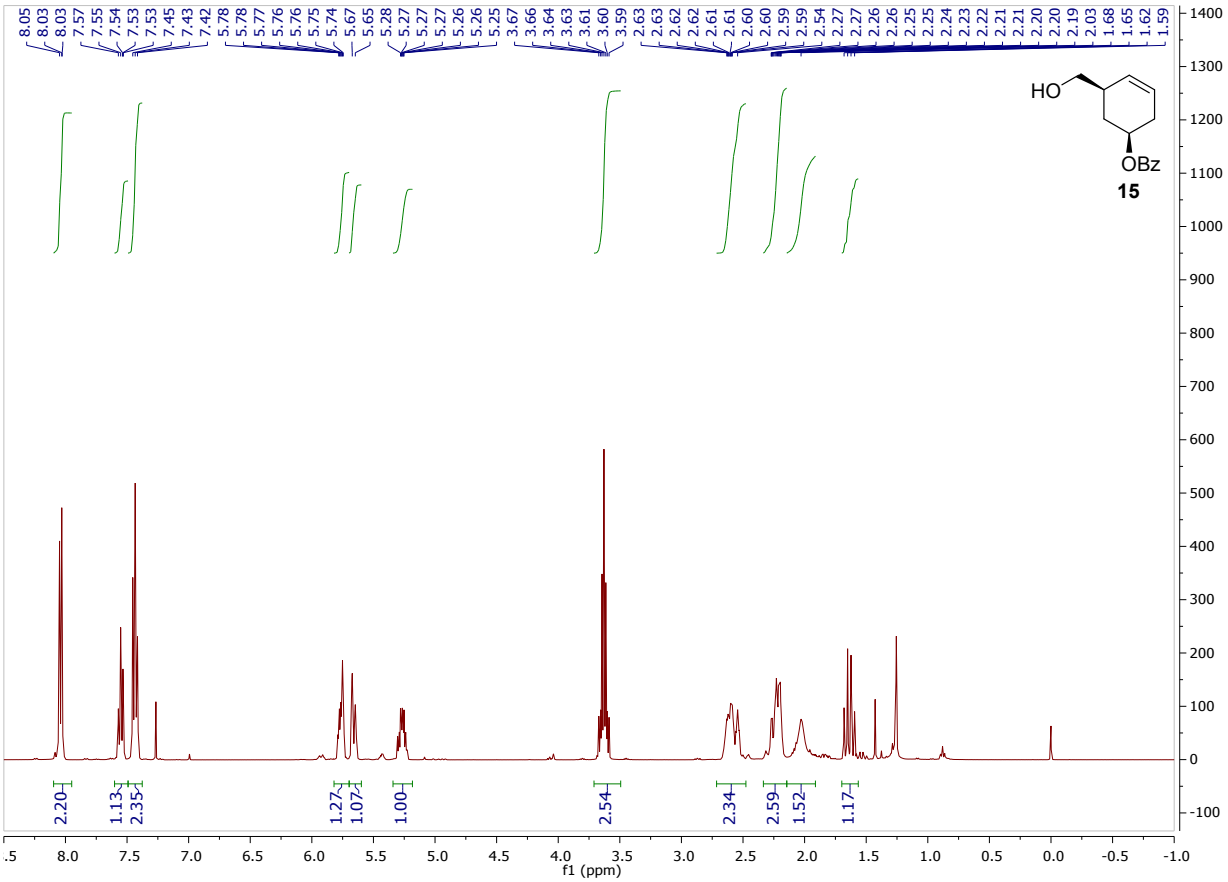

TrtO
O
13
f1 (ppm)

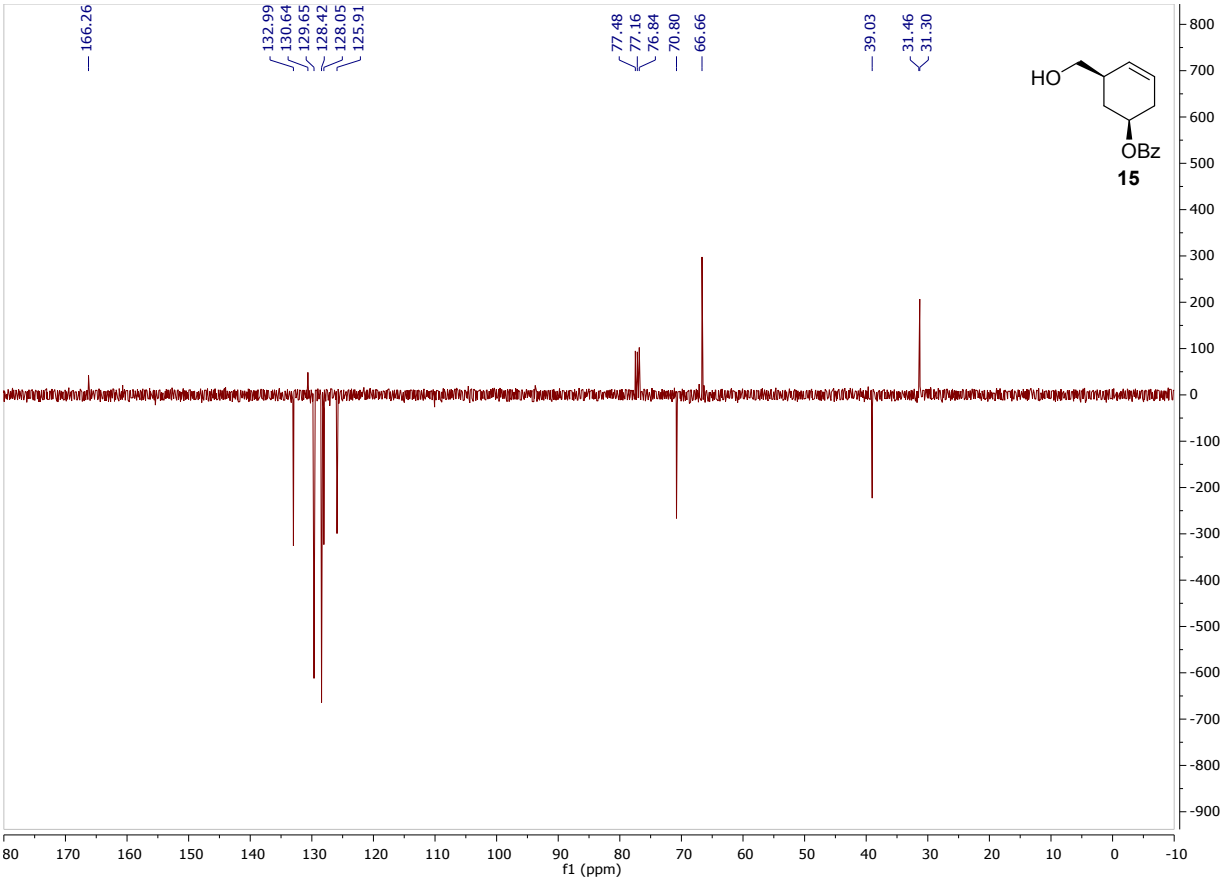

TrtO
OH
14
DCM
stereoisomer
7.24
11.12
1.98
0.16
1.00
2.30
0.98
1.05
1.17
1.27
1.61
3.41
f1 (ppm)

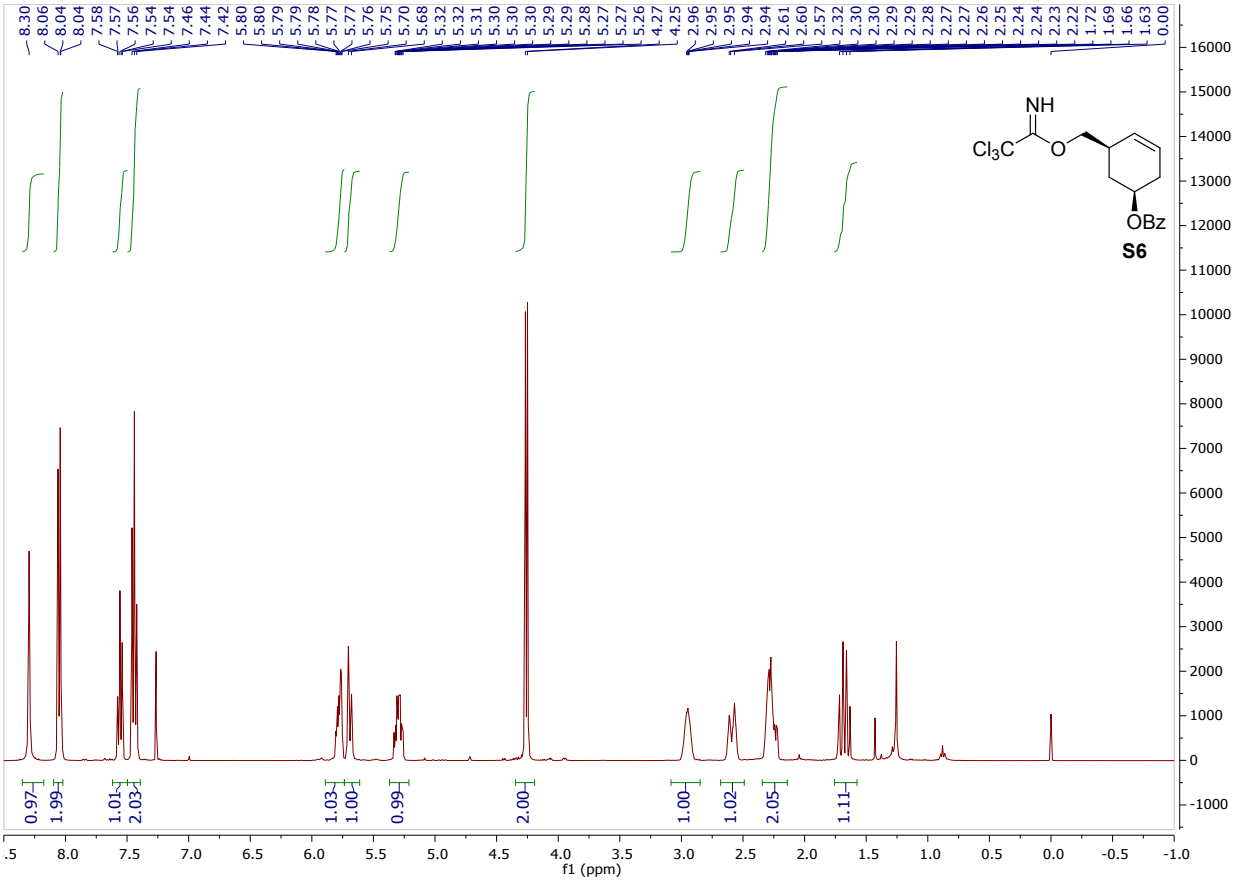

TrtO
OH
14
144.28
131.31
128.84
128.63
128.50
127.86
127.04
125.29
86.46
77.48
77.16
76.84
67.58
67.47
37.57
36.42
35.03
f1 (ppm)

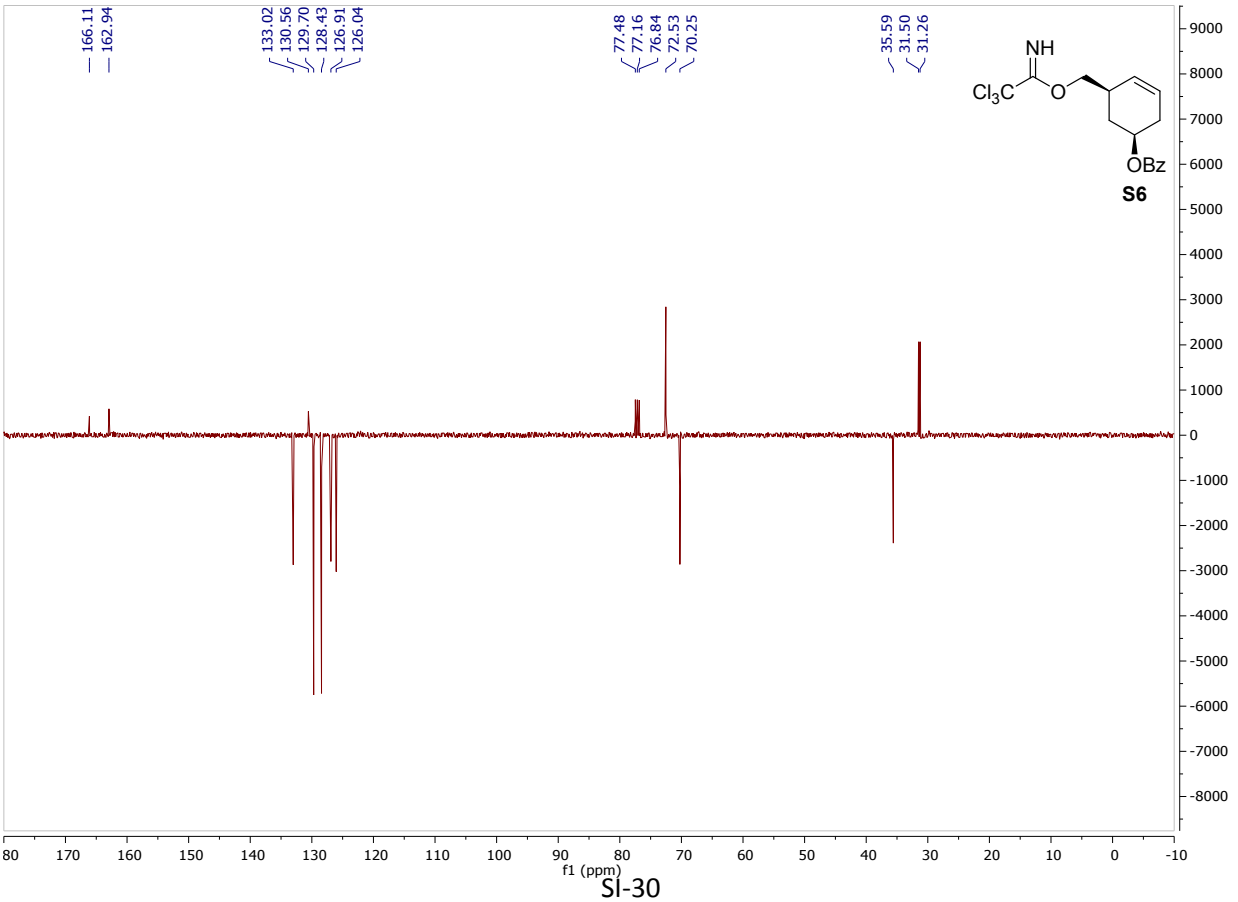

TrtO
OBz
S5
f1 (ppm)

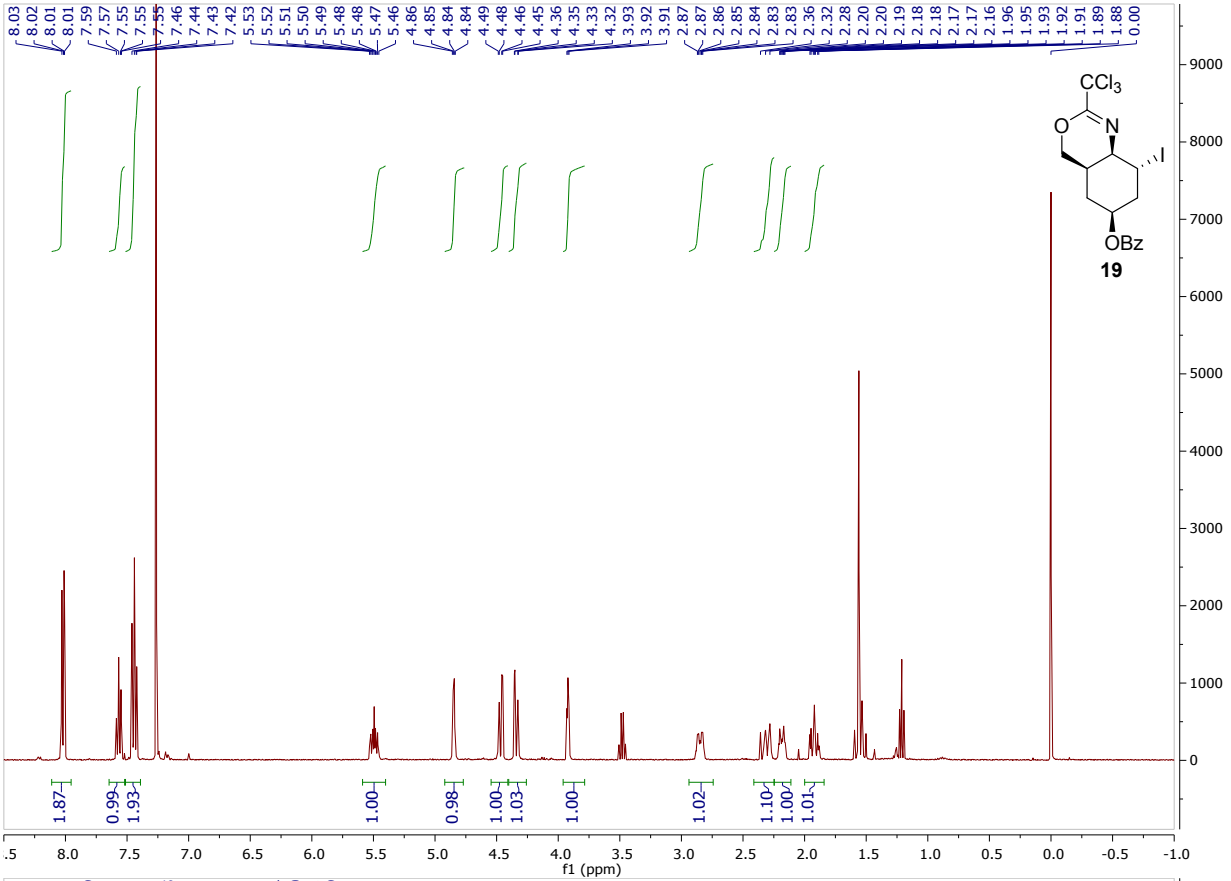

TrtO
OBz
S5
f1 (ppm)

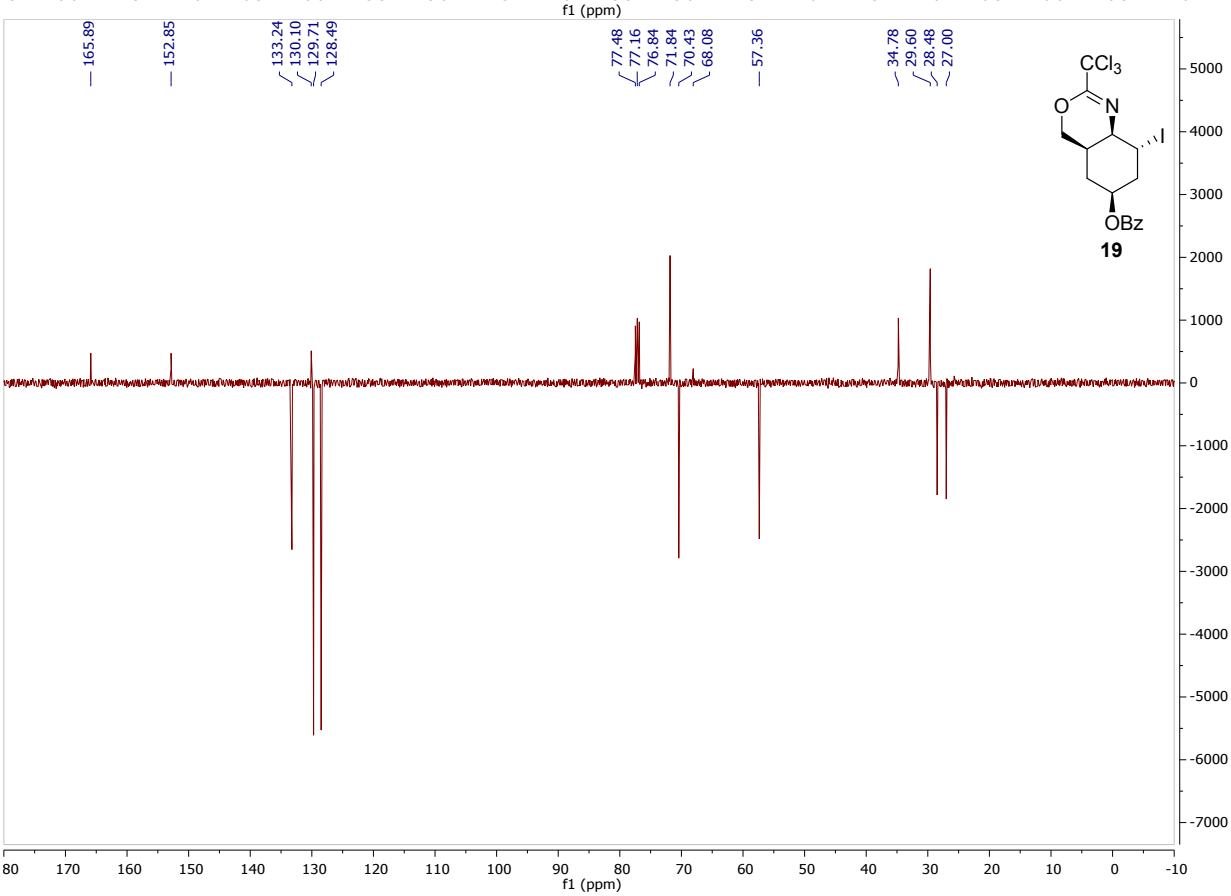

HO
OBz
15
f1 (ppm)

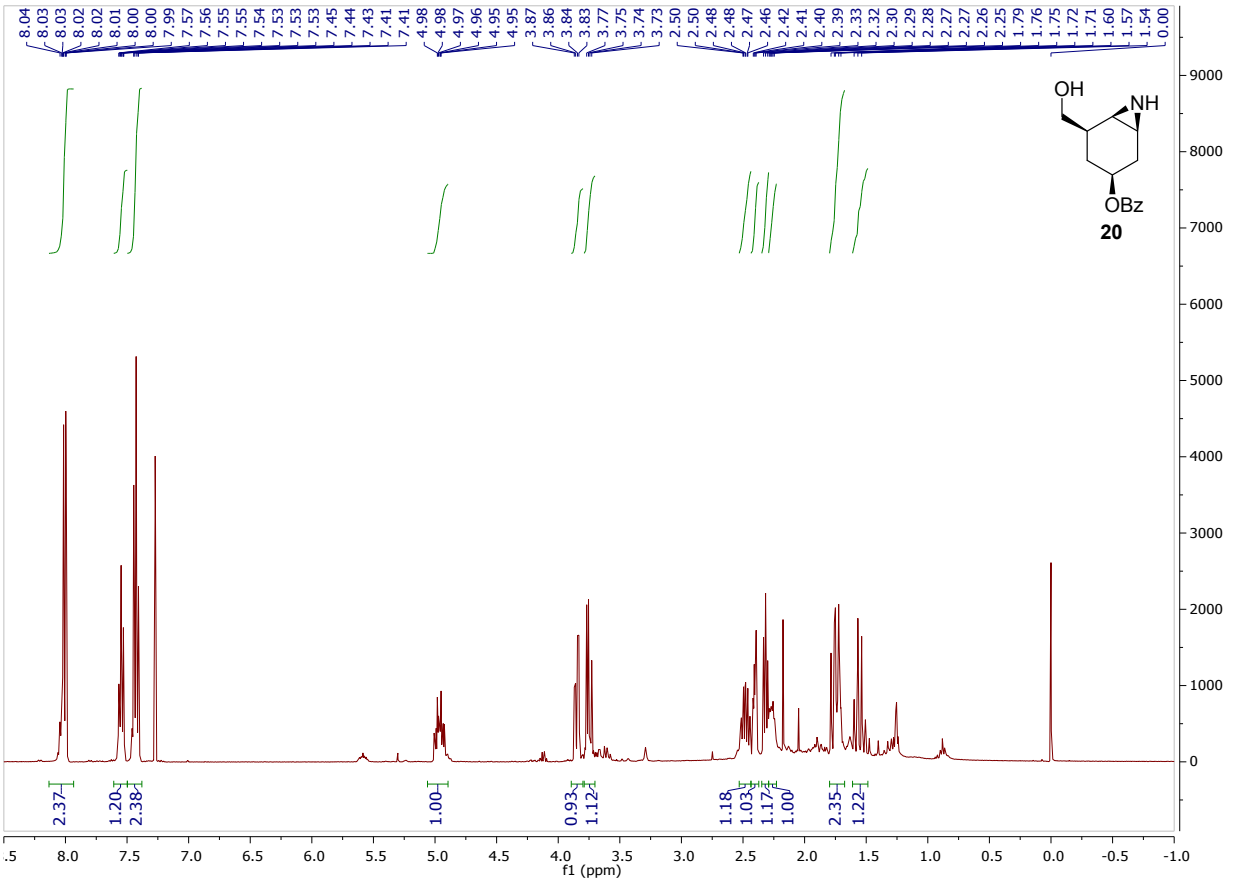

HO
OBz
15
f1 (ppm)

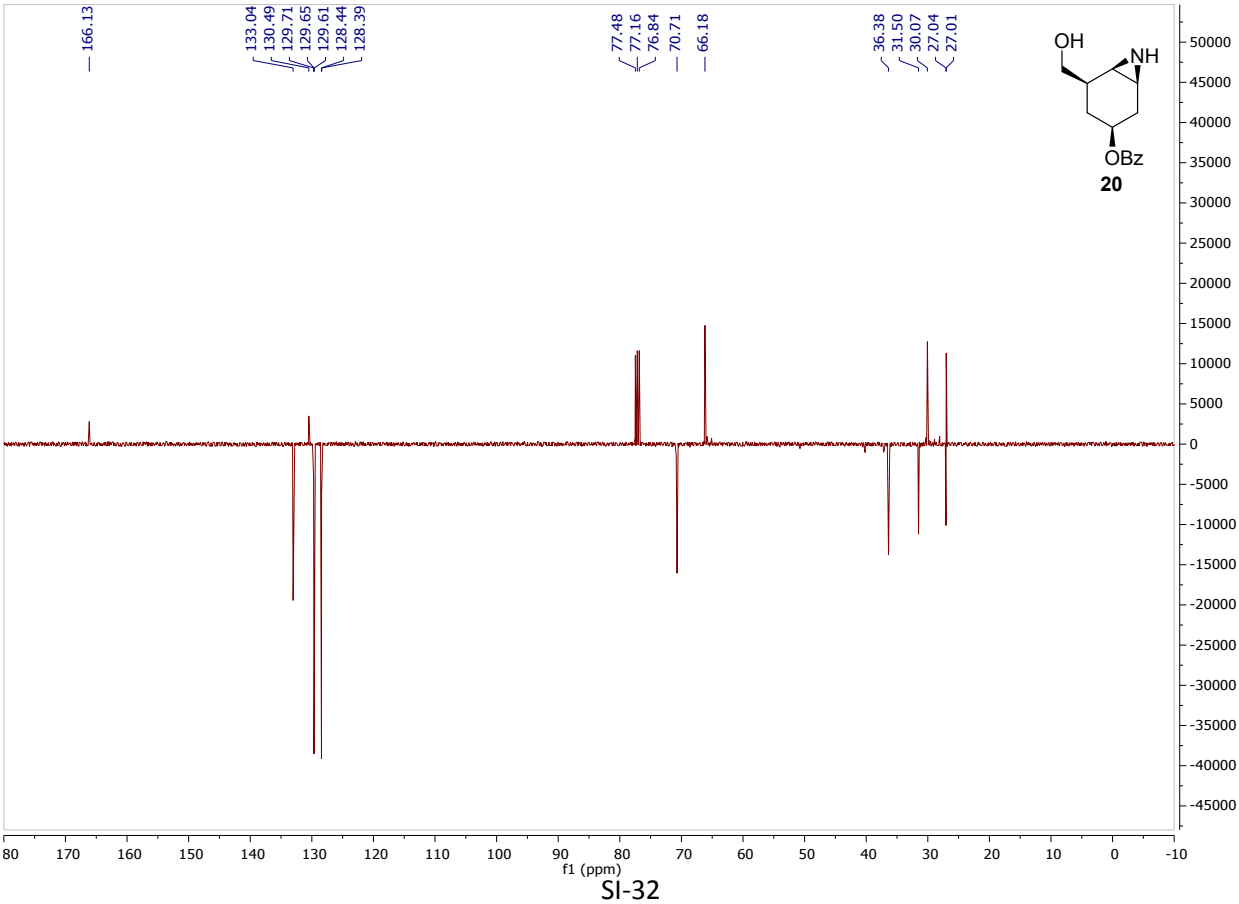

8.30 8.06 8.04 8.04 7.58 7.57 7.56 7.54 7.54 7.46 7.44 7.42 5.80 5.79 5.79 5.78 5.77 5.77 5.76 5.75 5.70 5.68 5.32 5.31 5.30 5.30 5.30 5.29 5.29 5.28 5.27 5.27 5.26 4.27 4.25 2.96 2.95 2.95 2.94 2.94 2.61 2.60 2.57 2.32 2.30 2.30 2.29 2.29 2.28 2.27 2.27 2.26 2.25 2.24 2.24 2.23 2.22 1.72 1.69 1.66 1.63 0.00
NH
Cl3C
O
OBz
S6
0.97 1.99 1.01 2.03 1.03 1.00 0.99 2.00 1.00 1.02 2.05 1.11
f1 (ppm)

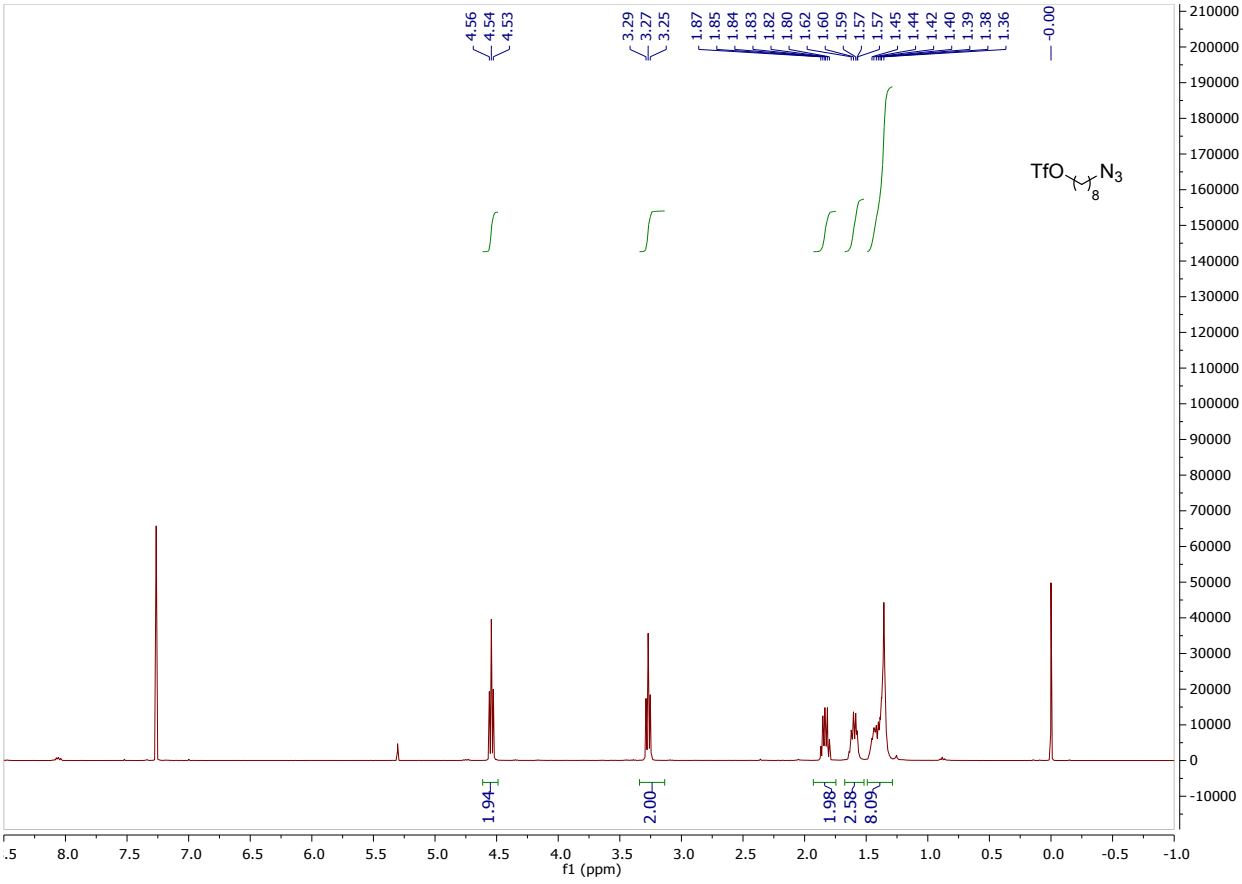

166.11 162.94 133.02 130.56 129.70 128.43 126.91 126.04 77.48 77.16 76.84 72.53 70.25 35.59 31.50 31.26
NH
Cl3C
O
OBz
S6
f1 (ppm)

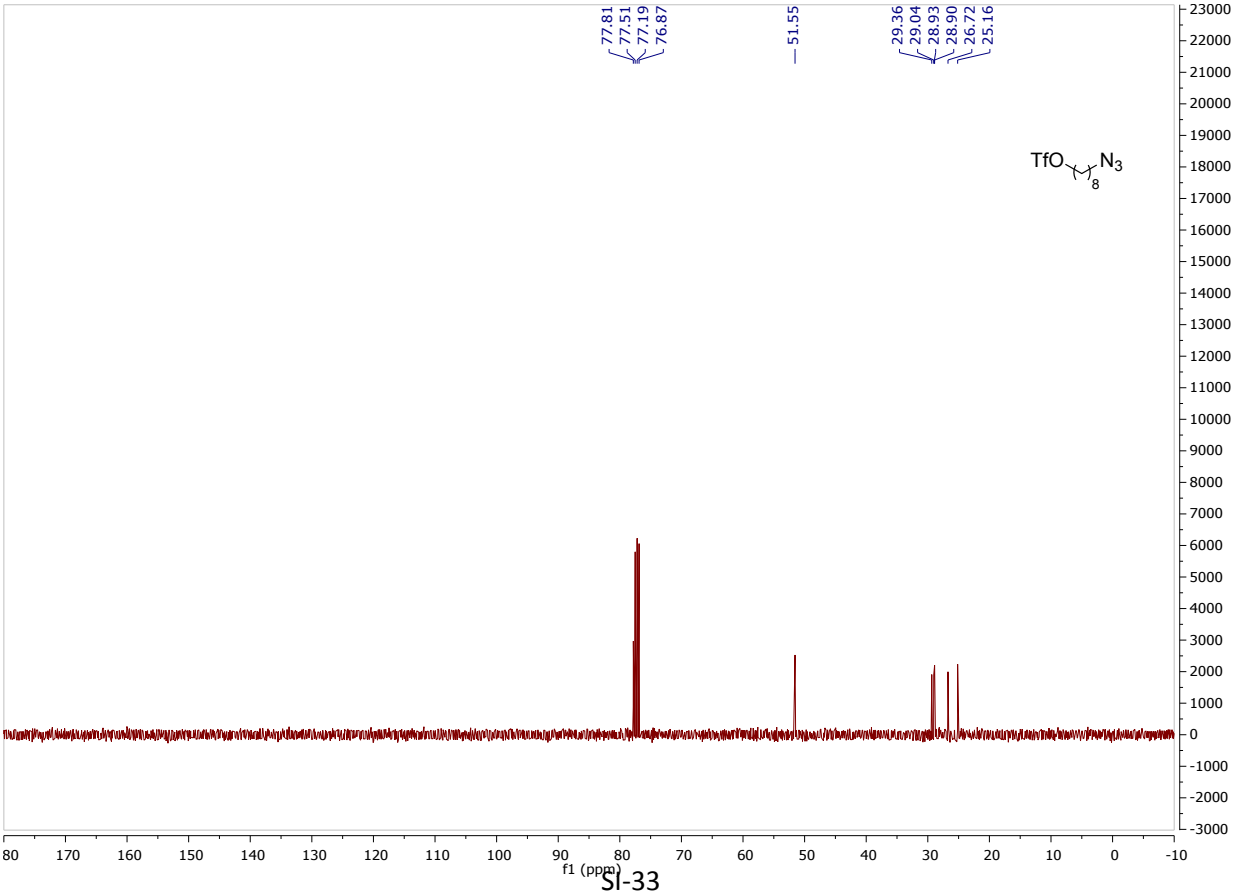

8.03 8.02 8.01 8.01 7.59 7.57 7.55 7.55 7.55 7.46 7.44 7.43 7.42 5.53 5.52 5.51 5.50 5.49 5.48 5.48 5.47 5.46 4.86 4.85 4.84 4.84 4.49 4.48 4.46 4.45 4.36 4.35 4.33 4.32 3.93 3.92 3.91 2.87 2.87 2.86 2.85 2.84 2.83 2.83 2.36 2.32 2.28 2.20 2.20 2.19 2.18 2.18 2.17 2.17 2.16 1.96 1.95 1.93 1.92 1.91 1.89 1.88 0.00
CCl3
O
N
I
OBz
19
1.87 0.99 1.93 1.00 0.98 1.00 1.03 1.00 1.02 1.10 1.00 1.01
f1 (ppm)

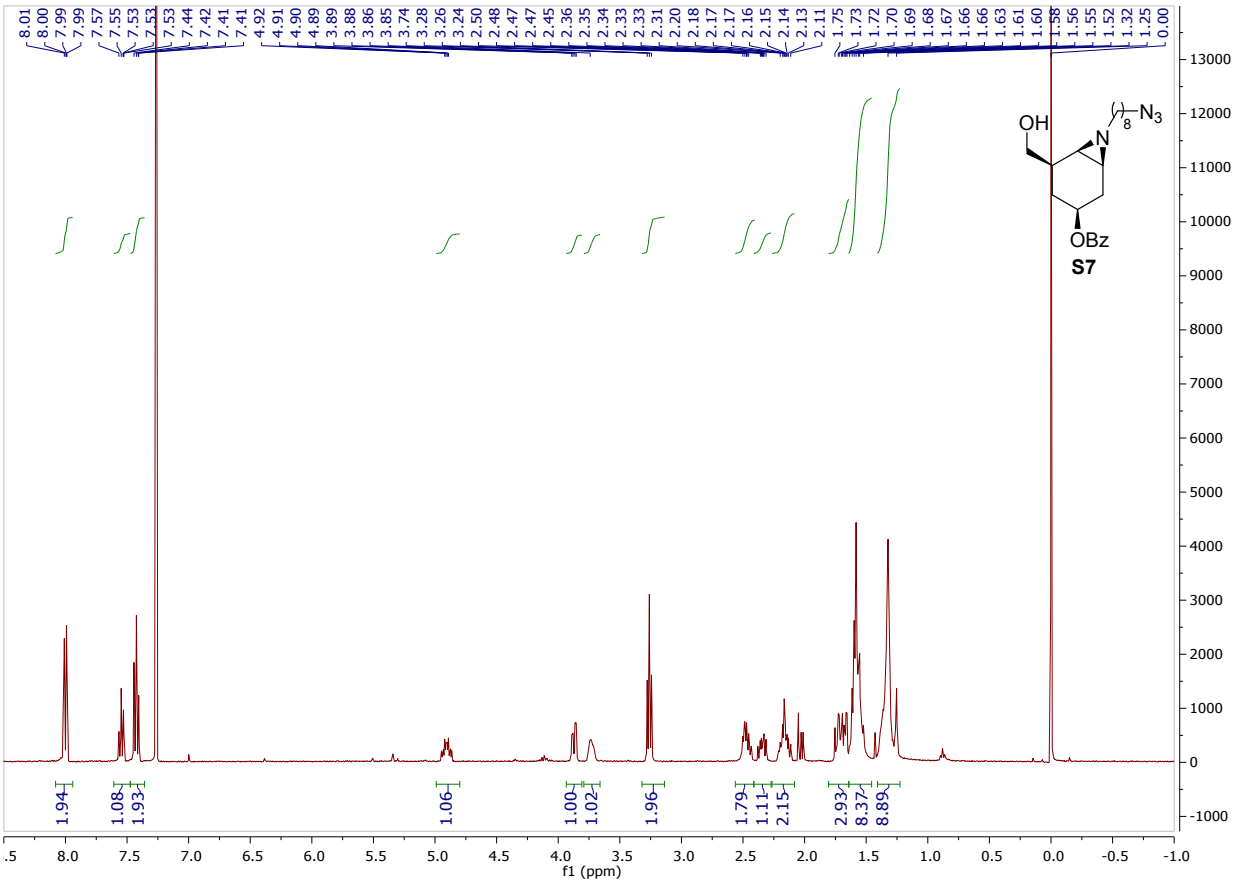

165.89 152.85 133.24 130.10 129.71 128.49 77.48 77.16 76.84 71.84 70.43 68.08 57.36 34.78 29.60 28.48 27.00
CCl3
O
N
I
OBz
19
f1 (ppm)

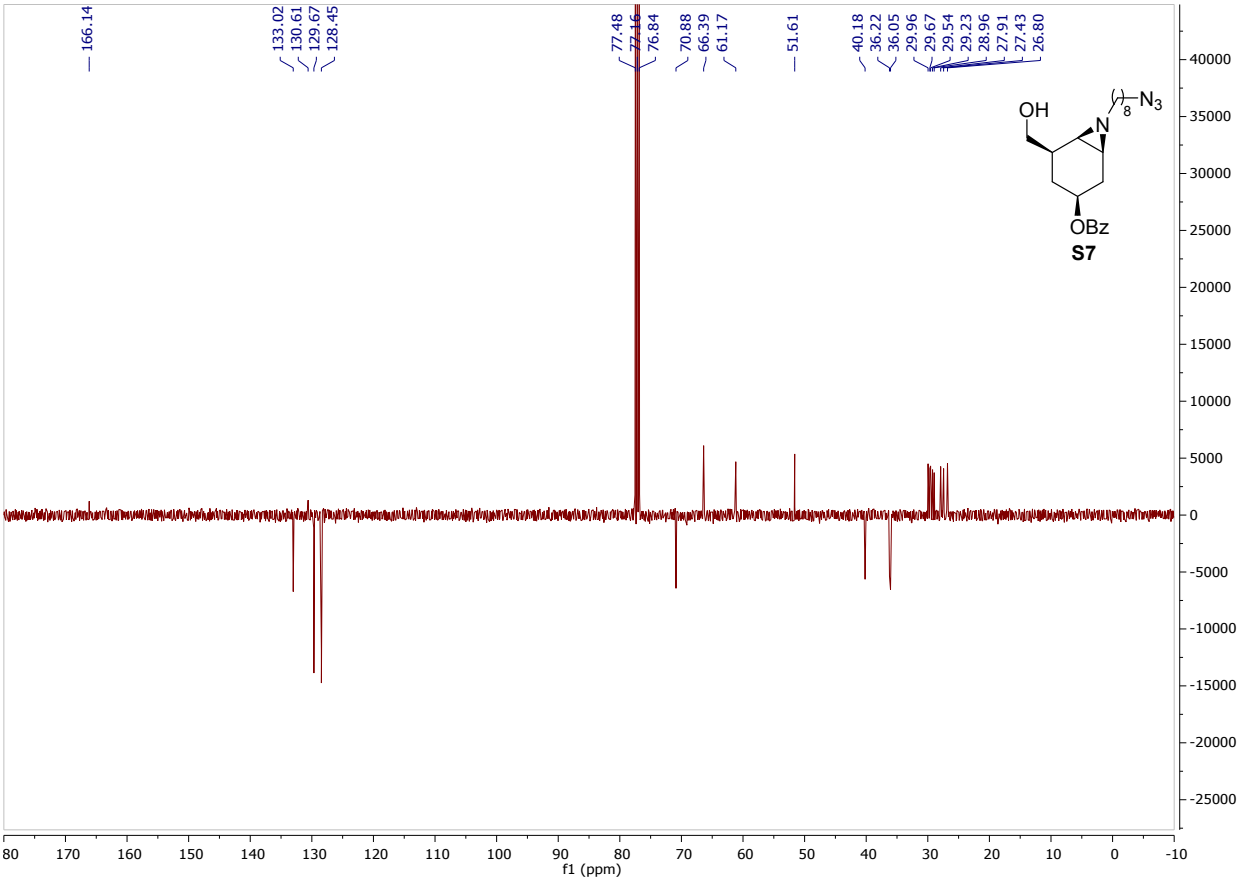

OH
NH
OBz
20
f1 (ppm)

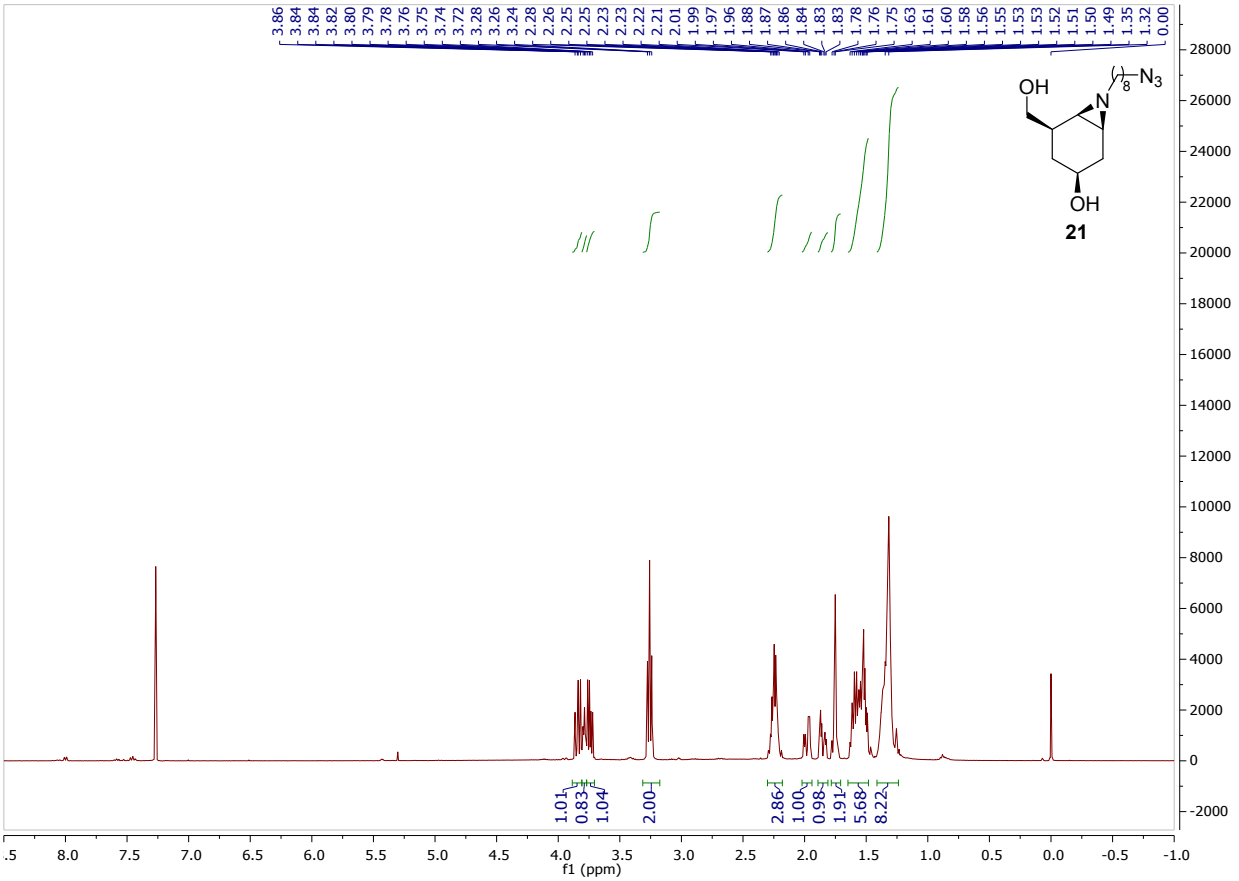

OH
NH
OBz
20
f1 (ppm)

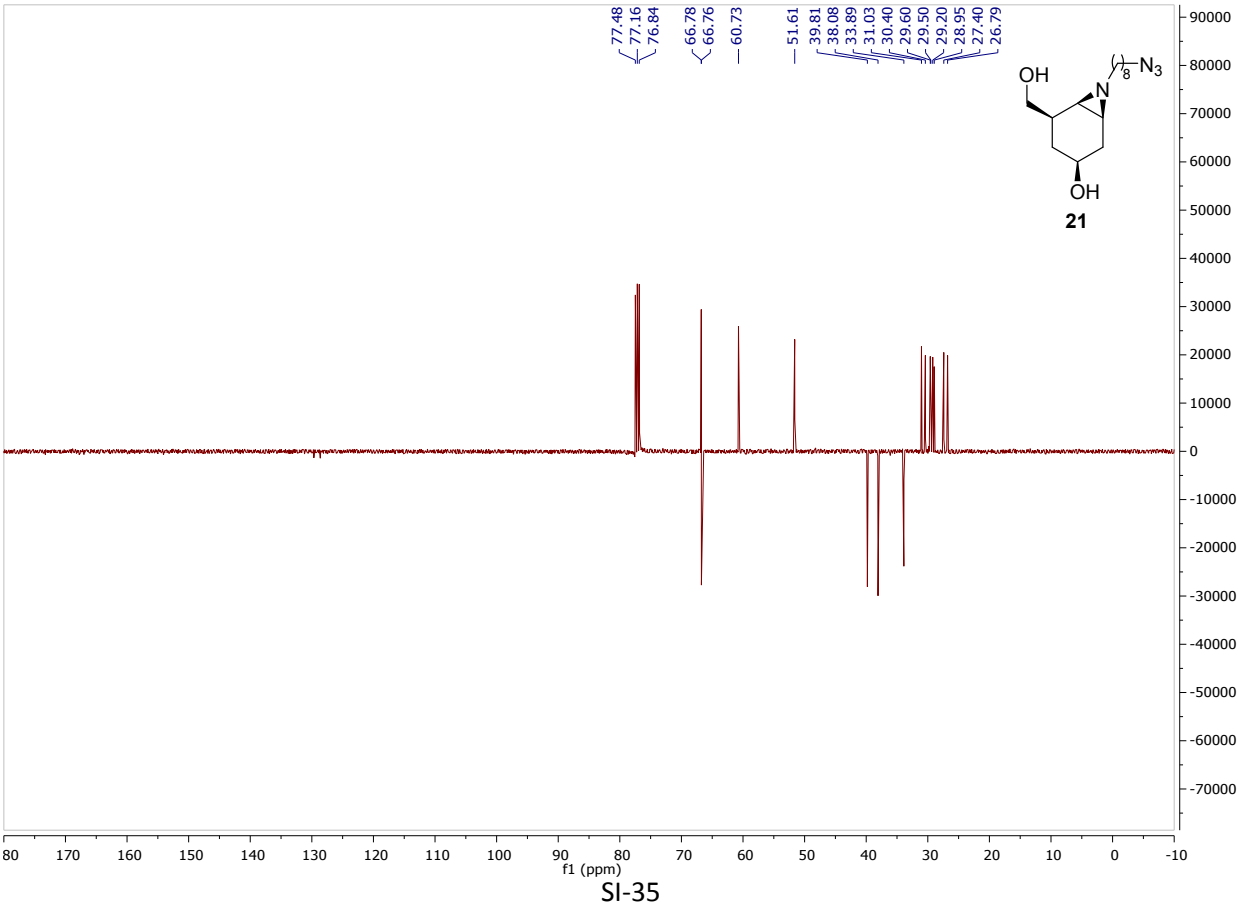

4.56
4.54
4.53
3.29
3.27
3.25
1.87
1.85
1.84
1.83
1.82
1.80
1.62
1.60
1.59
1.57
1.57
1.45
1.44
1.42
1.40
1.39
1.38
1.36
-0.00
TfO
N3
8
1.94
2.00
1.98
2.58
8.09
f1 (ppm)

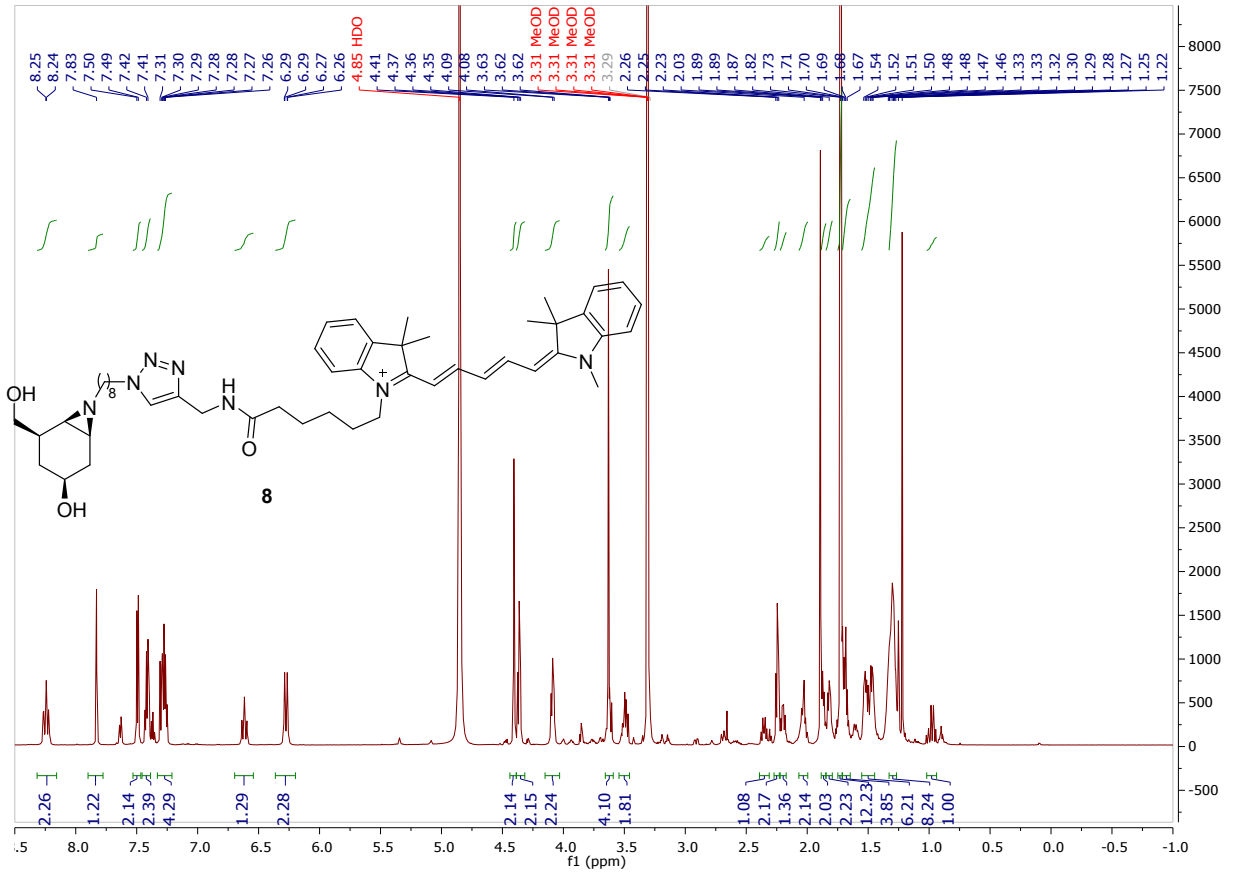

77.81
77.51
77.19
76.87
51.55
29.36
29.04
28.93
28.90
26.72
25.16
TfO
N3
8
f1 (ppm)

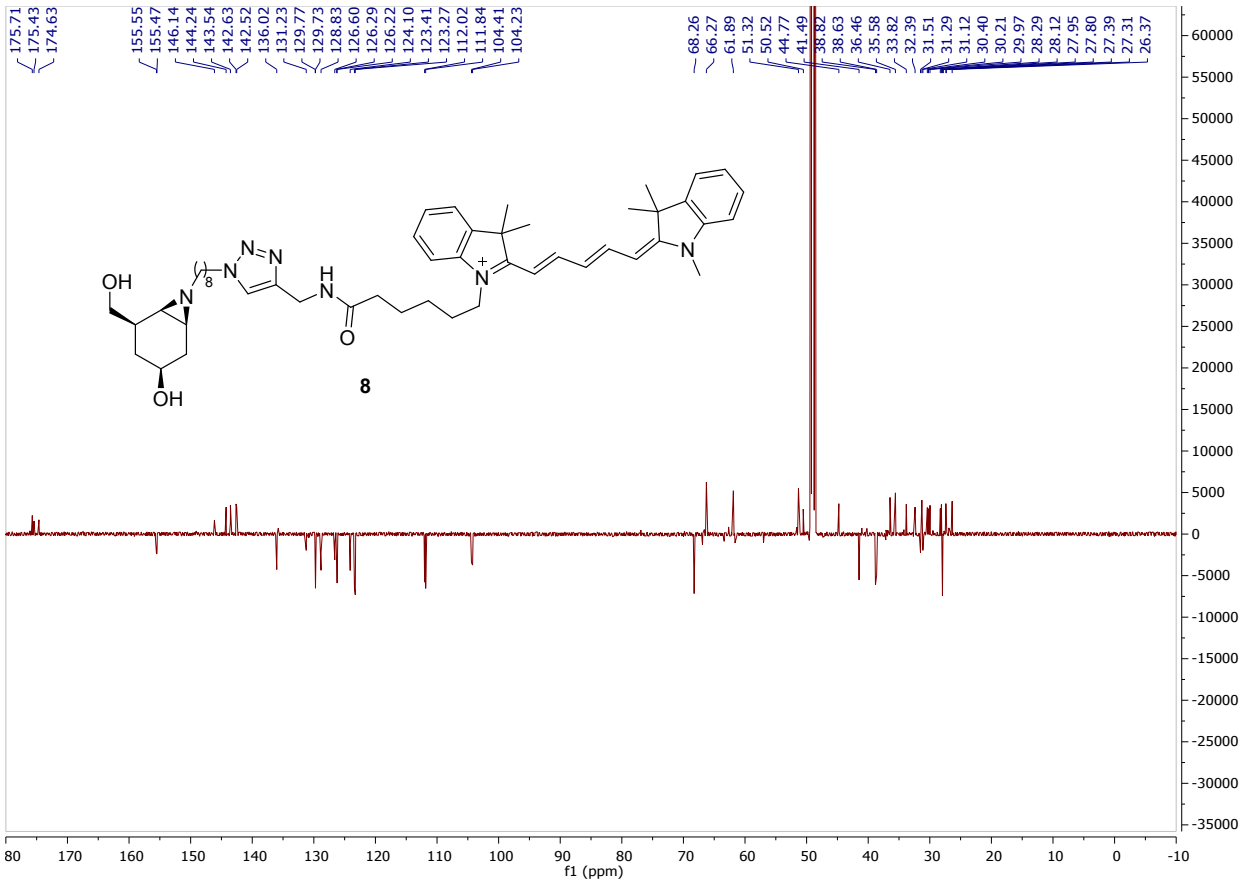

OH
N
N3
8
OBz
S7
f1 (ppm)

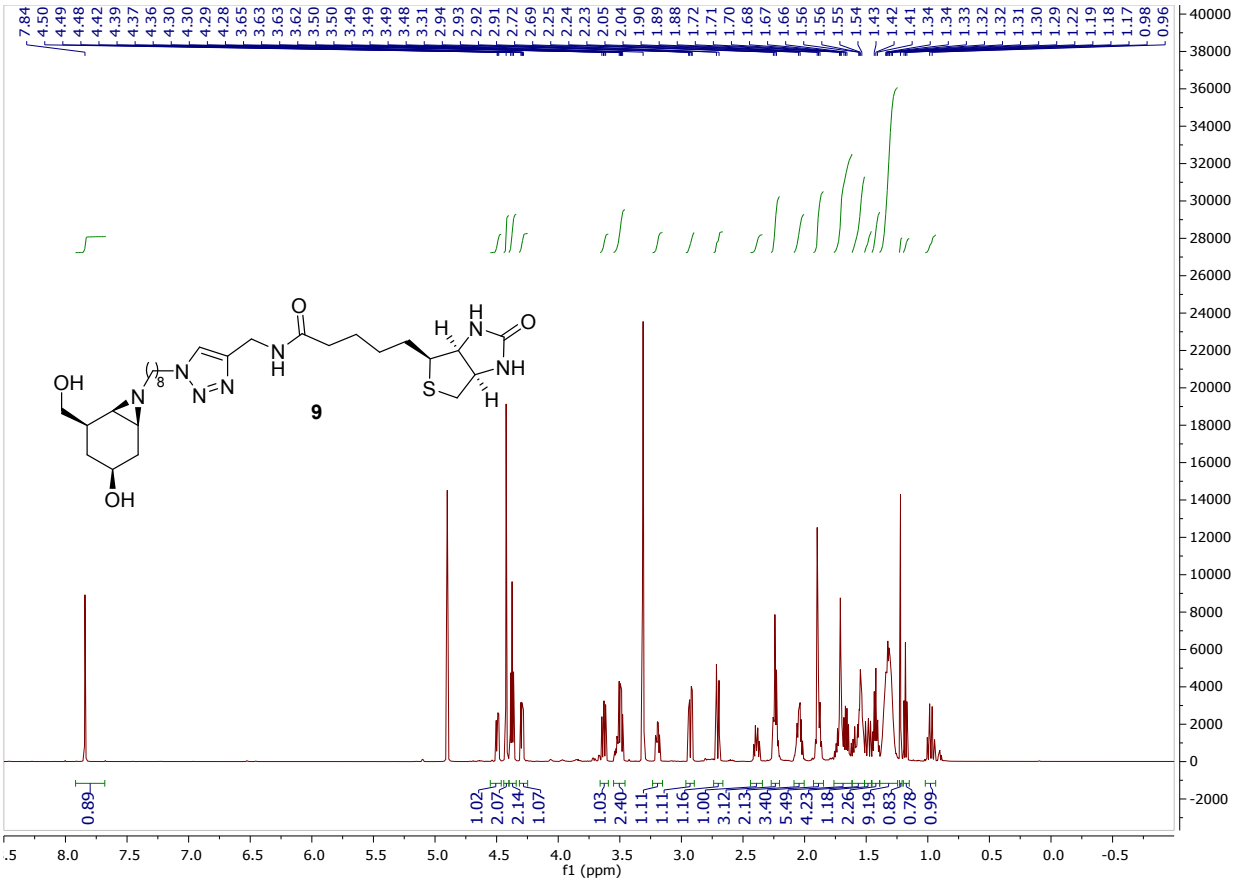

OH
N
N3
8
OBz
S7
f1 (ppm)

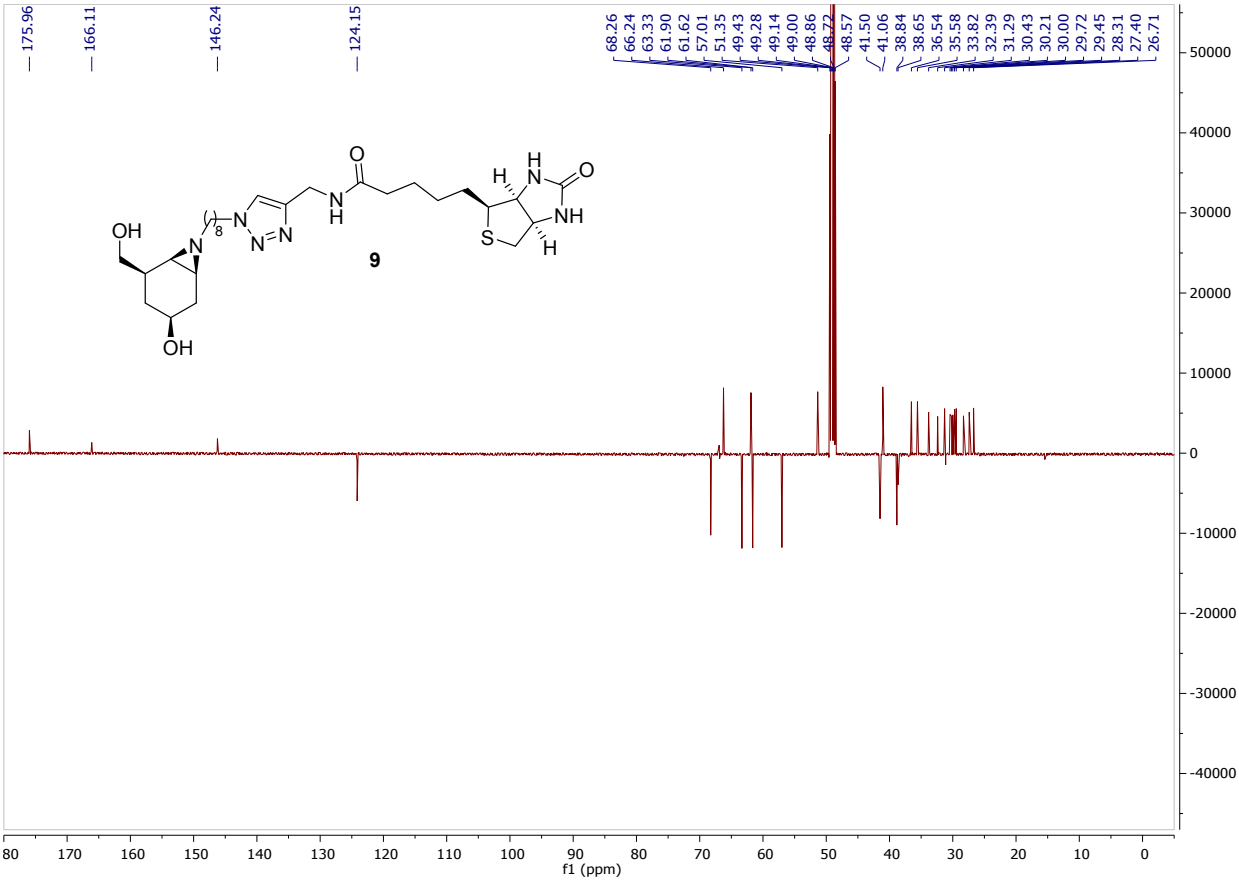

3.86 3.84 3.84 3.82 3.80 3.79 3.78 3.76 3.75 3.74 3.72 3.28 3.26 3.24 2.28 2.26 2.25 2.25 2.23 2.23 2.22 2.21 2.01 1.99 1.97 1.96 1.88 1.87 1.86 1.84 1.83 1.83 1.78 1.76 1.75 1.63 1.61 1.60 1.58 1.56 1.55 1.53 1.53 1.52 1.51 1.50 1.49 1.35 1.32 0.00
OH
N
8
N3
OH
21
1.01 0.83 1.04 2.00 2.86 1.00 0.98 1.91 5.68 8.22
f1 (ppm)


77.48 77.16 76.84 66.78 66.76 60.73 51.61 39.81 38.08 33.89 31.03 30.40 29.60 29.50 29.20 28.95 27.40 26.79
OH
N
8
N3
OH
21
f1 (ppm)


8.25
8.24
7.83
7.50
7.49
7.42
7.41
7.31
7.30
7.29
7.28
7.27
7.26
6.29
6.29
6.27
6.26
4.85 HDO
4.41
4.37
4.36
4.35
4.09
4.08
3.63
3.62
3.62
3.31 MeOD
3.31 MeOD
3.31 MeOD
3.31 MeOD
3.29
2.26
2.25
2.23
2.03
1.89
1.89
1.87
1.82
1.73
1.71
1.70
1.69
1.69
1.67
1.54
1.52
1.51
1.50
1.48
1.48
1.47
1.46
1.33
1.33
1.32
1.30
1.29
1.28
1.27
1.25
1.22
2.26
1.22
2.14
2.39
4.29
1.29
2.28
2.14
2.15
2.24
4.10
1.81
1.08
2.17
1.36
2.14
2.03
2.23
12.23
3.85
6.21
8.24
1.00
f1 (ppm)
OH
OH
N
N
N
H
N
O
N
N
8


175.71
175.43
174.63
155.55
155.47
146.14
144.24
143.54
142.63
142.52
136.02
131.23
129.77
129.73
128.83
126.60
126.29
126.22
124.10
123.41
123.27
112.02
111.84
104.41
104.23
68.26
66.27
61.89
51.32
50.52
44.77
41.49
38.82
38.63
36.46
35.58
33.82
32.39
31.51
31.29
31.12
30.40
30.21
29.97
28.29
28.12
27.95
27.80
27.39
27.31
26.37
f1 (ppm)
OH
OH
N
N
N
H
N
O
N
N
8


9
f1 (ppm)


9
f1 (ppm)

Biochemical experiments

Materials

All used epoxide inhibitors **1**[4], **2**[7] and **3**[8] were synthesized as previously described and obtained from the in-house Dutch Compound Library (DCL). Bicinchoninic acid (BCA) protein determination kit was acquired from Pierce Chemical Company (Rockford, IL). Recombinant β-glucocerebrosidase (GBA1) was obtained from Genzyme (Imiglucerase, Cerezyme). Overexpressed bacterial β-galactosidase (*Cj*GH35A from *Cellvibrio japonicus*) was obtained as previously described.[9] Recombinant β-mannosidase (*Helix pomatia*) was obtained from Sigma Aldrich. Primary human fibroblasts (product code CC-2511, lot no. C92030) were obtained from Lonza. Mouse kidneys and livers were obtained from Jackson's laboratories (C57Bl6/J). 4-methylumbelliferyl β-D-glucopyranoside was obained from Glycosynth Limited (Winwick Quay, Warrington, UK) and 4-methylumbelliferyl β-D-galactopyranoside was obtained from Sigma Aldrich. Trypsin was commercially available from Promega, DynaBead MyOne Straptavidin Beads C1 were obtained from Invitrogen, dithiothreitol (DTT) was obtained from Biochemica and Iodoacetamide (IAA) was obtained from Sigma Aldrich. All other chemicals were obtained from standard commercial sources.

Cell culture and lysate preparation

Primary human fibroblasts (product code CC-2511, lot no. C92030, Lonza) were grown in Dulbecco's modified Eagle's medium: Nutrient Mixture F-12 (DMEM/F-12, Invitrogen), containing 10% (v/v) heat-inactivated fetal calf serum, 200 µg/mL penicillin, 200 µg/mL streptomycin at 37 °C, 7% CO_2 (New Brunswick Galaxy 170R incubator). Fibroblasts were subcultured in the same medium at a ratio of 1:3 or 1:4 each one and a half week, grown to a 70-90% confluence and were detached from the plates by scraping in 5 mL phosphate-buffered saline (PBS). Collected fibroblasts were centrifuged for 5 min at 1500 rpm. Pellets were resuspended into ice-cold lysis buffer (25 mM phosphate buffer pH 6.5 supplemented with 0.1% (v/v) Triton X-100 and protease inhibitor cocktail tablet Roche (version 12)). The suspension was then vortexed vigorously and lysated on ice. To homogenize, lysates were freeze-thawed using liquid nitrogen. Obtained lysates were stored in aliquots at -80 °C. Concentrations of lysates were determined by using the BCA method.[10]

Lysate preparation of mouse tissue homologues

Mouse kidneys and mouse livers were cut fine and were grinded by a MP FastPrep-24 5G with a CoolPrep 24x2 mL Sample Holder in ice-cold lysis buffer (50 mM phosphate buffer pH 6.5 supplemented with 0.1% (v/v) Triton X-100 and protease inhibitor cocktail AMRESCO mammalian M250) by using 1.0 mm Glass Beads (Cat.No. 11079110, BioSpec Products). The obtained mixture was transferred and centrifuged for 15 minutes at 1200 rpm. The supernatant was transferred again and centrifuged for 20 minutes at full speed, to separate the membranes and cytosolic fraction. Concentrations of the supernatant were determined by using the BCA method.[10] Aliquots were stored at -80 °C.

Labelling and SDS-PAGE of recombinant enzymes

Recombinant glycosidases were labelled with ABPs **5**, **6** and **8** at 37˚C for 30 minutes with the optimized conditions for each individual enzyme; 2.5 pmol GBA1 (Cerezyme) in 150 mM McIlvaine buffer pH 5.2 supplemented with 0.1% Triton X-100, 0.2% (w/v) sodium taurocholate, 13 pmol *Cj*GH35A (β-galactosidase) in 150 mM McIlvaine buffer pH 4.5 and 10 pmol β-mannosidase (*Helix pomatia*) in 150 mM McIlvaine buffer pH 5.0 supplemented with 0.1% (w/v) bovine serum albumin (BSA) for stabilization of the recombinant proteins. Samples were denatured with sample buffer (4x Laemmli buffer, containing 50% (v/v) 1M Tris-HCl pH 6.8, 50% (v/v) 100% glycerol, 10% (w/v) Dithiothreitol (DTT), 10% (w/v) sodium dodecyl sulphate (SDS), 0.01% bromophenol blue) and heated at 100 ˚C for 5 minutes. Proteins were resolved by electrophoresis in sodium dodecylsulfate (SDS-PAGE) 10% polyacrylamide gels, running at a constant of 90V for 30 minutes followed by 120V for approximately 60 minutes. Wet slab gels were scanned on fluorescence using a Typhoon FLA9500 Imager (GE Healthcare) using λ_{EX} 635 nm; λ_{EM} > 665 nm. Images were acquired, processed and quantified with Image Quant (GE Healthcare).

Labelling and SDS-PAGE of recombinant enzymes competed with epoxide inhibitors or denaturation

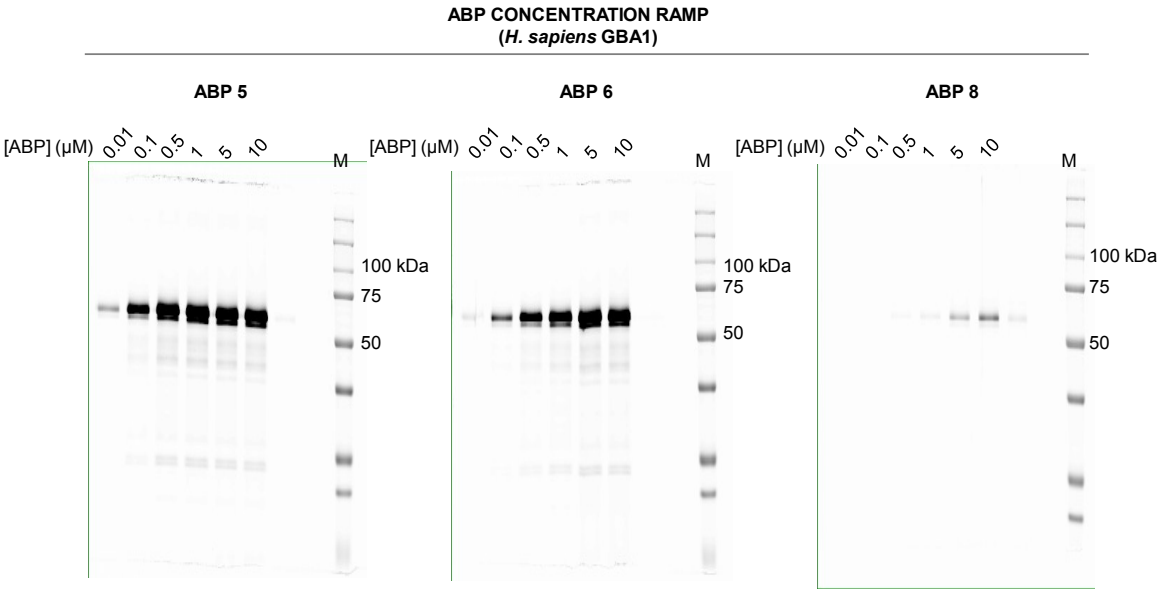
Recombinant glycosidases were dissolved in the appropriate buffer as described above, and incubated with 100 µM of matching cyclophellitol epoxides **1**, **2** or **3** at 37 ˚C for 1 hour. Then, the enzymes were labelled with the matching ABP **5** (100 nM), **6** (500 nM) or **8** (10 µM) at 37 ˚C for 30 minutes. Additionally a positive control was performed; the enzyme was incubated in buffer supplemented with 1% DMSO for 1 h at 37 °C, and subsequently labelled with the matching ABP with the concentration described above. A negative control was also performed, which was acquired by heat-inactivation of proteins (adding 4x Laemmli buffer and heating at 100 °C for 5 minutes) prior to ABP labelling. Proteins were denatured with sample buffer (4x Laemmli buffer, containing 50% (v/v) 1M Tris-HCl pH 6.8, 50% (v/v) 100% glycerol, 10% (w/v) Dithiothreitol (DTT), 10% (w/v) sodium dodecyl sulphate (SDS), 0.01% bromophenol blue) and heated at 100˚C for 5 minutes. Proteins were resolved by SDS-PAGE 10% polyacrylamide gels. Wet slab gels were scanned on fluorescence using a Typhoon FLA9500 Imager (GE Healthcare) using λ_{EX} 635 nm; λ_{EM} > 665 nm. Images were acquired, processed and quantified with Image Quant (GE Healthcare).

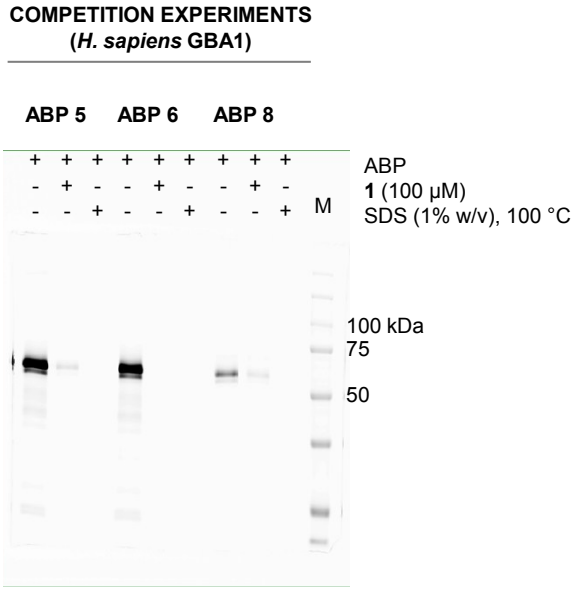
***In vitro* labelling of lysates and SDS-PAGE**

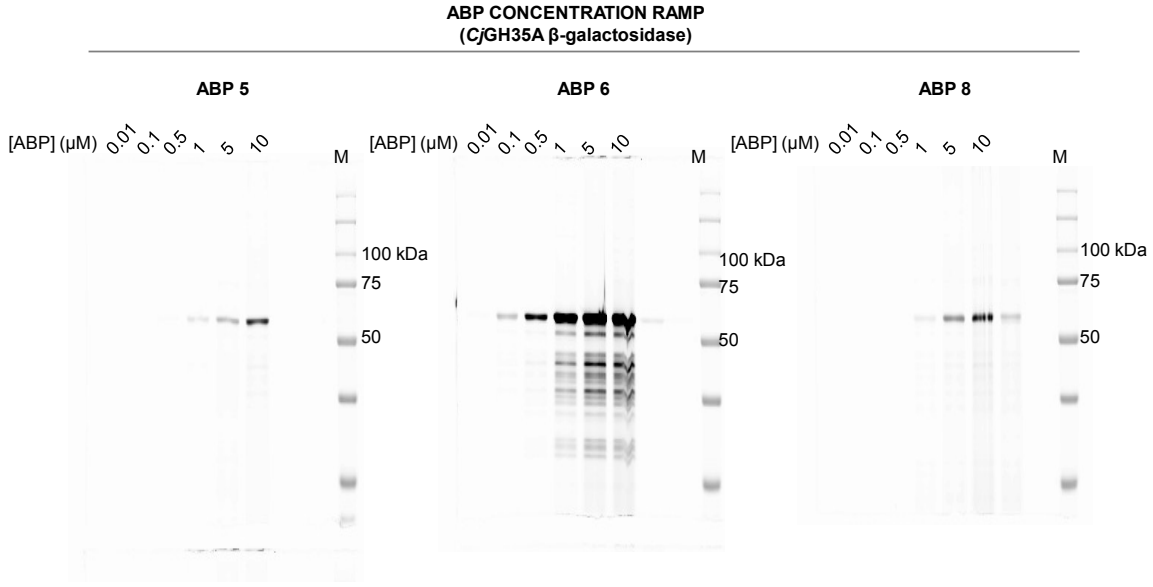
Mouse kidney lysates (50 µg total protein per sample), mouse liver lysates (40 µg total protein per sample) or human fibroblasts lysates (10 µg total protein per sample) were dissolved in McIlvaine buffer pH 5.0 for 5 minutes on ice. The samples were incubated with epoxide inhibitors **1** and/or **2** (100 µM), or DMSO (1% v/v) for 1 hour at 37˚C, followed by incubation with 1 µM ABP **5**, **6** or **8** for 30 minutes at 37˚C. Samples were denatured with sample buffer (4x Laemmli buffer, containing 50%

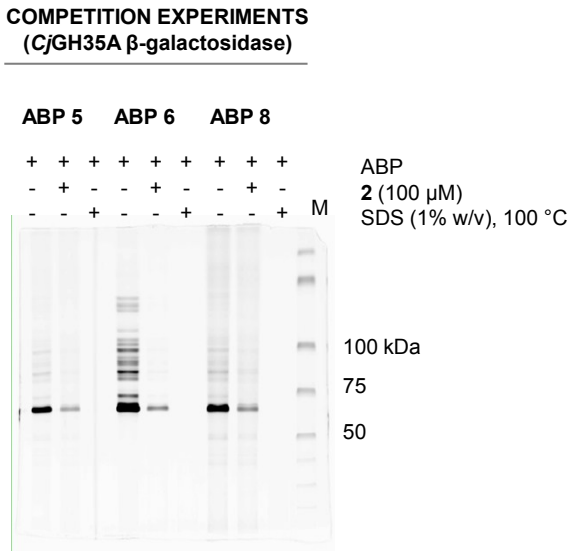
(v/v) 1M Tris-HCl pH 6.8, 50% (v/v) 100% glycerol, 10% (w/v) Dithiothreitol (DTT), 10% (w/v) sodium dodecyl sulphate (SDS), 0.01% bromophenol blue) and heated at 100 °C for 5 minutes. Proteins were resolved on 10% acrylamide SDS-PAGE. Wet slab gels were scanned on fluorescence using a Typhoon FLA9500 Imager (GE Healthcare) using λ_{EX} 635 nm; λ_{EM} > 665 nm. Images were acquired, processed and quantified with Image Quant (GE Healthcare). Gels were subsequently extensively stained with coomassie brilliant blue (CBB-G250) and de-stained with destaining solution (10:0.625:0.875 Water:methanol:acetic acid) until no background remained. Gels were scanned on a regular flatbed scanner in transparency modus. Images were acquired, processed and quantified with Image Quant (GE Healthcare).

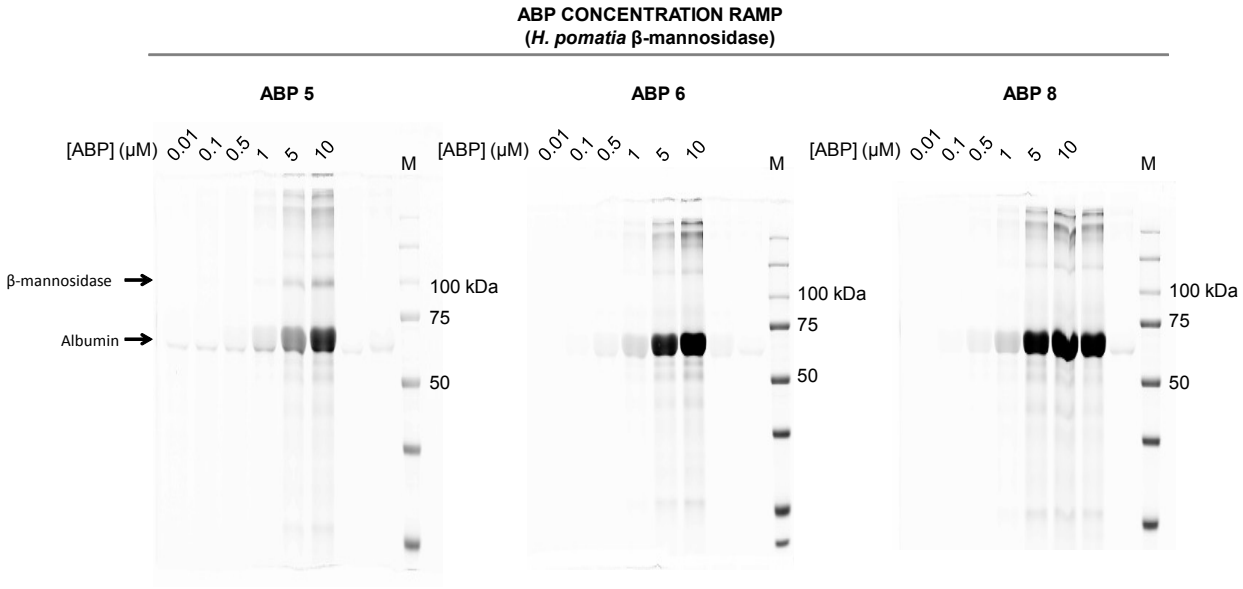
Recombinant enzyme SDS-PAGE gels

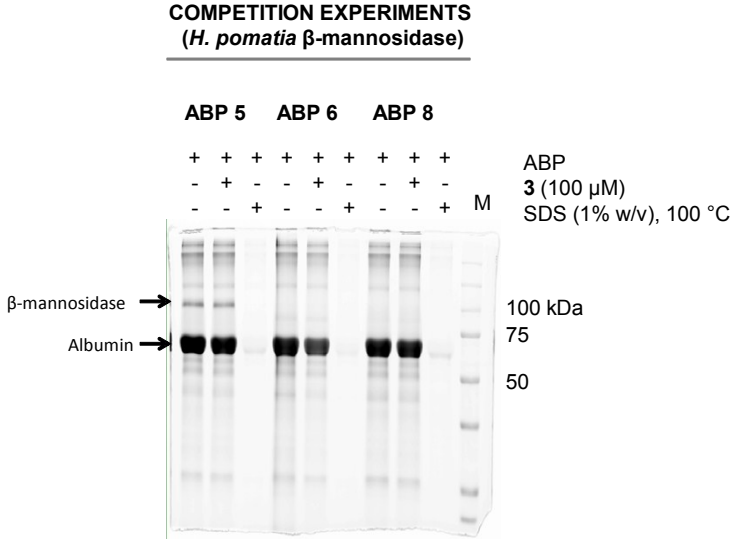




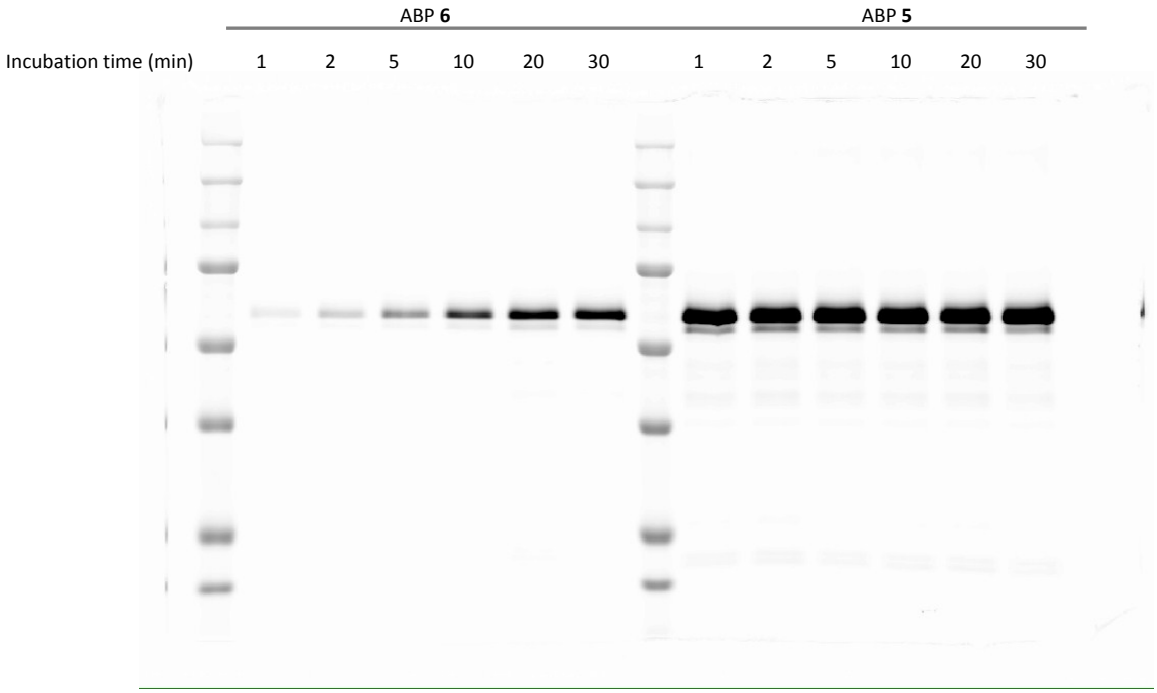

ABP CONCENTRATION RAMP
(CjGH35A β-galactosidase)
ABP 5
ABP 6
ABP 8
[ABP] (μM) 0.01 0.1 0.5 1 5 10
M
100 kDa
75
50


COMPETITION EXPERIMENTS
(CjGH35A β-galactosidase)
ABP 5 ABP 6 ABP 8
ABP
2 (100 μM)
SDS (1% w/v), 100 °C
M
100 kDa
75
50

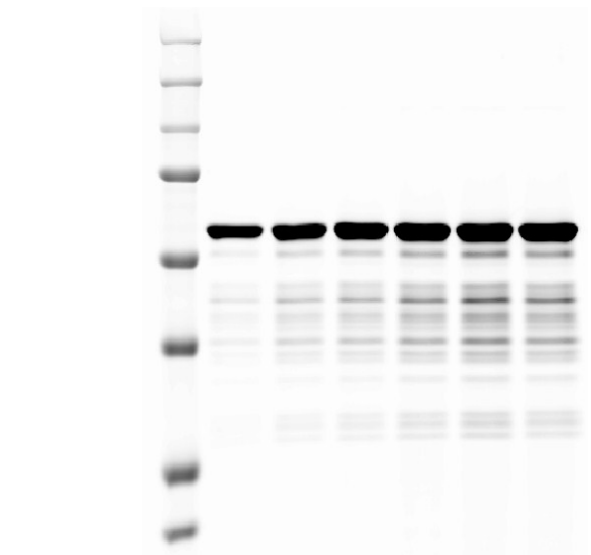

ABP CONCENTRATION RAMP
(H. pomatia β-mannosidase)
ABP 5
ABP 6
ABP 8
[ABP] (μM) 0.01 0.1 0.5 1 5 10
M
β-mannosidase
Albumin
100 kDa
75
50


COMPETITION EXPERIMENTS
(H. pomatia β-mannosidase)
ABP 5 ABP 6 ABP 8
ABP
3 (100 μM)
SDS (1% w/v), 100 °C
M
β-mannosidase
Albumin
100 kDa
75
50

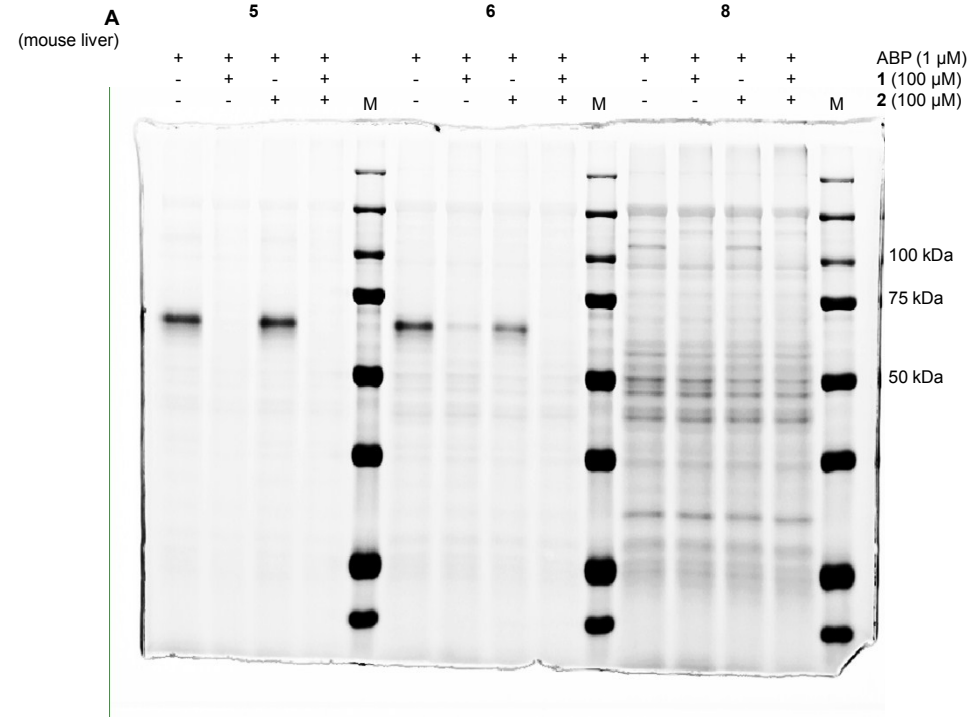
TIME DEPENDENT LABELING OF GBA1 WITH ABPs **5** AND **6**

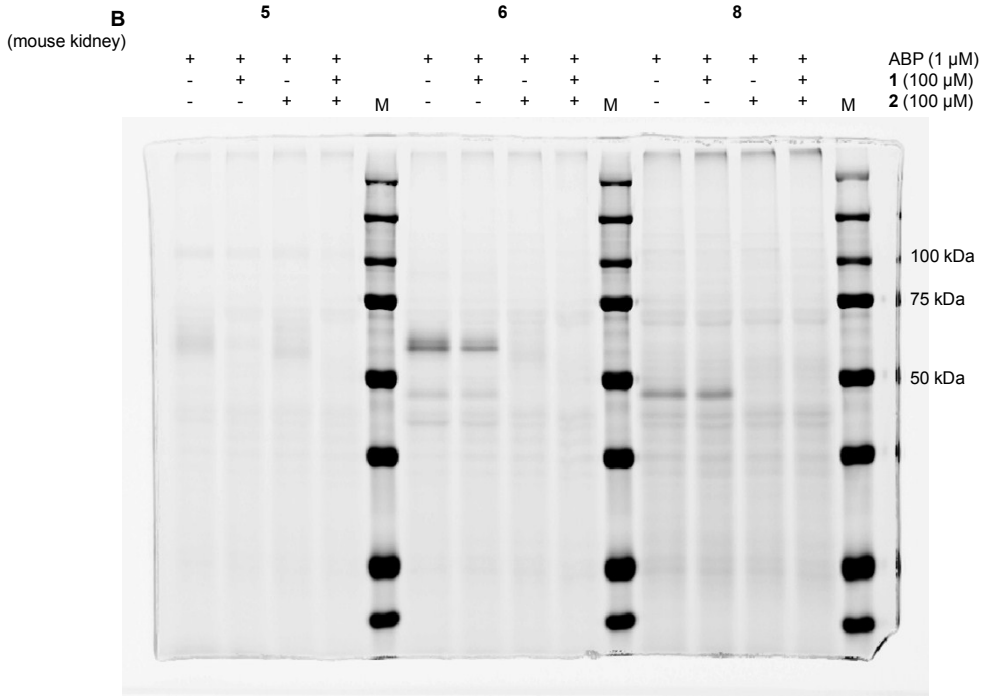


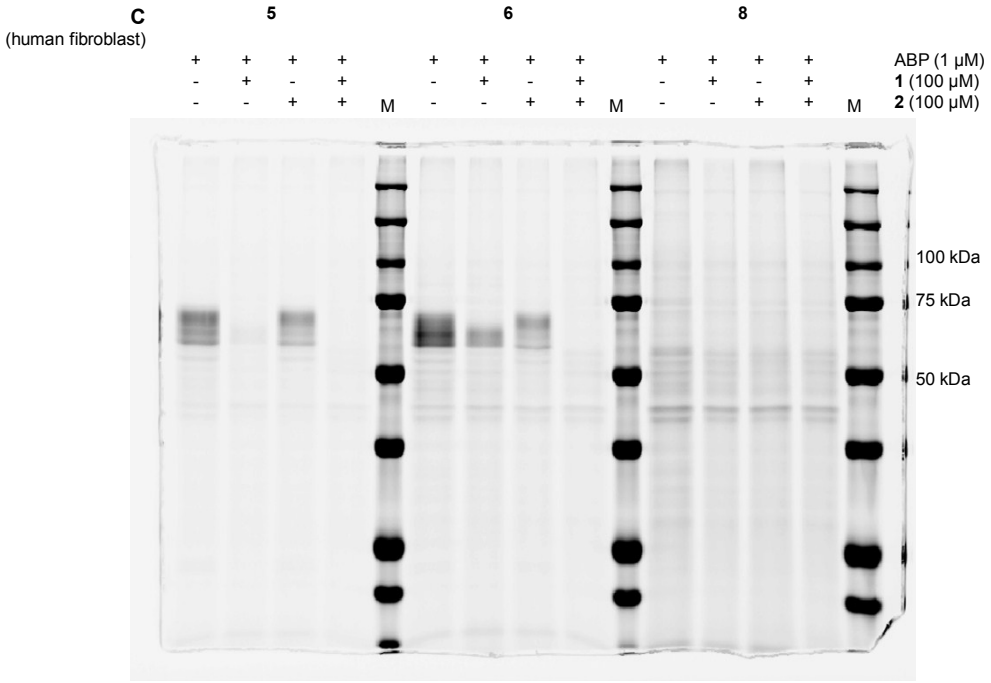
TIME DEPENDENT LABELING OF CjGH35 WITH ABP **6**



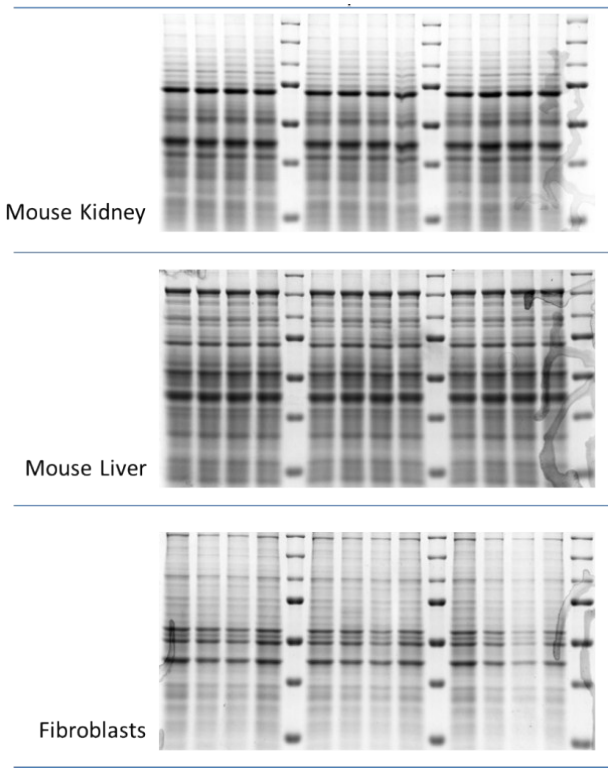
In vitro lysates SDS-PAGE gels







Coomassie stained gels



Kinetics

Due to the high potency of **5** towards recombinant GBA1, kinetics of all probes were determined following the method described by Wu *et al.*[11] Recombinant enzymes GBA1 (Cerezyme, β-glucosidase) and *Cj*GH35A (overexpressed bacterial β-galactosidase) were used for kinetic experiments. For GBA1 (3.6 nM), 150 mM McIlvaine buffer pH 5.2, supplemented with 0.2% Sodium Taurocholate, 0.1% Triton X-100 and 0.1% BSA was used. For *Cj*GH35A (4.8 nM), 150 mM McIlvaine buffer pH 4.5 supplemented with 0.1% BSA was used. Kinetics were measured by adding a series of concentrations of ABP **5**, **6** or **8** (162.5 µL, maximal reaction concentration was 20 µM due to limited availability of the ABPs) to β-4MU-Glc (1300 µL, reaction [S] 2345 µM) and β-4MU-Gal (1300 µL, reaction [S] 837 µM) at 37˚C. The t=0 samples were prepared by taking 112.5 µL from the [ABP + S] to a 96-well plate in duplo, after which stop buffer (200 µL, glycine/MeOH 1M, pH 10) and lastly the enzyme (12.5 µL) was added. Then, enzyme (137.5 µL) was added (t=0) to [ABP + S]. At t=2.5, 5.0, 7.5, 10.0 and 12.5 min, 125 µL of the reaction mixture was taken and added to stop buffer (200 µL, glycine/MeOH 1M, pH 10) in the 96-well plate. After 12.5 minutes, the plate was measured with a PerkinElmer Fluorescence Spectrometer LS-55 using BL Studio with excitation at 366 nm and emission at 445 nm. Kinetic experiments were performed in duplo or triplo with duplo measurements each run. Obtained values were plotted as a one phase exponential association and transferred to k_{obs} values. Which, according to Michaelis-Menten kinetics can be plotted against the concentration to determine the K_I, k_{inact} and their ratio.[11] The K_m of each enzyme was determined by incubating the same enzyme concentration as the kinetic experiments with a series of substrate concentrations for 12.5 minutes at 37 °C. Then, the reaction was stopped by addition to stop buffer and measured on fluorescence. The results were then plotted and K_m was determined according to Michaelis-Menten K_m = ½V_{max}.[12] K_m measurement was performed in a single measurement with a triplo of each sample. GBA1 K_m=1423 µM; *Cj*GH35 K_m=499 µM. Results were processed and analysed using GraphPad Prism 6.0.

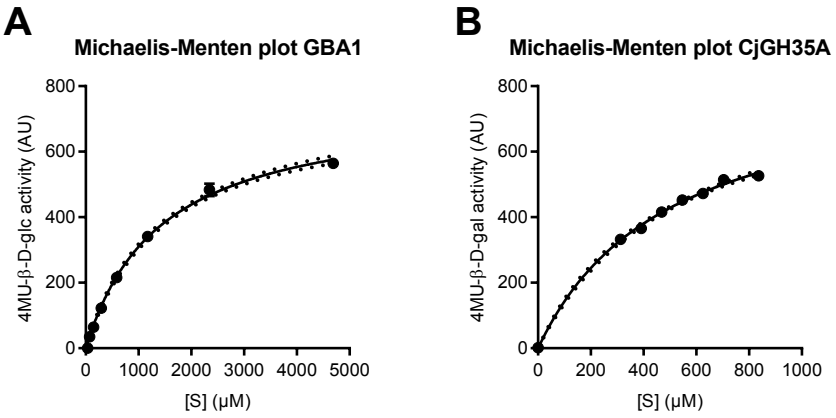


Figure 1 K_m plots for GBA1 and *Cj*GH35.

Table 1 Kinetic parameters of ABPs **5**, **6** and **8** for GBA1 and *Cj*GH35 beta-galactosidase.

	GBA1			***Cj*GH35A**		
ABP	**K_I (μM)**	**k_{inact} (min^{-1})**	**k_{inact}/K_I (μM^{-1} min^{-1})**	**K_I (μM)**	**k_{inact} (min^{-1})**	**k_{inact}/K_I (μM^{-1} min^{-1})**
5	nd[a]	nd[a]	27.51 ± 0.85	nd[b] (20%[c])	nd[b]	nd[b]
6	8.64 ± 1.58	1.18 ± 0.13	0.14 ± 0.08	3.09 ± 0.61	0.49 ± 0.04	0.16 ± 0.07
8	nd[b] (6%[c])	nd[b]	nd[b]	nd[b] (10%[c])	nd[b]	nd[b]

[a] Could not be determined under our assay conditions due to fast inhibition. [b] Could not be determined under our assay conditions due to low potency. [c] Percentage of enzyme inhibition reached at highest concentration of ABP (20 μM).

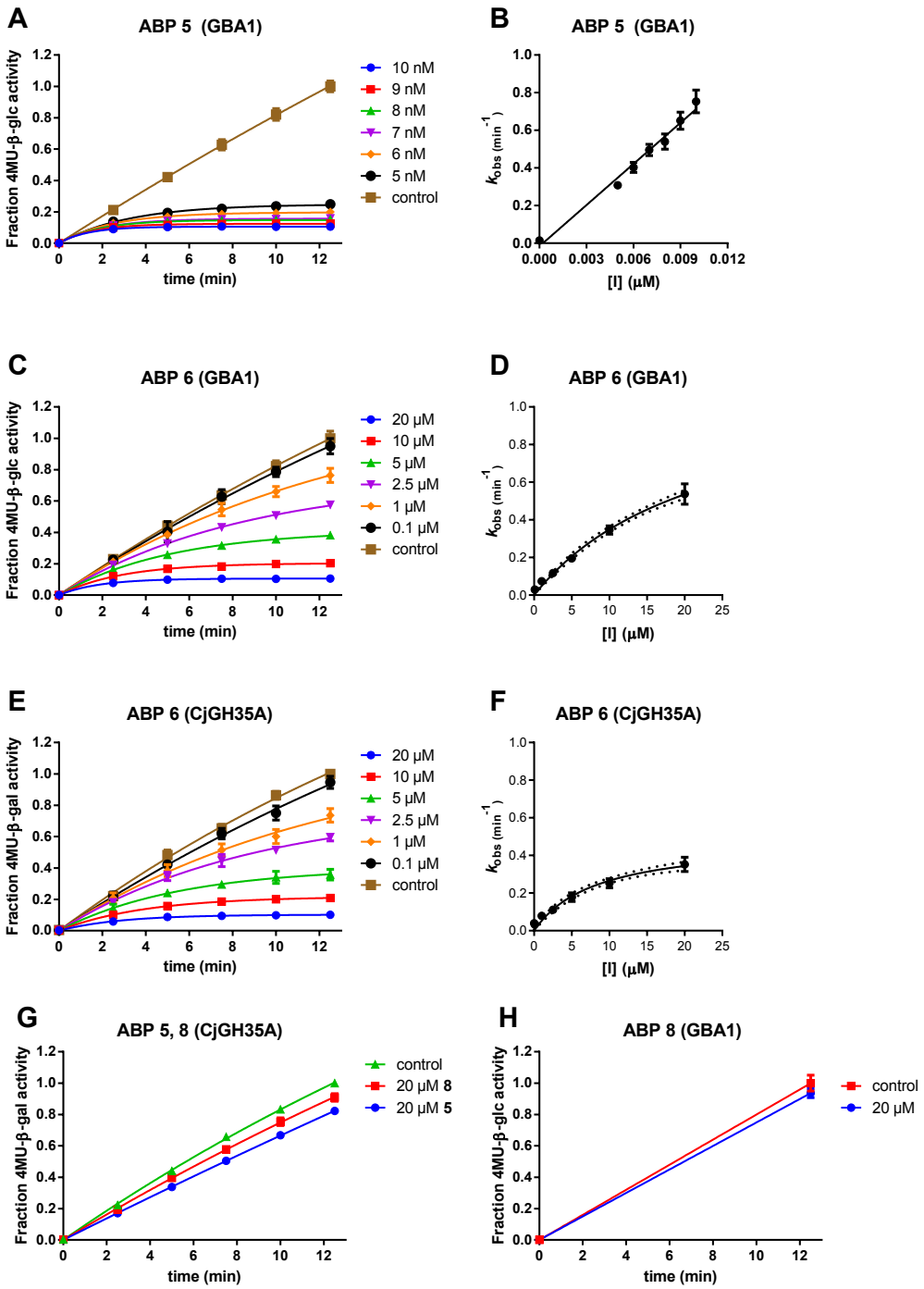


Figure 2 Kinetic plots of GBA1 and *Cj*GH35A inhibition by **5**, **6** and **8** with 4-Methylumbelliferyl ß-D-glucopyranoside and ß-D-galactopyranoside substrates. (A,C,E) Plots of relative substrate activity versus time at a series of different concentrations of **5** or **6** for different enzymes. (B,D,F) Plot of pseudo first-order rate constants from plots A,C and E respectively vs concentration of **5** or **6**. (G) ABPs **5** and **8** do not display significant inhibition of *Cj*GH35 beta-galactosidase at the highest inhibitor concentration in this assay. (H) ABP **8** does not display significant inhibition of GBA1 at the highest inhibitor concentration in this assay.

Pull-down and LC-MS/MS analysis

All pull-down experiments were performed in duplicates. Mouse kidney lysates, mouse liver lysates (1.0 mg total protein) or human fibroblasts lysates (250 µg total protein) were incubated with either 0.25% (v/v) DMSO, ABP **7** or **9** (50 µM) for 3 hours at 37 ˚C, in a total volume of 200 µL McIlvaine buffer pH 5.0, followed by denaturation by addition of 10% (w/v) SDS 50 µL and boiling for 5 minutes at 100 °C. Samples were then reduced with the use of dithiothreitol (DTT), alkylated by iodoacetamide (IAA) and further prepared for pull-down with DynaBead MyOne Straptavidin Beads C1 as published previously.[13] After pull-down, all samples were used for on-bead digestion. The samples were treated with on-bead digestion buffer (100 mM Tris-HCl (pH 7.5), 100 mM NaCl, 1 mM $CaCl_2$, 2% (v/v) acetonitrile and 10 ng/µL trypsin) and incubated overnight in a shaker at 37 ˚C. The supernatant, containing tryptic-digested peptides was then desalted using StageTips. Consequently, the acetonitrile was evaporated using a SpeedVac at 45˚C followed by addition of 20 µL of LC-MS sample solution (95:3:0.1, H_2O:acetonitrile:formic acid) for LC-MS analysis. All peptide samples were analysed with a two hour gradient of 5% to 25% acetonitrile on nano-LC, hyphenated to an LTQ-Orbitrap and identified using the Mascot protein search engine.[14] Raw data was calculated using MaxQuant against Uniprot of human (for fibroblasts) or mouse (for mouse kidney and mouse liver) proteome database to obtain an identification list of found proteins. The abundance of the protein hits was quantified as previously described,[15] in unsupervised mode using the default settings of the PLGS (Waters) and IsoQuant software.

RAW and quantified data is available as separate excel files.

Crystal structure of TB562 in *Cj*GH35

*Cj*GH35A was purified and crystallized as described previously[9]. A crystal was soaked in the presence of a speck of TB562 powder for 70 hours. The crystal was fished directly into liquid nitrogen without the need for additional cryoprotectant. Data were collected on beamline IO2 at the Diamond Light Source at wavelength 0.97950 Å, and were processed using *DIALS*[16] and scaled with *AIMLESS*[17] to 1.6 Å.

In contrast to previous *Cj*GH35A structures, the space group was P1 and the unit cell dimensions, 98.9, 115.8, 116.0 Å, and angles, 90.2, 90.2, 90.4° and was twinned. The structure was solved using programs from the *CCP4* suite[18]*;* molecular replacement was performed the native coordinates, PDB entry 4D1I, as the model and with refinement using twinned intensities. Full details of data quality and refinement statistics are given in the, header information to PDB 5JAW.

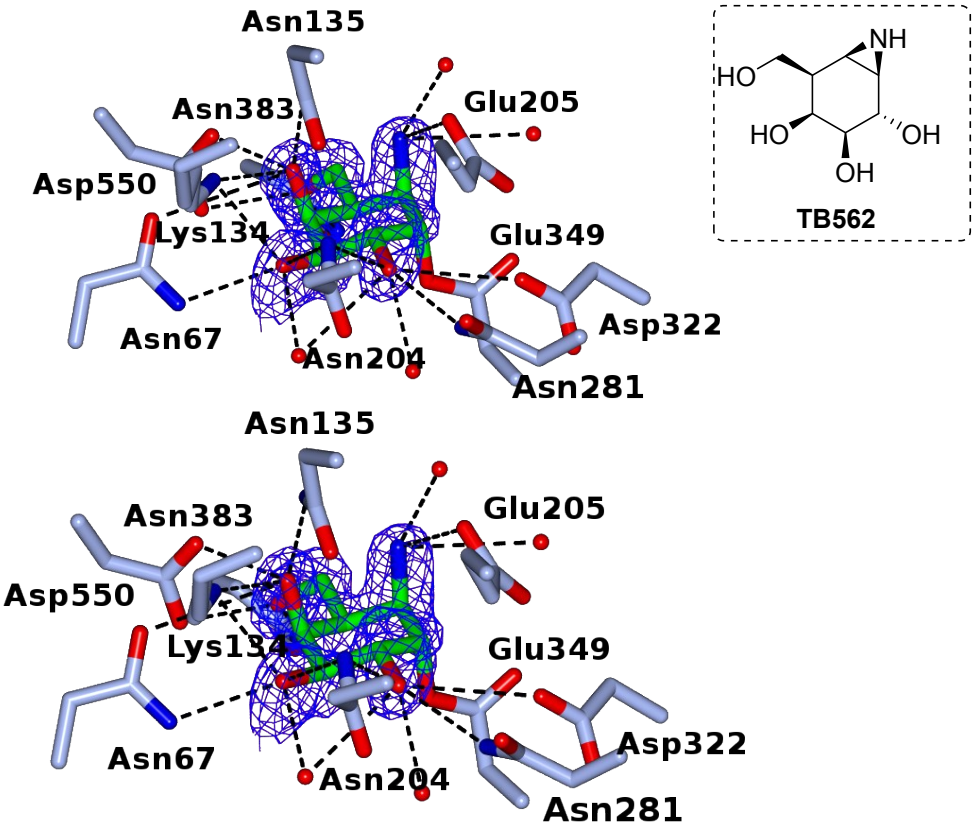


Figure 3 3-D structure of trapped β-galactosidase aziridine complex (PDB code 5JAW). Stereo view (divergent, "wall-eyed") of TB562[7] (β-galacto configured cyclophellitol aziridine) bound to the active centre of the *Cellvibrio japonicus* β-galactosidase *Cj*GH35A. The map shown is a F_o-F_c map, with phases calculated prior to the inclusion of ligand in the refinement, contoured at 3σ. Carbon atoms are coloured green for the ligand and ice blue for the side chains. The interacting residues are annotated, including catalytic residues Glu349 (nucleophile) and Glu205 (acid/base). The ligand TB562,[7] in $^4C_1$ conformation, is covalently attached to the catalytic nucleophile Glu349, as expected. The sugar hydroxyl groups interact extensively with the side chains of active site residues Asn204, Lys134, Asn67 and Asn383 as well as with Asp550 on another molecule in the asymmetric unit. Of note are the extensive interactions at O4, with Lys134, Asp550 and Asn135. The figure was prepared using CCP4MG[19].

References

1 O. Kaczmarek, H. A. Scheidt, A. Bunge, D. Föse, S. Karsten, A. Arbuzova, D. Huster and J. Liebscher, *European J. Org. Chem.*, 2010, 1579–1586.

2 F. M. Cordero, P. Bonanno, M. Chioccioli, P. Gratteri, I. Robina, A. J. Moreno Vargas and A. Brandi, *Tetrahedron*, 2011, **67**, 9555–9564.

3 J. Jiang, T. J. M. Beenakker, W. W. Kallemeijn, G. A. Van Der Marel, H. Van Den Elst, J. D. C. Codée, J. M. F. G. Aerts and H. S. Overkleeft, *Chem. - A Eur. J.*, 2015, **21**, 10861.

4 F. G. Hansen, E. Bundgaard and R. Madsen, *J. Org. Chem.*, 2005, **70**, 10139–10142.

5 A. G. Volbeda, H. A. V. Kistemaker, H. S. Overkleeft, G. A. Van Der Marel, D. V. Filippov and J. D. C. Codée, *J. Org. Chem.*, 2015, **80**, 8796–8806.

6 G. Franck, K. Brödner and G. Helmchen, *Org. Lett.*, 2010, **12**, 3886–9.

7 L. I. Willems, T. J. M. Beenakker, B. Murray, B. Gagestein, H. Van Den Elst, E. R. Van Rijssel, J. D. C. Codée, W. W. Kallemeijn, J. M. F. G. Aerts, G. A. Van Der Marel and H. S. Overkleeft, *European J. Org. Chem.*, 2014, 6044.

8 V. W. F. T. Tony K. M. Shing, *J. Chem. Soc. Perkin Trans. 1*, 1994, 2017–2025.

9 J. Larsbrink, A. J. Thompson, M. Lundqvist, J. G. Gardner, G. J. Davies and H. Brumer, *Mol. Microbiol.*, 2014, **94**, 418.

10 P. K. Smith, R. I. Krohn, G. T. Hermanson, A. K. Mallia, F. H. Gartner, M. D. Provenzano, E. K. Fujimoto, N. M. Goeke, B. J. Olson and D. C. Klenk, *Anal. Biochem.*, 1985, **150**, 76–85.

11 L. Wu, J. Jiang, Y. Jin, W. W. Kallemeijn, C.-L. Kuo, M. Artola, W. Dai, C. van Elk, M. van Eijk, G. A. van der Marel, J. D. C. Codée, B. I. Florea, J. M. F. G. Aerts, H. S. Overkleeft and G. J. Davies, *Nat. Chem. Biol.*, 2017, 1–10.

12 K. A. Johnson and R. S. Goody, *Biochemistry*, 2011, **50**, 8264–8269.

13 N. Li, C.-L. Kuo, G. Paniagua, H. van den Elst, M. Verdoes, L. I. Willems, W. A. van der Linden, M. Ruben, E. van Genderen, J. Gubbens, G. P. van Wezel, H. S. Overkleeft and B. I. Florea, *Nat. Protoc.*, 2013, **8**, 1155–1168.

14 Y. Ishihama, *Mol. Cell. Proteomics*, 2005, **4**, 1265–1272.

15 J. Kuharev, P. Navarro, U. Distler, O. Jahn and S. Tenzer, *Proteomics*, 2015, **15**, 3140–3151.

16 D. G. Waterman, G. Winter, R. J. Gildea, J. M. Parkhurst, A. S. Brewster, N. K. Sauter and G. Evans, *Acta Crystallogr. Sect. D, Struct. Biol.*, 2016, **72**, 558–575.

17 P. R. Evans and G. N. Murshudov, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2013, **69**, 1204–1214.

18 M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B.

Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin and K. S. Wilson, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2011, **67**, 235–242.

19 S. McNicholas, E. Potterton, K. S. Wilson and M. E. M. Noble, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2011, **67**, 386–394.