Electronic Supplementary Material (ESI) for Chemical Communications

Supporting Information

Unexpected Structure of a *C. difficile* Toxin A Ligand Necessitates an Annotation Correction in a Popular Screening Library

Ping Zhang,[a,b] *Nahid Razi,*[c] *Luiz Eugenio,*[b] *Messele Fentabil,*[d] *Elena Kitova,*[d] *John Klassen,*[d] *David R. Bundle,*[d] *Kenneth K.-S. Ng*[b]* *and Chang-Chun Ling*[a]*

[a] Department of Chemistry, Alberta Ingenuity Centre for Carbohydrate Science, University of Calgary, Calgary AB T2N1N4 Canada. Fax: 403-289-9488; Tel: 403-220-2768; E-mail: ccling@ucalgary.ca

[b] Department of Biological Sciences, Alberta Ingenuity Centre for Carbohydrate Science, University of Calgary, Calgary AB T2N1N4 Canada. Fax: 403-289-9311; Tel: 403-220-4320; E-mail: ngk@ucalgary.ca

[c] Department of Chemical Physiology, The Scripps Research Institute, La Jolla, CA 92037, USA.

[d] Department of Chemistry, Alberta Ingenuity Centre for Carbohydrate Science, University of Alberta, Edmonton AB T6G2G2 Canada.

Table of Content

General methods

Optical rotations were determined in a 5 cm cell at 25 ± 2°C. $[a]_D^{25}$ values are given in units of 10^{-1} deg cm^2 g^{-1}. Analytical TLC was performed on Silica Gel 60-F_{254} (Merck, Darmstadt) with detection by quenching of fluorescence and/or by charring with 5% sulfuric acid in water or with a ceric ammonium molybdate dip. All commercial reagents were used as supplied unless otherwise stated. Column chromatography was performed on Silica Gel 60 (Silicycle, Ontario). Size exclusion chromatography was performed using Sephadex G-15 resin. Molecular sieves were stored in an oven at 100° C and flame-dried under vacuum before use. Organic solutions from extractions were dried with anhydrous Na_2SO_4 prior to concentration under vacuum at < 40 °C (bath). ^{1}H NMR spectra were recorded at 400 MHz on Bruker spectrometers. The first order proton chemical shifts δ_H and δ_C are reported in δ (ppm) and referenced to either residual $CHCl_3$ (δ_H 7.24, δ_C 77.0, $CDCl_3$) or residual CD_2HOD (δ_H 3.30, δ_C 49.5, CD_3OD). ^{1}H and ^{13}C NMR spectra were assigned with the assistance of GCOSY, GHSQC spectra. Microanalyses and ESI-MS were performed by the analytical services of the Department of Chemistry, University of Calgary. For high resolution mass determination, spectra were obtained by voltage scan over a narrow range at a resolution of approximately 10000 and recorded by the analytical services of the Department of Chemistry, University of Alberta.

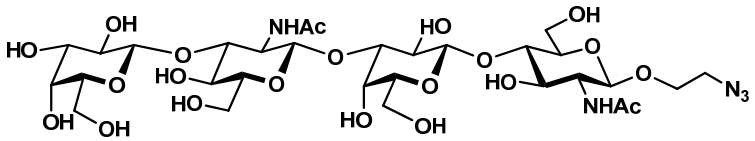
A. Chemical and Enzymatic Synthesis

Experimental details for the expression of the recombinant human FUT III in Sf9 insect cells using gene encoded in the baculoviral construct.

The baculoviral construct encoded the human α1-3/4 fucosyltransferase III (FUT-III) gene was kindly contributed by NEOSE, Technologies Inc. (currently Novo Nordisk, Princeton, NJ, USA). The viruses were purified by conventional plaque-forming assay to determine the viral titers. Viral amplification was performed in Sf9 cells by infecting 1 × 10^6 cells with multiplicity of infection, MOI = 0.5–1. The viral production was initiated

in T-flasks (20 mL) and further adapted to 500 mL shaker flask (150 mL culture), incubating at 27 °C for 48 h. The enzyme expression was also performed in Sf9 cells in 350 mL culture in 1 L shaker flasks with serum-free media in the presence of 10% conditioned media to maintain the optimal cell growth and protein expression. Sf9 cells were infected ($1.5 – 2 \times 10^6$ cells/mL) with the desired purified viral stock with the MOI = 5. Incubation continued with vigorous shaking (120 rpm) at 27 °C. The enzyme expression was monitored for up to 7 days, testing by general fucosyltransferase assay with ^{14}C-GDP-fucose, according to literature procedure.[1] The culture medium was collected by centrifugation for 30 min at 5000 rpm when the cell viability decreased to about 50%, and the assay showed a decrease in the activity. Collected media was concentrated up to 20 fold by tangential flow filtration and used in synthesis.

B. Chemical and Enzymatic Synthesis

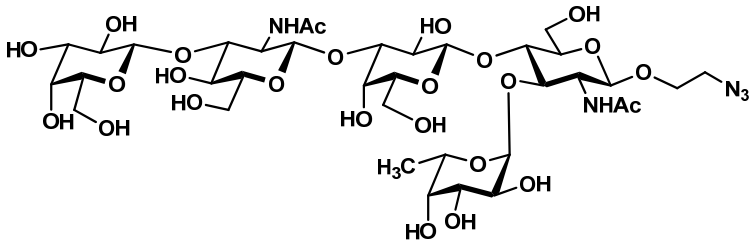


5

2-Azidoethyl β-D-galactopyranosyl-(1→3)-2-acetamido-2-deoxy-β-D-glucopyranosyl-(1→3)-β-D-galactopyranosyl-(1→4)-2-acetamido-2-deoxy-β-D-glucopyranoside (**5**): 2-Azidoethyl 2-acetamido-2-deoxy-β-D-glucopyranosyl-(1→3)-β-D-galactopyranosyl-(1→4)-2-acetamido-2-deoxy-β-D-glucopyranoside (**3**) (1.0 g, 1.52 mmol) and UDP-glucose (1.02 g, 1.67 mmol) were dissolved in a solution containing cacodylate buffer (25 mL, 100 mM, pH 7.5) and $MnCl_2$ (100 mM), and a solution of β(1,3)-galactosyltransferase V (GalT V, 20 mL, 0.4 U/mL) was added followed by another solution containing the UDP-Gal 4-epimerase (GalE, 15 mL, 42.5 U/mL). The reaction was slowly stirred at room temperature for 48 hours. The reaction mixture was centrifuged and went through gel filtration column Sephadex G15 twice. Collected fractions were freeze dried to give tetrasaccharide **5** (1.2 g, 1.46 mmol).

^{1}H NMR (400 MHz, D_2O) δ 4.64 (d, overlapped, 1H, H-1–GlcN-I)), 4.52 (d, 1H, *J* 8.1 Hz, H-1_GlcN_II), 4.38 (d, 1H, *J* 8.1 Hz, H-1_Gal_I), 4.35 (d, 1H, *J* 8.1 Hz, H-1_Gal_II), 4.07 (dd, 1H, *J* 3.0, 1.0 Hz, H-4_Gal_I), 3.96 (m, 1H, OC*H*$_a$H$_b$CH$_c$H$_d$N$_3$), 3.91 (dd, 1H, *J* 12.5, 1.9 Hz, H-6a_Glc_II), 3.85–3.37 (m, 24H, H-2_GlcN_I + H-3_GlcN_I + H-4_GlcN_I + H-5_GlcN_I + H-6a_GlcN_I + H-6b_GlcN_I + H-2_GlcN_II + H-3_GlcN_II + H-4_GlcN_II + H-5_GlcN_II + H-6b_GlcN_II + H-2_Gal_I + H-3_Gal_I + H-5_Gal_I + H-6a_Gal_I + H-6b_Gal_I + H-2_Gal_II + H-3_Gal_II + H-4_Gal-II + H-5_Gal_II + H-6a_Gal_II + H-6b_Gal_II + OCH$_a$*H*$_b$CH$_c$H$_d$N$_3$ + OCH$_a$H$_b$C*H*$_c$H$_d$N$_3$), 3.33 (m, 1H, OCH$_a$H$_b$CH$_c$*H*$_d$N$_3$), 1.95 (s, 3H, AcNH), 1.94 (s, 3H, AcNH).

^{13}C NMR (400 MHz, D_2O, from HSQC) δ_H 103.23 (C-1_Gal_II), 102.68 (C-1_Gal_I), 102.31 (C-1_GlcN_II), 100.70 (C-1_GlcN_I), 81.95, 81.74, 78.25, 74.89, 74.82 (× 2), 74.54, 72.17 (× 2), 70.47, 69.72, 68.58 (O*C*H$_2$CH$_2$N$_3$), 68.42, 68.28, 67.92, 60.58 (× 2), 60.07, 59.75, 54.66 (C-2_GlcN_I), 54.35 (C-2_GlcN_II), 50.06 (OCH$_2$*C*H$_2$N$_3$), 22.03 (Ac), 21.92 (Ac).



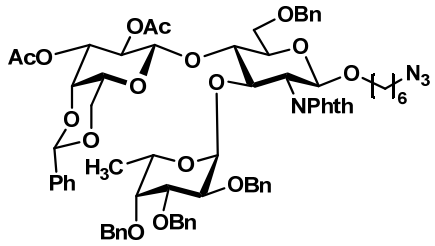
7 (S259-1)

2-Azidoethyl β-D-galactopyranosyl-(1→3)-2-acetamido-2-deoxy-β-D-glucopyranosyl-(1→3)-β-D-galactopyranosyl-(1→4)-[α-L-fucopyranosyl-(1→3)]-2-acetamido-2-deoxy-β-D-glucopyranoside (**7**): Tetrasaccharide **5** (400 mg, 48.9 umol) and GDP-Fucose (929 mg, 40%; 58.7 umol) were dissolved in cacodylate buffer (5 mL, pH 7.5) containing $MnCl_2$ (30 mM). A solution of α(1,3)-fucosyltransferase (FUT III, 8 mL, 0.35 U/mL) was added and the reaction was slowly stirred at room temperature for 48 hrs. TLC and MS showed monofucosylated product. The mixture was centrifuged. The solution (~50 mL) was loaded onto a column of Sephadex G-15 (5 × 140 cm) pre-equilibrated with 5% n-BuOH-H_2O, and eluted with the same eluent. Fractions

containing S259-1 (**7**) were combined and evaporated (~320 mg, 80% purity). 100 mg of this were re-purified by gel-filtration chromatography on Sephadex G-15 as above. Appropriate fractions were collected to afford pentasaccharide S259-1 (26 mg, > 90% purity). The less pure portion was collected and re-purified as above to afford more pure S259-1 (58 mg, 95% purity). Total amount of S259-1 obtained was 84 mg.

^{1}H NMR (400 MHz, D_2O) δ_H 5.01 (1H, d, J = 4.0 Hz, H-1_Fuc), 4.72 (1H, overlapped with HDO, H-5_Fuc), 4.64 (1H, d, J = 8.4 Hz, GlcN_I), 4.53 (1H, d, J = 8.1 Hz, GlcN_II), 4.36 (1H, d, J = 7.7 Hz, H-1_Gal_II), 4.35 (1H, d, J = 7.8 Hz, H-1_Gal_I), 4.01 (1H, dd, J = 3.4, ~1 Hz, H-4_Gal_II), 3.96 (1H, ddd, J = 3.1, 5.4, 11.3 Hz, OC*Ha*Hb), 3.92 (1H, dd, J = 2.0, 12.3 Hz, H-6a_GlcN_II), 3.88 - 3.58 (19H, m, H-2_GlcN_I + H-2_GlcN_II + H-4_Gal_I + H-3_Fuc + OC*Ha*Hb + H-3_GlcN_II + H-4_GlcN_II + H-3_GlcN_I + H-3_Fuc + H-3_Gal_II + H-5_Gal_I + H-2_Fuc + H-6b_GlcN_II + H-6a_Gal_I + H-6b_Gal_I + H-6a_GlcN_I + H-6b_GlcN_I + H-6a_Gal_II + H-6b_Gal_II), 3.56 (1H, dd, 1H, J = 3.5, 10.0 Hz, H-3_Gal_I), 3.54 – 3.30 (9H, m, H-4_GlcN_I + OCHa*Hb* + H-5_Gal_II + H-5_GlcN_II + H-2_Gal_I + H-2_Gal_II + H-5_Gal_I + C*Ha*HbN3 + + CHa*Hb*N_3), 1.95 (3H, s, NHAc), 1.94 (3H, s, NHAc), 1.07 (3H, d, J = 6.8 Hz, H-6_Fuc).

^{13}C NMR (100 MHz, D_2O, from HSQC) δ 103.36 (C-1_Gal_II), 102.51 (C-1_GlcN_I), 101.80 (C-1_Gal_I), 100.64 (C-1_GlcN_II), 98.64 (C-1_Fuc), 81.93, 81.49, 75.11, 75.03, 74.77 (× 3), 73.01, 72.23, 71.71, 70.41 (× 2), 69.00, 68.65, 68.30, 68.04, 67.35, 66.48, 61.14, 60.71, 60.10, 59.66, 55.04 (C-2_GlcN_II), 54.69 (C-2_GlcN_I), 50.21 (CH_2N_3), 21.90 (NHAc × 2), 14.80 (C-6_Fuc).



10

6-Azidohexyl 2,3-di-*O*-acetyl-4,6-O-benzylidene-β-D-galactopyranosyl-(1→3)-[2,3,4-tri-*O*-benzyl-α-L-fucopyranosyl–(1→3)]-6-*O*-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranoside (10). Compound **8** (200 mg, 0.233 mmol) and thiofucoside **9** (395 mg, 0.704 mmol) were dissolved in anhydrous CH_2Cl_2 (3.5 mL) under Ar, and 4 Å molecular sieves (700 mg) were added. After stirring at rt for 1 h, the mixture was cooled to -78° C, and NIS (139 mg, 0.583 mmol) was added. After another 5 mins, a saturated solution of TfOH in anhydrous CH_2Cl_2 (20 μL) was added and the temperature was slowly warmed to -10° C. The mixture was neutralized with Et_3N (1.0 mL). The insoluble material was filtered off and washed with more EtOAc. The combined organic solution was washed with an 1 : 1 mixture of 10% aqueous $Na_2S_2O_3$ and 5% $NaHCO_3$ (2 × 40 mL), dried over anhydrous Na_2SO_4 and evaporated. The residue was purified by column chromatography on silica gel using 10% ethyl acetate – toluene as eluent to afford the desired glycoside **10** (240 mg, yield: 81%).

R_f: 0.69 (40% ethyl acetate – toluene).

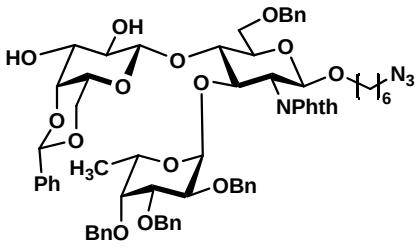
$[\alpha]_D$ -28.5° (*c* 0.8, $CHCl_3$).

^{1}H NMR (400 MHz, $CDCl_3$) δ_H 7.85 - 7.60 (4H, m, NPhth), 7.55 - 7.50 (2H, m, Ar), 7.46 – 7.40 (4H, m, Ar), 7.40 – 7.07 (14H, m, Ar), 7.05 – 6.96 (4H, m, Ar), 6.93 – 6.88 (2H, m, Ar), 5.54 (1H, s, PhC*H*), 5.35 (1H, dd, *J* = 10.3, 8.3 Hz, H-2_Gal), 5.03 (1H, d, *J* = 8.6 Hz, H-1_GlcN), 4.88 (1H, d, *J* = 11.9 Hz, Bn), 4.83 – 4.64 (5H, m, H-1_Fuc + H-5_Fuc + H-1_Gal + H-3_Gal + H-3_GlcN), 4.55 (2H, s, Bn), 4.50 (1H, d, *J* = 12.1 Hz, Bn), 4.46 – 4.35 (3H, m, H-2_Glc + H-6a_Gal + Bn), 4.31 – 4.24 (2H, m, Bn + H-4_Gal), 4.21 (1H, dd, *J* = 9.3, 9.3 Hz, H-4_GlcN), 4.05 (1H, d, *J* = 11.2 Hz, Bn), 4.02 – 3.95 (2H, m, H-6b_Gal + H-6a_GlcN), 3.92 – 3.77 (3H, m, H-3_Fuc + H-6b_Glc + OC*Ha*Hb), 3.61 (1H, dd, *J* = 10.4, 3.8 Hz, H-2_Fuc), 3.53 (1H, m, H-5_GlcN), 3.43 (1H, d, *J* = 11.2 Hz, Bn), 3.32 (1H, ddd, *J* = 9.7, 7.7, 5.5 Hz, OCHa*Hb*), 3.14 (1H, dd, *J* = 1.8, ~1 Hz, H-4_Fuc), 3.06 (1H, m, H-5_Gal), 3.00 (2H, t, *J* = 7.0 Hz, CH_2N_3), 2.13 (3H, s, Ac), 2.10 (3H, s, Ac), 1.49 – 1.00 (11H, m, 4 × CH_2 + H-6_Fuc).

^{13}C NMR (100 MHz, $CDCl_3$) δ 170.63 (CO), 168.71 (CO), 139.52, 139.45, 138.27, 138.12, 137.66, 134.14, 128.87, 128.56, 128.15, 128.03, 128.00, 127.90, 127.88, 127.77, 127.72, 127.57, 127.33, 126.93, 126.73, 125.81, 99.67 (Ph*C*H), 99.61 (C-1_Gal), 98.51 (C-1_GlcN), 97.56 (C-1_Fuc), 78.99, 78.87, 75.52, 75.01, 74.75, 73.79, 73.45, 73.28,

72.80, 72.18, 71.62, 71.36, 69.32, 68.94, 68.75, 67.82, 66.35, 66.27, 56.54 (C-2_GlcN), 51.12 (CH_2N_3), 29.07 (CH_2_chain), 28.54 (CH_2_chain), 26.17 (CH_2_chain), 25.41 (CH_2_chain), 20.87 (Ac), 20.82 (Ac), 16.06 (C-6_Fuc).

ESI high resolution MS: $[C_{71}H_{78}N_4O_{18} + Na]^+$ calcd 1297.5209, found 1297.5202.



11

6-Azidohexyl 4,6-O-benzylidene-β-D-galactopyranosyl-(1→3)-[2,3,4-tri-*O*-benzyl-α-L-fucopyranosyl–(1→3)]-6-*O*-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranoside (11). Trisaccharide **10** (119 mg, 0.093 mmol) was dissolved in a guanidine/guanidinium chloride solution (3.0 mL), and the reaction was stirred at rt for 30 mins. CH_3CO_2H (1 mL) was added to quench the reaction, and the mixture was concentrated under reduced pressure. The residue was purified by column chromatography on silica gel using 30% ethyl acetate – toluene as the eluent to give the desired disaccharide **11** (107 mg, 96% yield) as a white solid.

R_f: 0.26 (40% ethyl acetate – toluene).

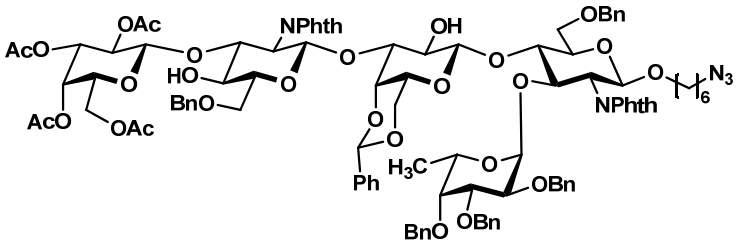
$[\alpha]_D$ -44.7° (*c* 0.3, $CHCl_3$).

^{1}H NMR (400 MHz, $CDCl_3$) δ_H 7.82 – 7.63 (4H, m, Phth), 7.52 – 7.48 (2H, m, Ar), 7.43 – 7.10 (18H, m, Ar), 7.08 - 7.00 (4H, m, Ar), 6.96 – 6.91 (2H, m, Ar), 5.57 (1H, s, PhC*H*), 5.04 (1H, d, *J* = 8.1 Hz, H-1_GlcN), 4.81 (1H, d, *J* = 12.3 Hz, Bn), 4.79 – 4.67 (3H, m, H-3_GlcN + H-1_Fuc + H-5_Fuc), 4.59 – 4.49 (4H, m, H-1_Gal + 3 × Bn), 4.46 – 4.37 (2H, m, Bn + H-2_GlcN), 4.32 (1H, dd, *J* = 12.4, ~1 Hz, H-6a_Gal), 4.28 – 4.18 (3H, m, 2 × Bn + H-4_GlcN), 4.12 (1H, dd, *J* = 11.2, 2.9 Hz, H-6a_Glc), 4.07 (1H, dd, *J* = 3.7, ~1 Hz, H-4_Gal), 3.98 (1H, dd, *J* = 12.4, 1.6 Hz, H-6b_Gal), 3.90 (2H, m, H-3_Fuc + H-6b_GlcN), 3.83 (1H, ddd, *J* = 9.9, 5.9, 5.9 Hz, OC*Ha*Hb), 3.66 (4H, m, H-5_GlcN + Bn + H-2_Gal + H-2_Fuc), 3.48 (1H, ddd, *J* = 3.3, 8.1, 8.1 Hz, H-3_Gal), 3.35

(1H, ddd, *J* = 9.9, 7.5, 5.6 Hz, OCHa*Hb*), 3.28 (1H, dd, *J* = 1.6, ~1 Hz, H-4_Fuc), 3.07 (1H, m, H-5_Gal), 3.01 (3H, m, OH-2_Gal + CH_2N_3), 2.58 (1H, d, *J* = 7.9 Hz, OH-3_Gal), 1.49 – 0.95 (11H, m, 4 × CH_2 + H-6_Fuc).

^{13}C NMR (100 MHz, $CDCl_3$) δ 139.27, 139.25, 138.31, 138.23, 137.66, 134.09, 128.97, 128.40, 128.17, 128.10, 127.98, 127.86, 127.83, 127.67, 127.63, 127.61, 127.35, 127.05, 126.91, 125.82, 101.25 (C-1_Gal), 100.13 (Ph*C*H), 98.52 (C-1_GlcN), 97.95 (C-1_Fuc), 79.01, 78.54, 75.61, 75.25, 75.17, 74.76, 73.95, 73.16, 72.90, 71.97, 71.60, 69.33, 69.23, 68.50, 66.54, 66.40, 56.53 (C-2_GlcN), 51.12 (CH_2N_3), 29.09 (CH_2), 28.55 (CH_2), 26.18 (CH_2), 25.42 (CH_2), 16.44 (C-6_Fuc).

ESI high resolution MS: $[C_{67}H_{74}N_4O_{16} + Na]^+$ calcd 1213.4992, found 1213.4991.



13

6-Azidohexyl 2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl-(1→3)-6-*O*-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl-(1→3)-4,6-*O*-benzylidene-β-D-galactopyranosyl-(1→4)-[2,3,4-tri-*O*-benzyl-α-L-fucopyranosyl–(1→3)]-6-*O*-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranoside (13). Compound **11** (50 mg, 0.042 mmol) and thiodisaccharide **12** (50 mg, 0.063 mmol) were dissolved in anhydrous CH_2Cl_2 (1.0 mL) under Ar, and 4 Å molecular sieves (200 mg) were added. After stirring at rt for 1 h, the mixture was cooled to -78° C, and NIS (24 mg, 0.105 mmol) was added. After another 5 mins, a saturated solution of TfOH in anhydrous CH_2Cl_2 (5 μL) was added and the temperature was slowly warmed to -10° C. The mixture was neutralized with Et_3N (0.5 mL). The insoluble material was filtered off and washed with more EtOAc. The combined organic solution was washed with an 1 : 1 mixture of 10% aqueous $Na_2S_2O_3$ and 5% $NaHCO_3$ (2 × 15 mL), dried over anhydrous Na_2SO_4 and evaporated. The

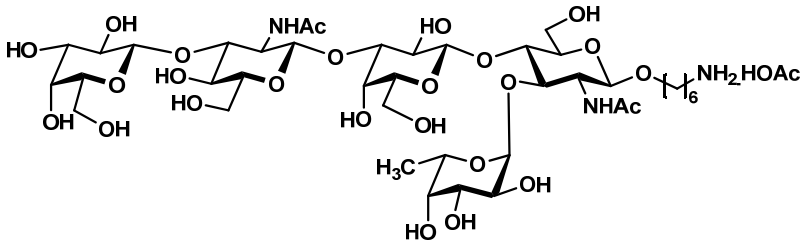
residue was purified by column chromatography on silica gel using 35% ethyl acetate – toluene as eluent to afford the desired glycoside **13** (62 mg, yield: 77%).

R_f: 0.42 (50% ethyl acetate – toluene).

$[\alpha]_D$ +40° (*c* 0.7, $CHCl_3$);

^{1}H NMR (400 MHz, $CDCl_3$) δ_H 7.94 – 7.56 (8H, m, 2 × Phth), 7.51 – 7.06 (31H, m, Ar), 6.98 (4H, m, Ar), 6.83 (2H, m, Ar), 5.44 (1H, d, *J* = 8.4 Hz, H-1_GlcN_I), 5.41 (1H, s, PhCH), 5.32 (1H, dd, *J* = 4.2, ~1 Hz, H-4_Gal_I), 5.17 (1H, dd, *J* = 10.5, 8.0 Hz, H-2_Gal_I), 4.96 (1H, d, *J* = 8.5 Hz, H-1_GlcN_II), 4.87 (1H, dd, *J* = 10.5, 3.5 Hz, H-3_Gal_I), 4.76 (1H, d, *J* = 12.1 Hz, Bn), 4.68 – 4.61 (3H, m, 2 × Bn + H-3_GlcN_II), 4.58 – 4.34 (10H, m, H-1_Gal_I + H-1_Fuc + H-5_Fuc + H-2_GlcN_I + H-2_GlcN_II + 4 × Bn + H-4_GlcN_I), 4.27 (1H, d, *J* = 7.8 Hz, H-1_Gal_II), 4.25 – 3.25 (29H, m), 2.99 (2H, t, *J* = 7.0 Hz, CH_2N_3), 2.96 (1H, dd, *J* = 2.5, ~1 Hz, H-4_Gal_I), 2.86 (1H, m, H-5_Gal_II), 2.75 (1H, s, OH-2_Gal_II), 2.14 (3H, s, Ac), 2.06 (3H, s, Ac), 1.90 (3H, s, Ac), 1.44 (3H, s, Ac), 1.35 – 1.09 (11H, m, 4 × CH_2 + H-6_Fuc).

ESI high resolution MS: $[C_{102}H_{111}N_5O_{31} + Na]^+$ calcd 1924.7155, found 1924.7150.



14

6-Aminohexyl β-D-galactopyranosyl-(1→3)-2-acetamido-2-deoxy-β-D-glucopyranosyl-(1→3)-β-D-galactopyranosyl-(1→4)-[α-L-fucopyranosyl–(1→3)]-2-acetamido-2-deoxy-β-D-glucopyranoside, acetic acid salt (14). A solution of pentasaccharide **13** (52 mg, 0.027 mmol) was dissolved in EtOH (4.0 mL); a solution of hydrazine hydrate (50%, 0.5 mL) was added and the mixture was heated to reflux for 48 h. The solution was evaporated to dryness under vacuum. The residue was acetylated using a mixture of 1 : 1 acetic anhydride – pyridine (1.5 mL). After evaporation, the residue was purified by chromatography on silica gel using 2% MeOH – CH_2Cl_2 as

eluent to afford the intermediate (~31.4 mg), which was redissolved in anhydrous MeOH (3.0 mL) and a solution of NaOMe in MeOH (1.5 M, 20 μL) was added; the reaction was continued for 2 hrs at room temperature. The solution was neutralized with acetic acid, and the evaporated to dryness. The residue was purified by column chromatography on silica gel using 20% MeOH-CH_2Cl_2 to afford an intermediate polyol (25 mg) as a colorless solid. The solid was dissolved in 10% AcOH-MeOH (4.0 mL), and hydrogenated with 20% $Pd(OH)_2$ on charcoal (20 mg) for 3 days at rt. The catalyst was removed by filtration and the solution was concentrated. The residue was dissolved in deionized H_2O, purified on a C18 Sep-Pak cartridge using a gradient of MeOH - H_2O (0% → 5%) as eluent to give the desired pentasaccharide **14** as a white solid after a lyophilization (16 mg, 56% yield for 4 steps).

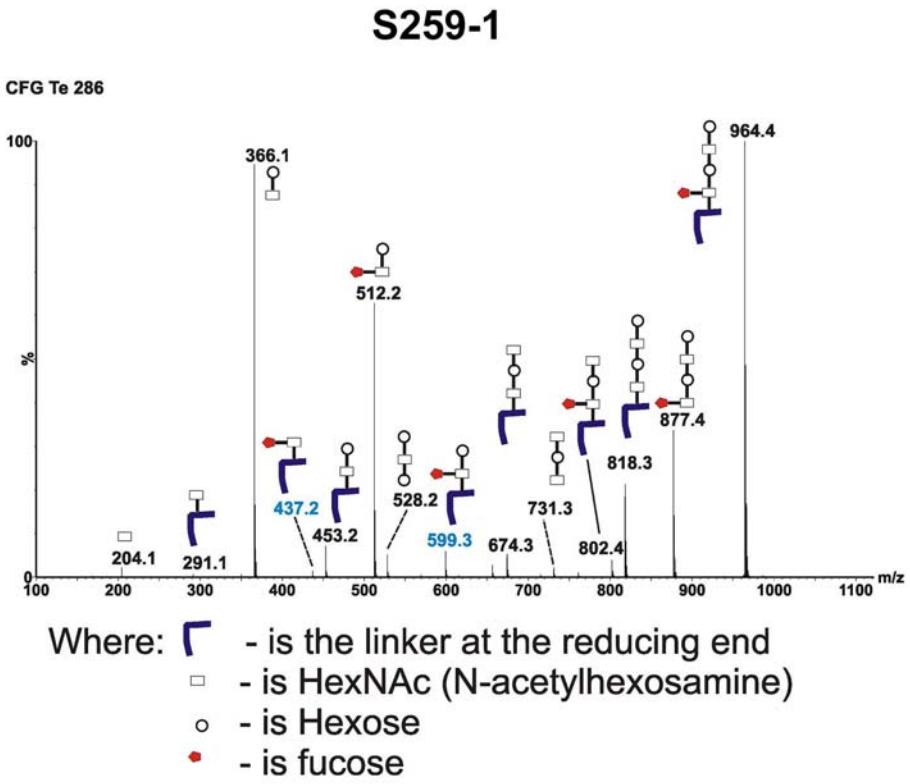
$[\alpha]_D$ -43° (*c* 0.3, MeOH);

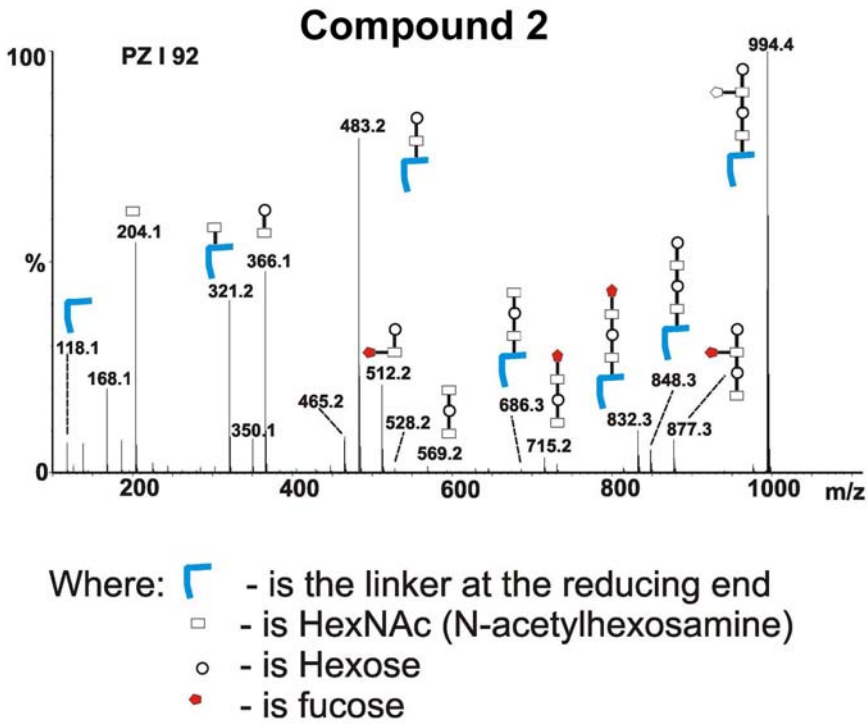
^{1}H NMR (400 MHz, D_2O) δ_H 5.02 (1H, d, *J* = 4.0 Hz, H-1_Fuc), 4.74 (1H, overlapped with HDO, H-5_Fuc), 4.65 (1H, d, *J* = 8.4 Hz, GlcN_I), 4.45 (1H, d, *J* = 8.1 Hz, GlcN_II), 4.37 (1H, d, *J* = 7.7 Hz, H-1_Gal_II), 4.36 (1H, d, *J* = 7.6 Hz, H-1_Gal_I), 4.02 (1H, dd, *J* = 3.3, ~1 Hz, H-4_Gal_II), 3.91 (1H, dd, *J* = 2.1, 12.4 Hz, H-6a_GlcN_II), 3.87 - 3.59 (19H, m, H-2_GlcN_I + H-2_GlcN_II + H-4_Gal_I + H-3_Fuc + OC*Ha*Hb + H-3_GlcN_II + H-4_GlcN_II + H-3_GlcN_I + H-3_Fuc + H-3_Gal_II + H-5_Gal_I + H-2_Fuc + H-6b_GlcN_II + H-6a_Gal_I + H-6b_Gal_I + H-6a_GlcN_I + H-6b_GlcN_I + H-6a_Gal_II + H-6b_Gal_II), 3.57 (1H, dd, 1H, *J* = 3.5, 9.8 Hz, H-3_Gal_I), 3.54 – 3.37 (7H, m, H-4_GlcN_I + OCHa*Hb* + H-5_Gal_II + H-5_GlcN_II + H-2_Gal_I + H-2_Gal_II + H-5_Gal_I), 2.91 (2H, t, *J* = 7.6 Hz, C$H_2$$NH_2$), 1.95 (6H, s, 2 × NHAc), 1.87 (3H, s, HOAc), 1.58 (2H, pentet, *J* = 7.6 Hz, CH_2), 1.48 (2H, pentet, *J* = 7.6 Hz, CH_2), 1.36 – 1.22 (4H, m, 2 × CH_2), 1.08 (3H, d, *J* = 6.8 Hz, H-6_Fuc).

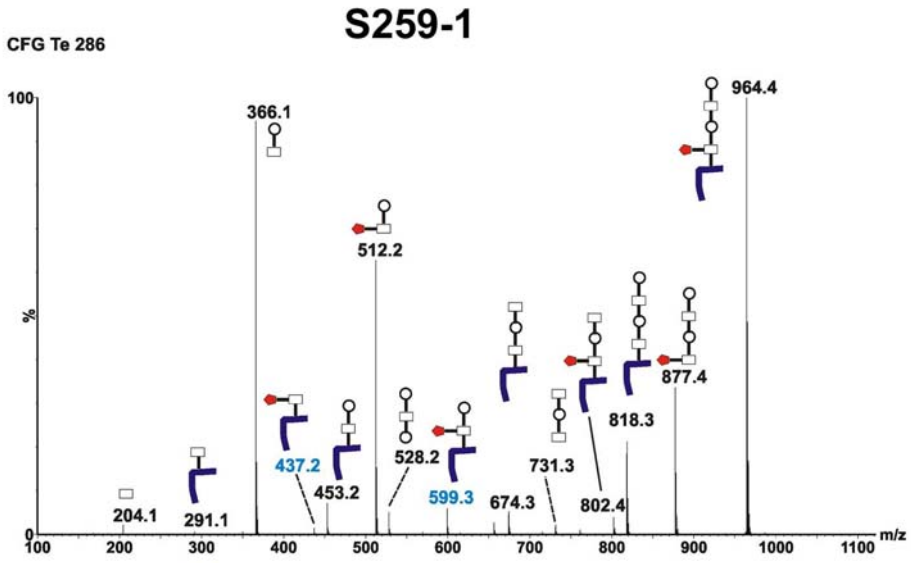
^{13}C NMR (100 MHz, D_2O) δ 174.91 (CO), 174.13 (CO), 103.44 (C-1_Gal_II), 102.46 (C-1_GlcN_I), 101.75 (C-1_Gal_I), 100.96 (C-1_GlcN_II), 98.70 (C-1_Fuc), 82.01, 81.57, 75.32, 75.26, 75.14, 74.93, 74.44, 73.10, 72.45, 71.84, 70.66, 70.52, 70.46, 69.18, 68.51, 68.42, 68.20, 67.64, 66.69, 61.42, 61.01, 60.47, 59.77, 55.85 (C-2_GlcN_II), 54.67 (C-2_GlcN_I), 39.39 (CH_2NH_2), 28.36 (CH_2), 26.63 (CH_2), 25.22 (CH_2), 24.62 (CH_2), 22.23 (NHAc), 22.20 (NHAc), 15.24 (C-6_Fuc).

ESI high resolution MS: $[C_{41}H_{71}N_3O_{25} + H]^+$ calcd 994.4455, found 994.4449.

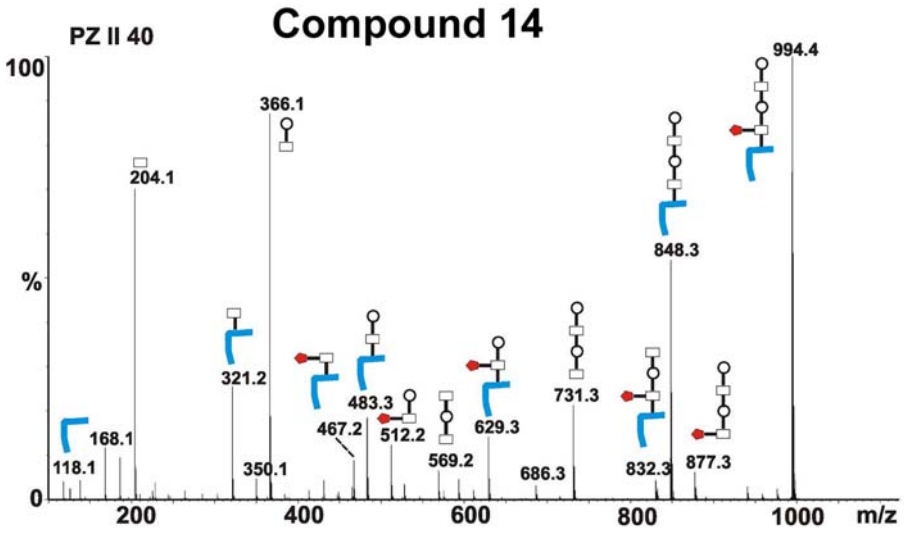
C. Collision-Induced Dissociation (CID) MS Analysis of Synthesized Oligosaccharides







Where: - is the linker at the reducing end
- is HexNAc (N-acetylhexosamine)
- is Hexose
- is fucose



Where: - is the linker at the reducing end
- is HexNAc (N-acetylhexosamine)
- is Hexose
- is fucose

D. Direct ESI-MS assay

Apparent association constants ($K_{a,app}$) for the fragments TcdA-A2 binding to the compound **S259-1** and pentasaccharides **2** were measured using the direct ESI-MS assay. The ESI-MS measurements were carried out using a 9.4T Apex II Fourier-transform ion cyclotron resonance (FTICR) MS (Bruker-Daltonics, Billerica, MA). Complete details of the experimental methodology and data analysis are described elsewhere.[2,3]

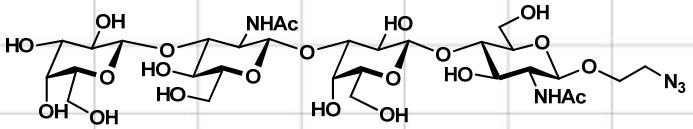
Reference

[1] Vasiliu, D.; Razi, N.; Zhang, Y.; Jacobsen, N.; Allin, K.; Liu, X.; Hoffmann, J.; Bohorov, O.; Blixt, O. *Carbohydr. Res.* **2006**, *341*, 1447–1457.

[2] Sun, J.; Kitova, E. N.; Wang, W.; Klassen, J. S. *Anal. Chem*. **2006**, *78*, 3010-3018.

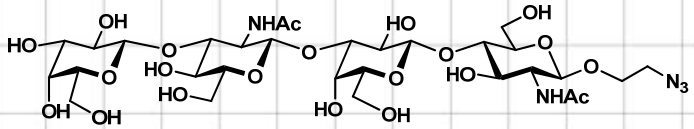
[3] Wang, W.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2003**, *75*, 4945-4955.

^{1}H NMR in D_2O, 400 MHz



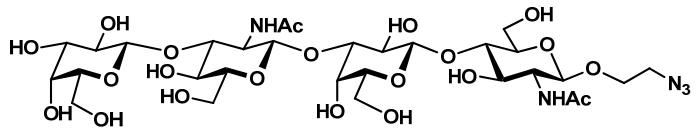
5

^{1}H NMR in D_2O, 400 MHz

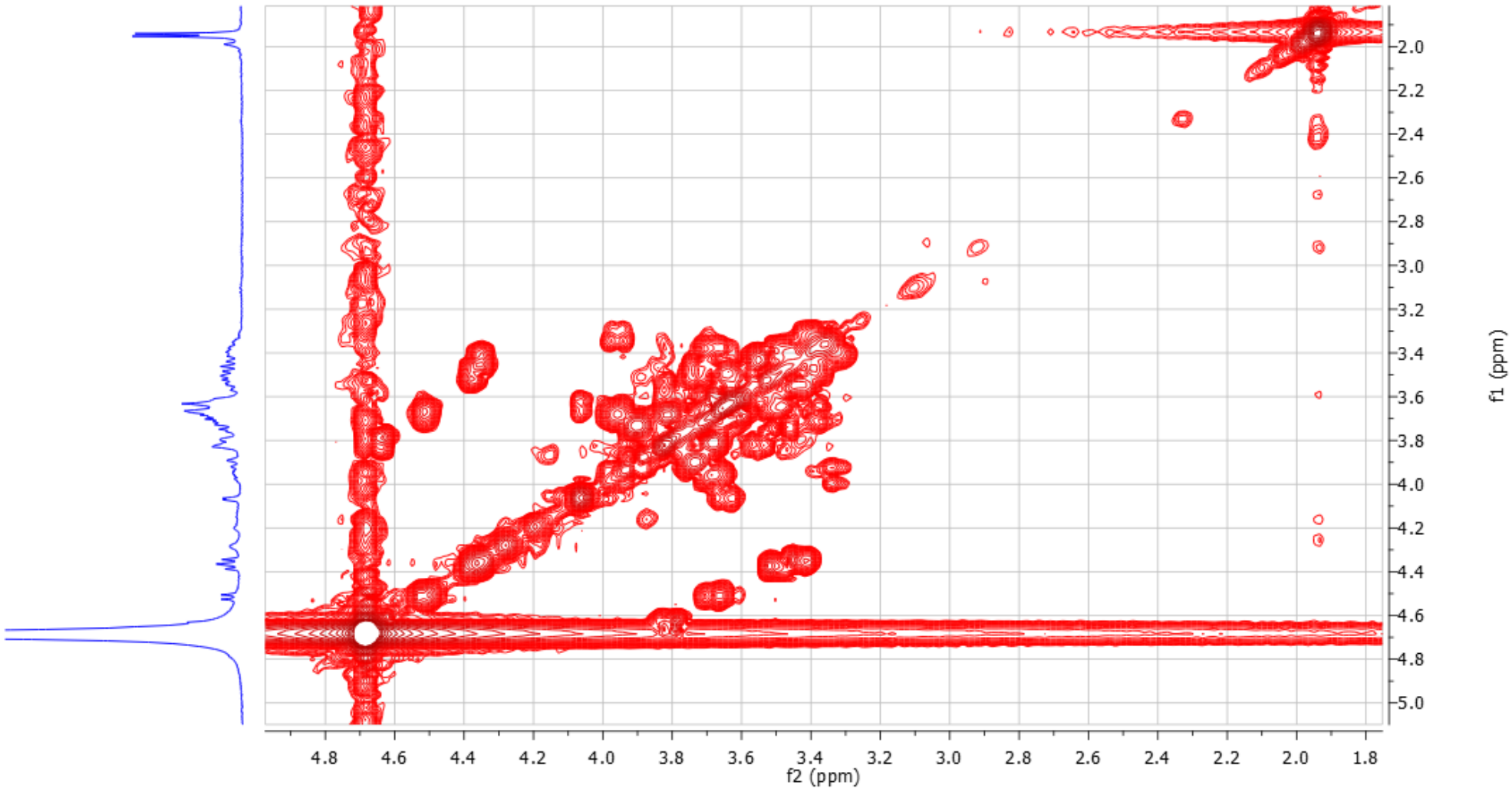


5

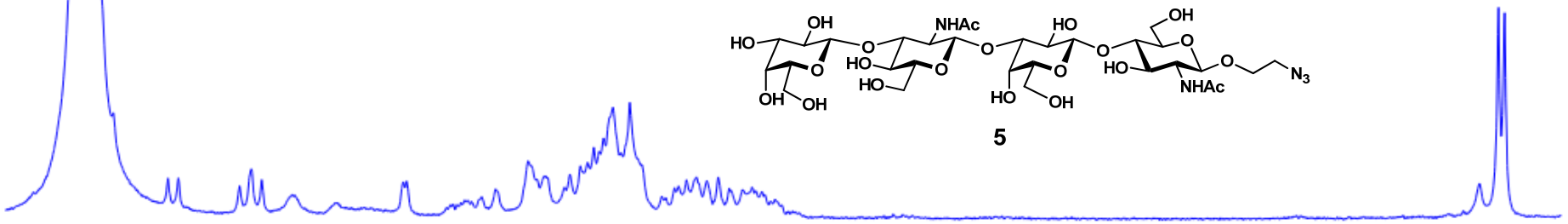
^{1}H-^{1}H GCOSY NMR in D_2O, 400 MHz



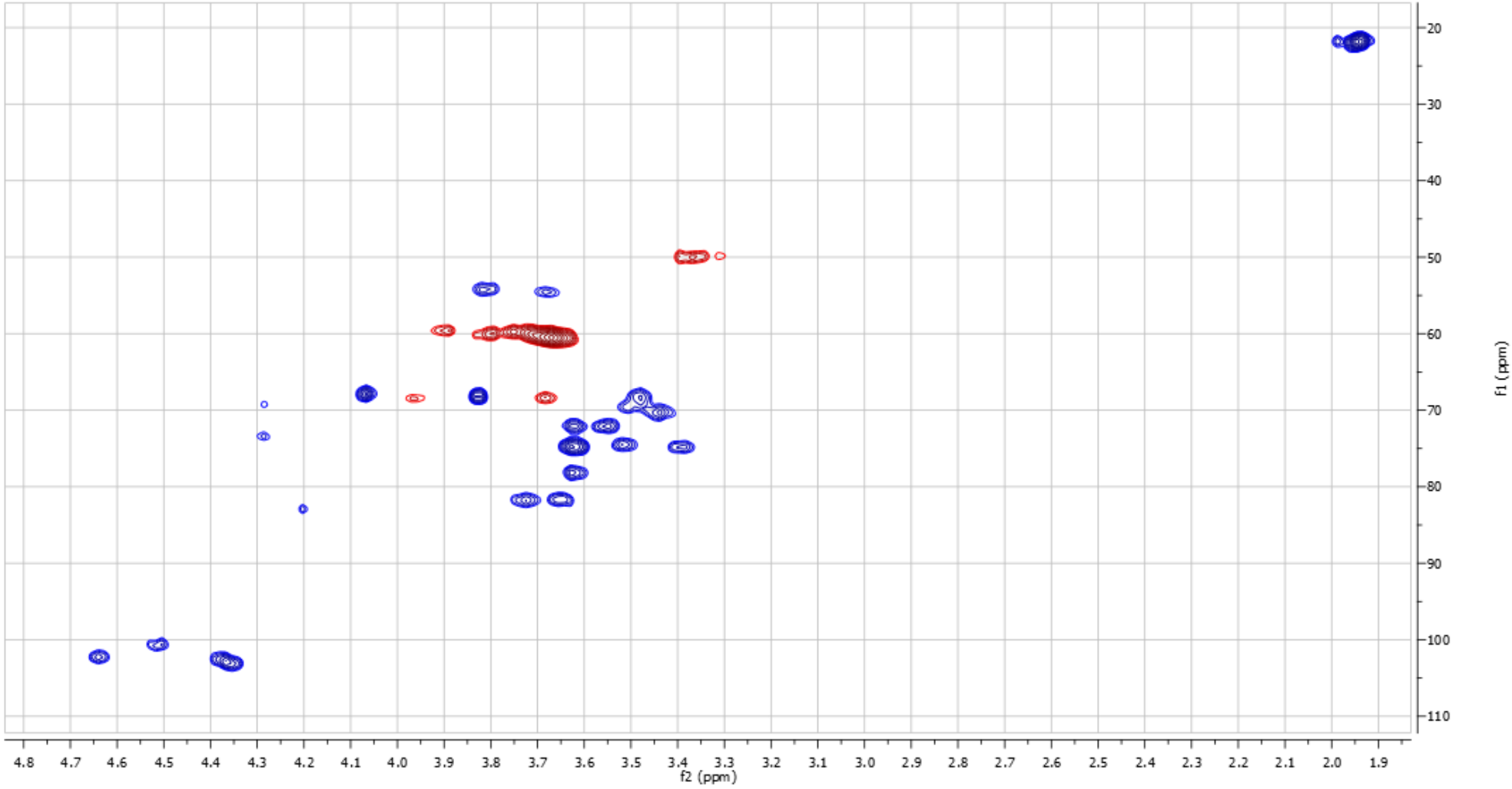
5



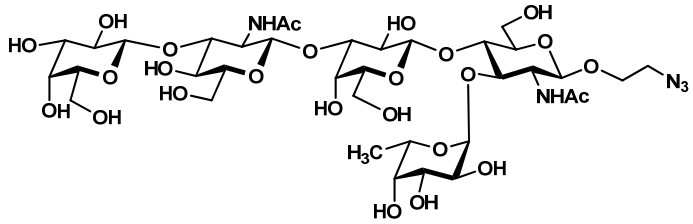
^{1}H-^{13}C GHSQC NMR in D_2O, 400 MHz



5

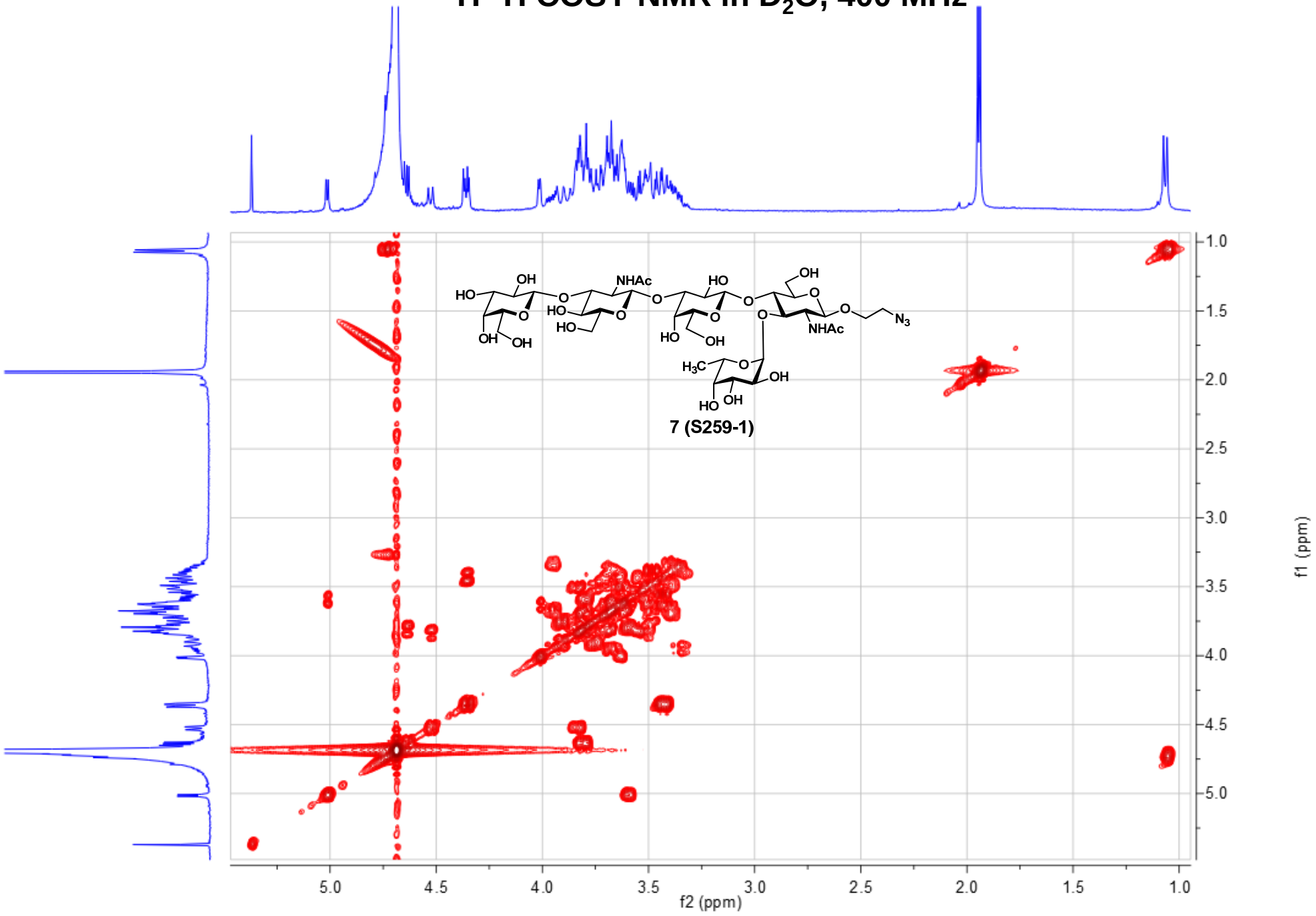


^{1}H NMR in D_2O, 400 MHz

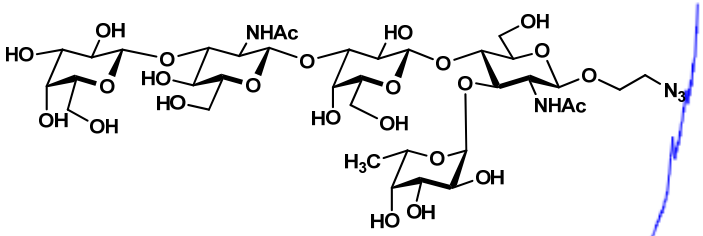


7 (S259-1)

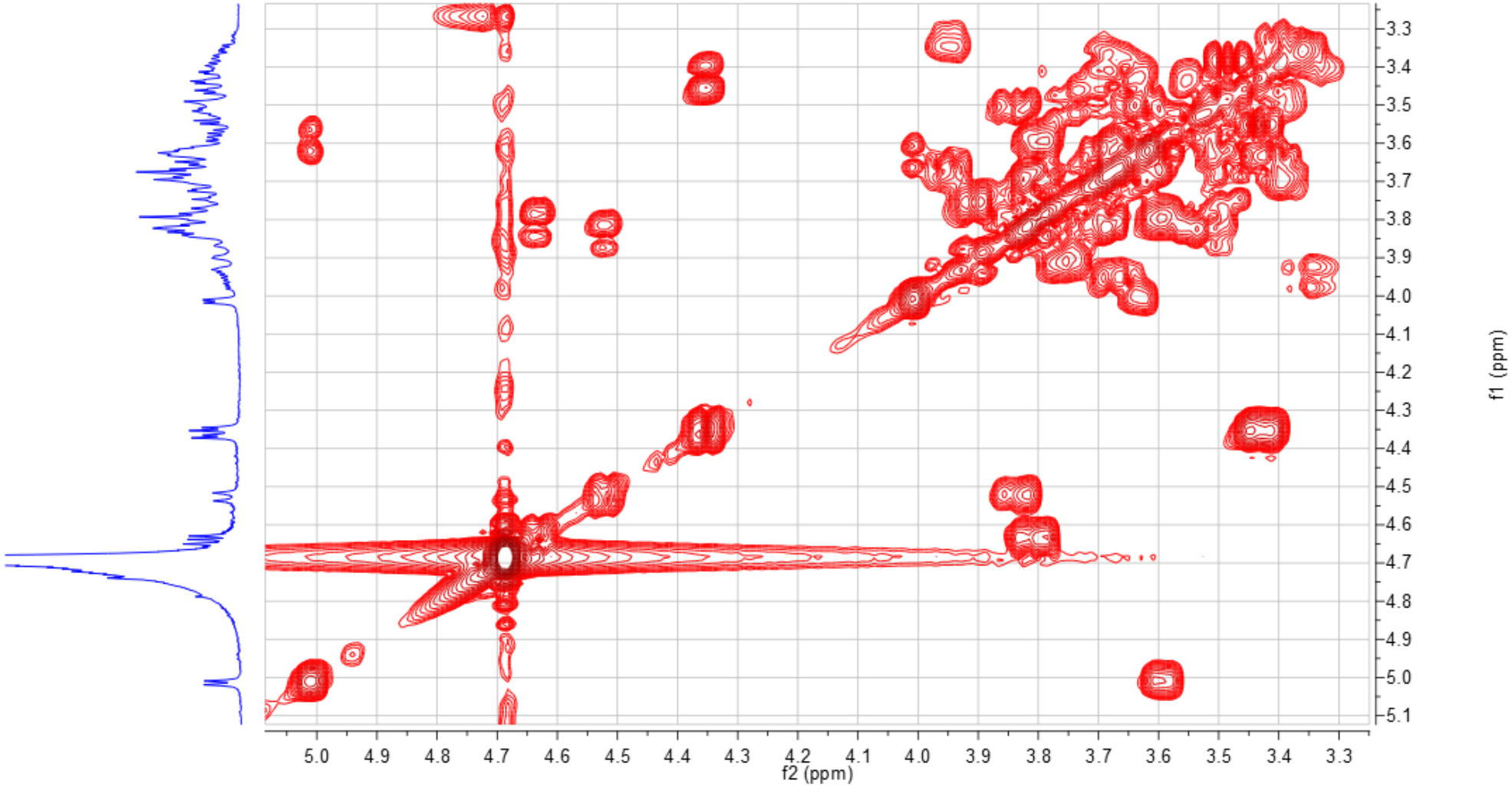
^{1}H-^{1}H COSY NMR in D_2O, 400 MHz



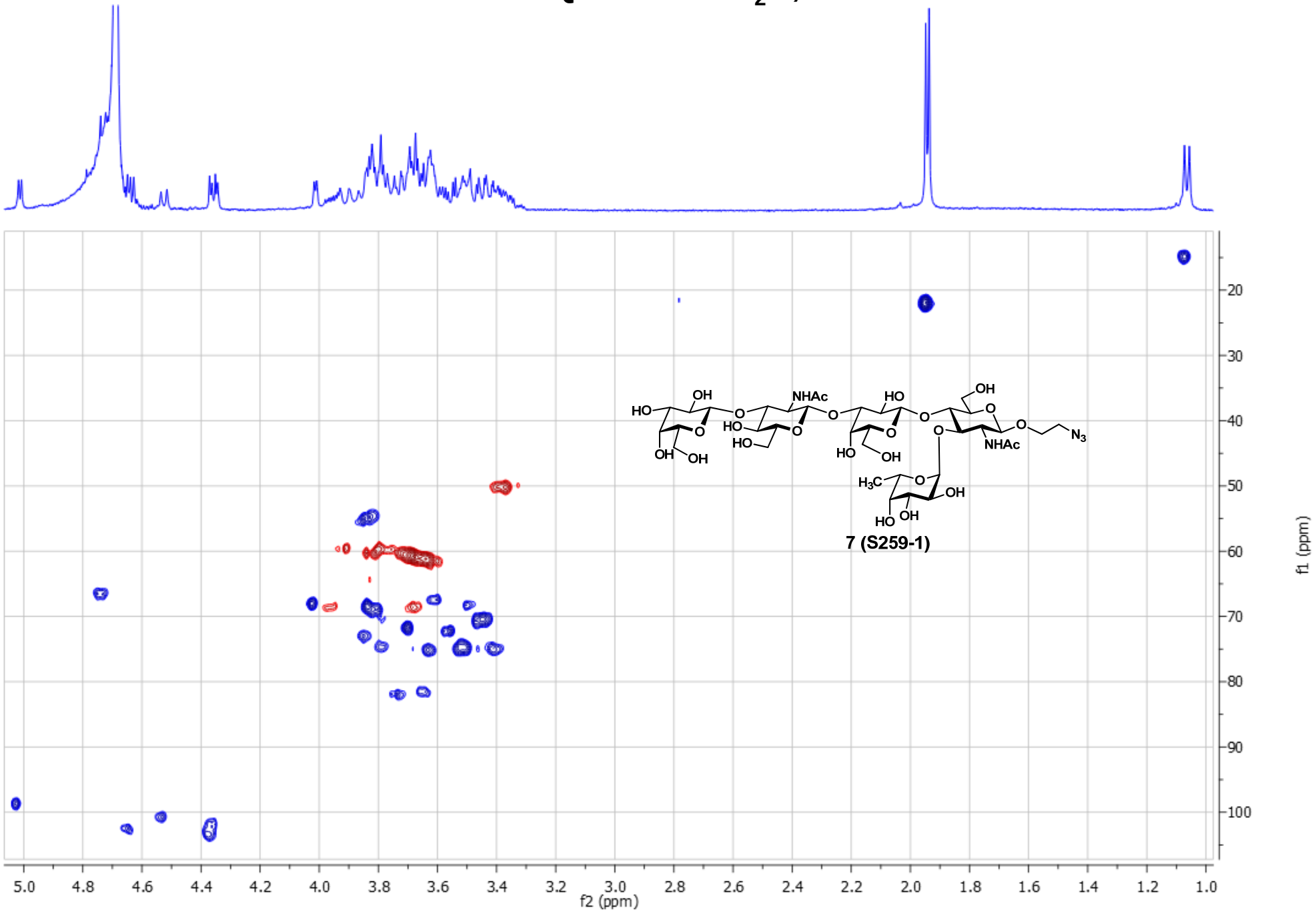
^{1}H-^{1}H COSY NMR in D_2O, 400 MHz



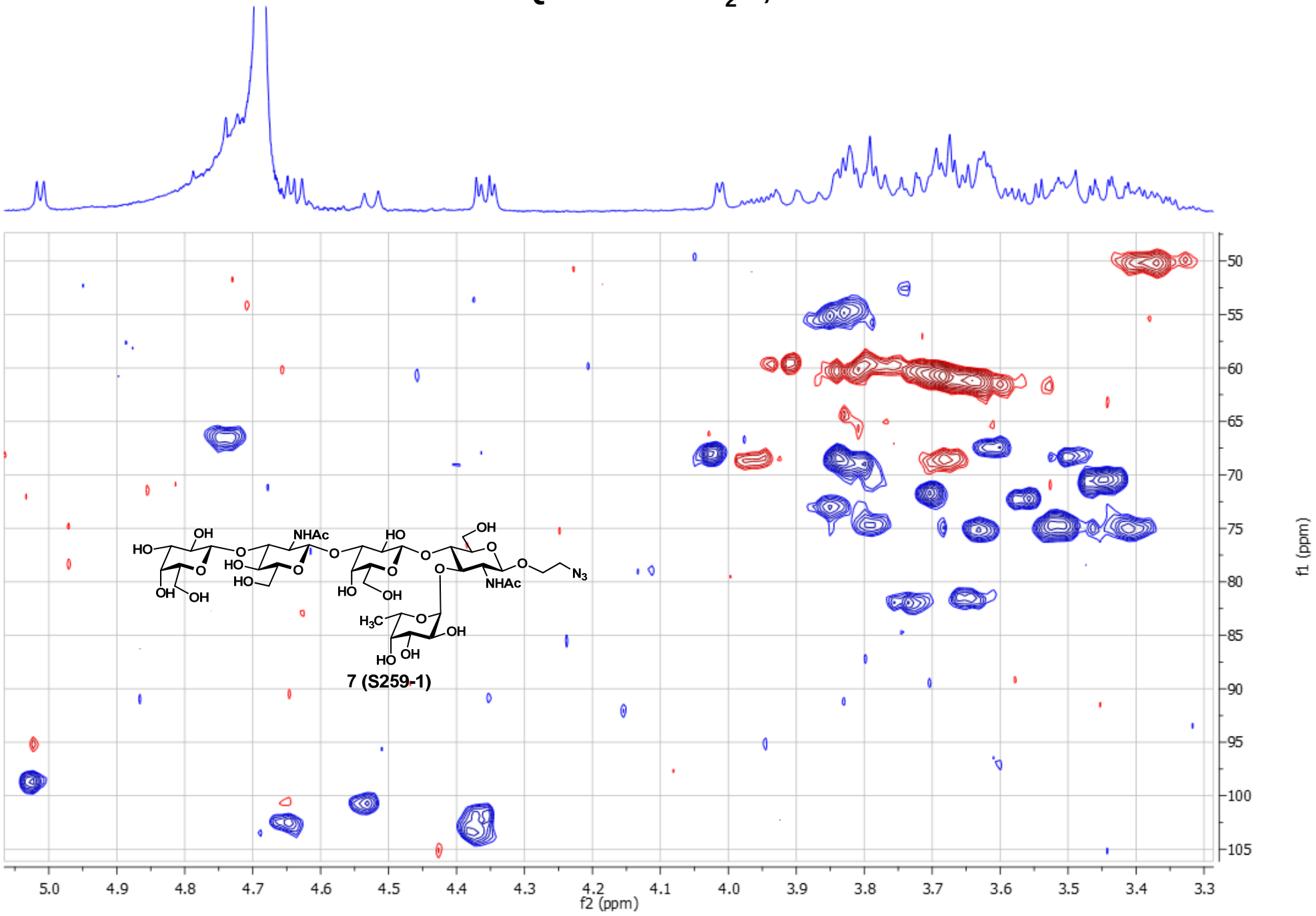
7 (S259-1)



^{1}H-^{13}C GHSQC NMR in D_2O, 400 MHz

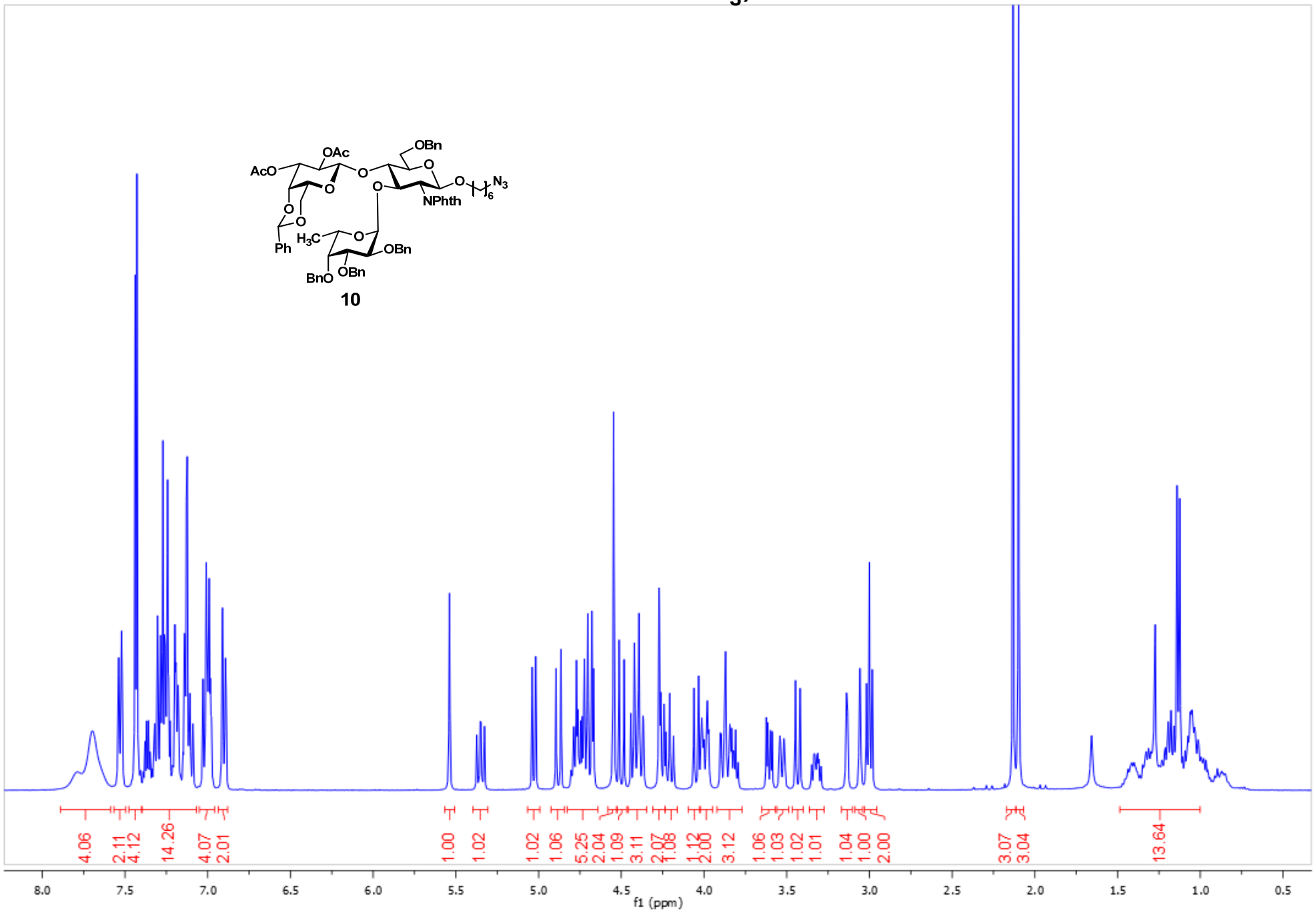


^{1}H-^{13}C GHSQC NMR in D_2O, 400 MHz

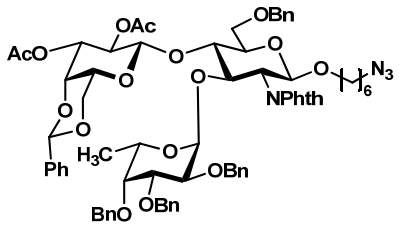


7 (S259-1)

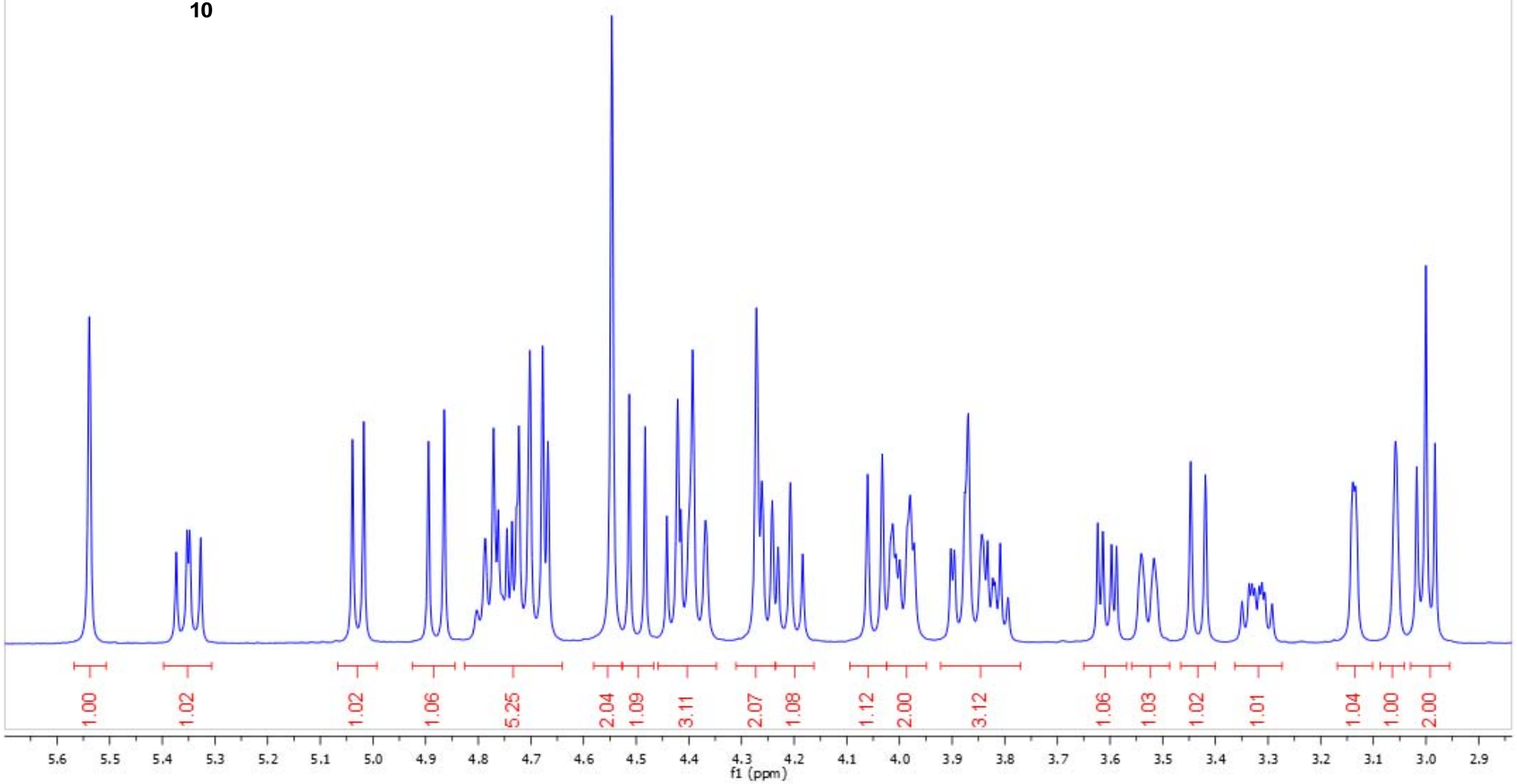
^{1}H NMR in $CDCl_3$, 400 MHz



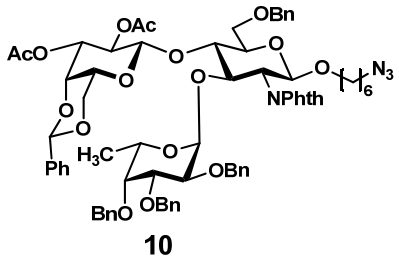
^{1}H NMR in $CDCl_3$, 400 MHz



10

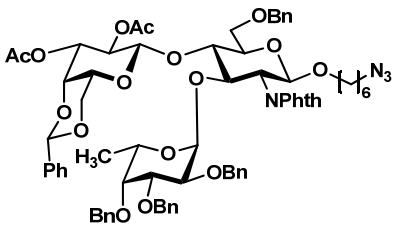


^{1}H-^{1}H COSY NMR in $CDCl_3$, 400 MHz



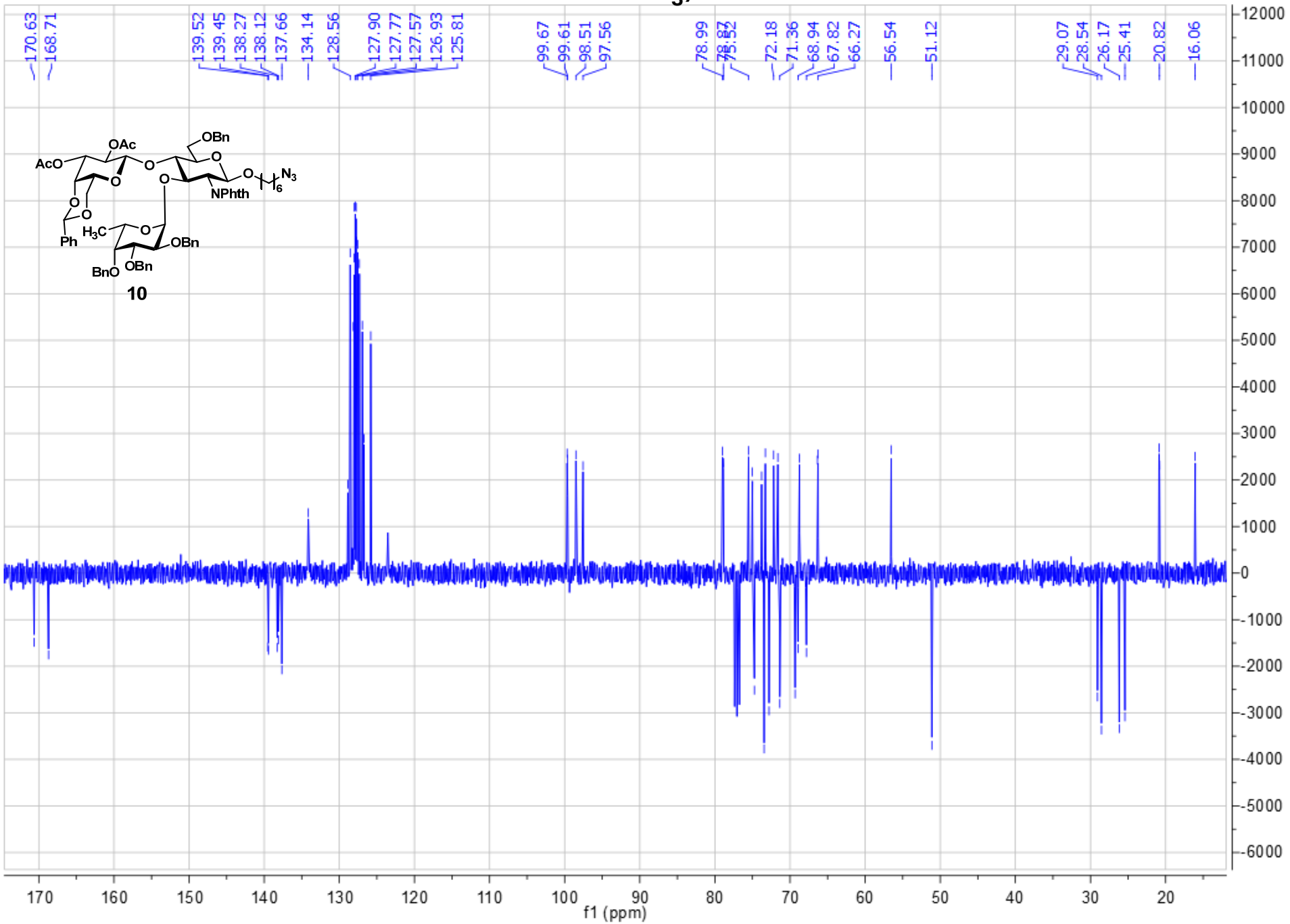
10

^{1}H-^{1}H COSY NMR in $CDCl_3$, 400 MHz

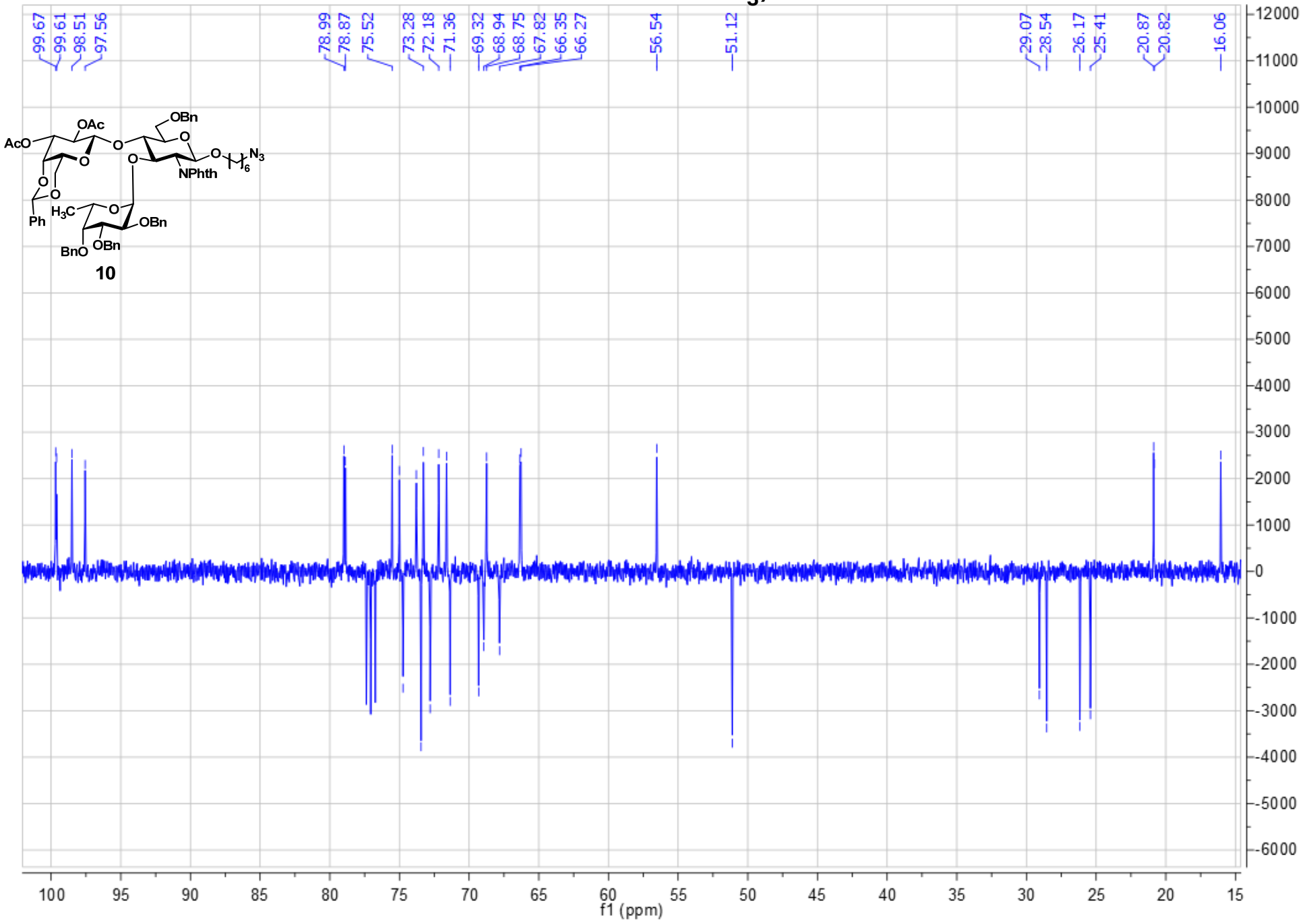


10

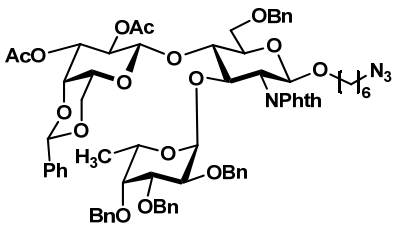
^{13}C NMR in $CDCl_3$, 100 MHz



^{13}C NMR in $CDCl_3$, 100 MHz

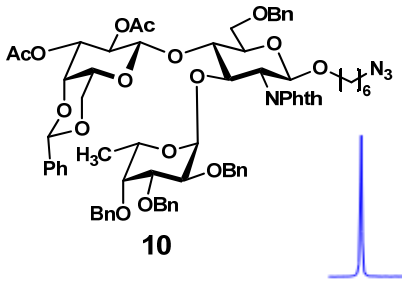


^{1}H-^{13}C GHSQC NMR in $CDCl_3$, 400 MHz



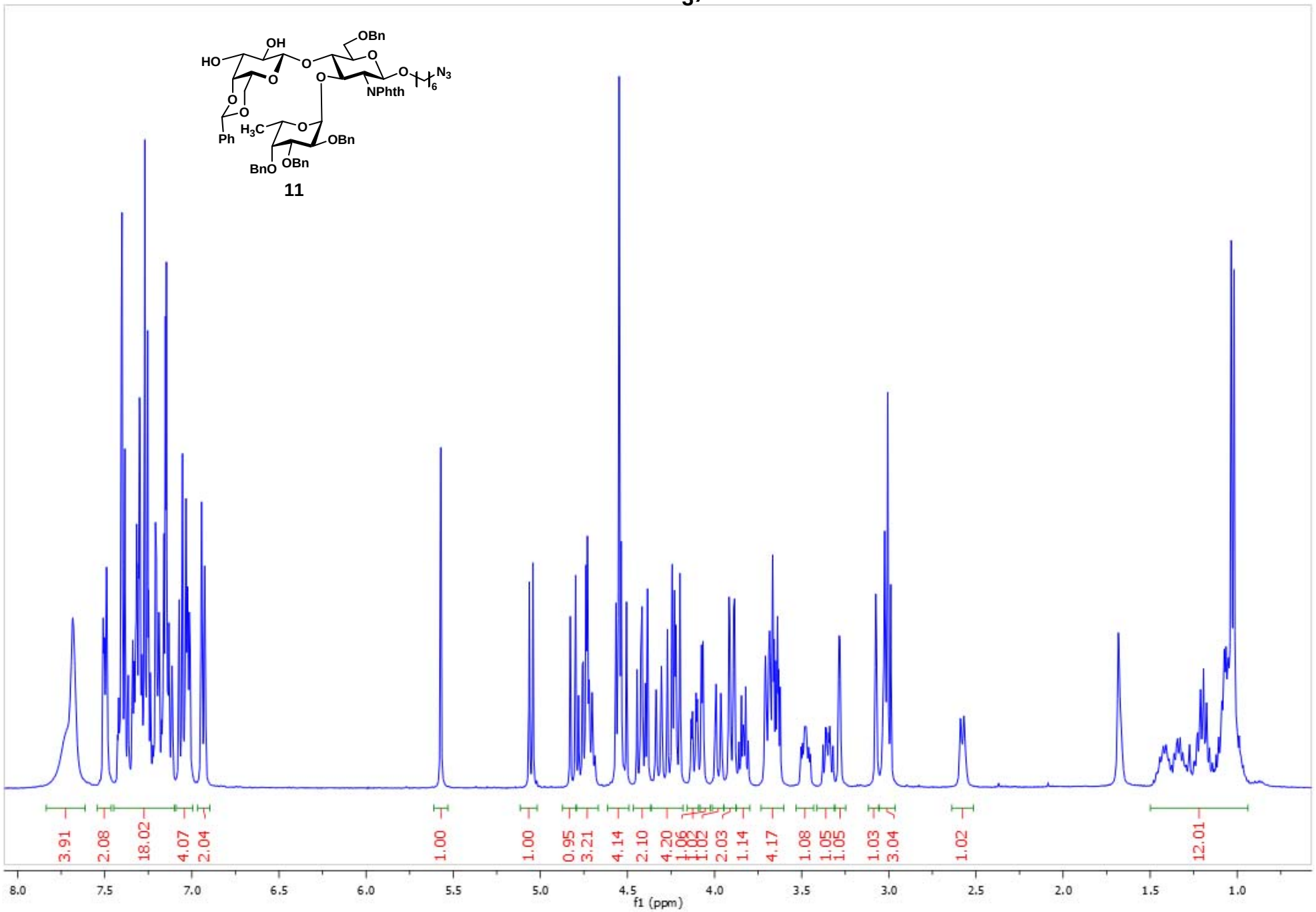
10

^{1}H-^{13}C GHSQC NMR in $CDCl_3$, 400 MHz

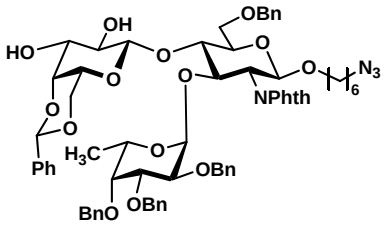


10

^{1}H NMR in $CDCl_3$, 400 MHz

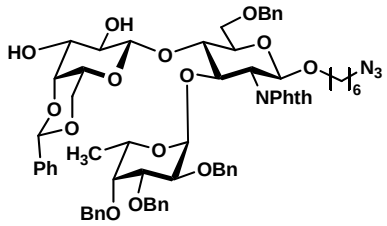


^{1}H NMR in $CDCl_3$, 400 MHz



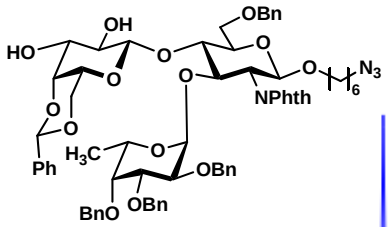
11

^{1}H-^{1}H COSY NMR in $CDCl_3$, 400 MHz



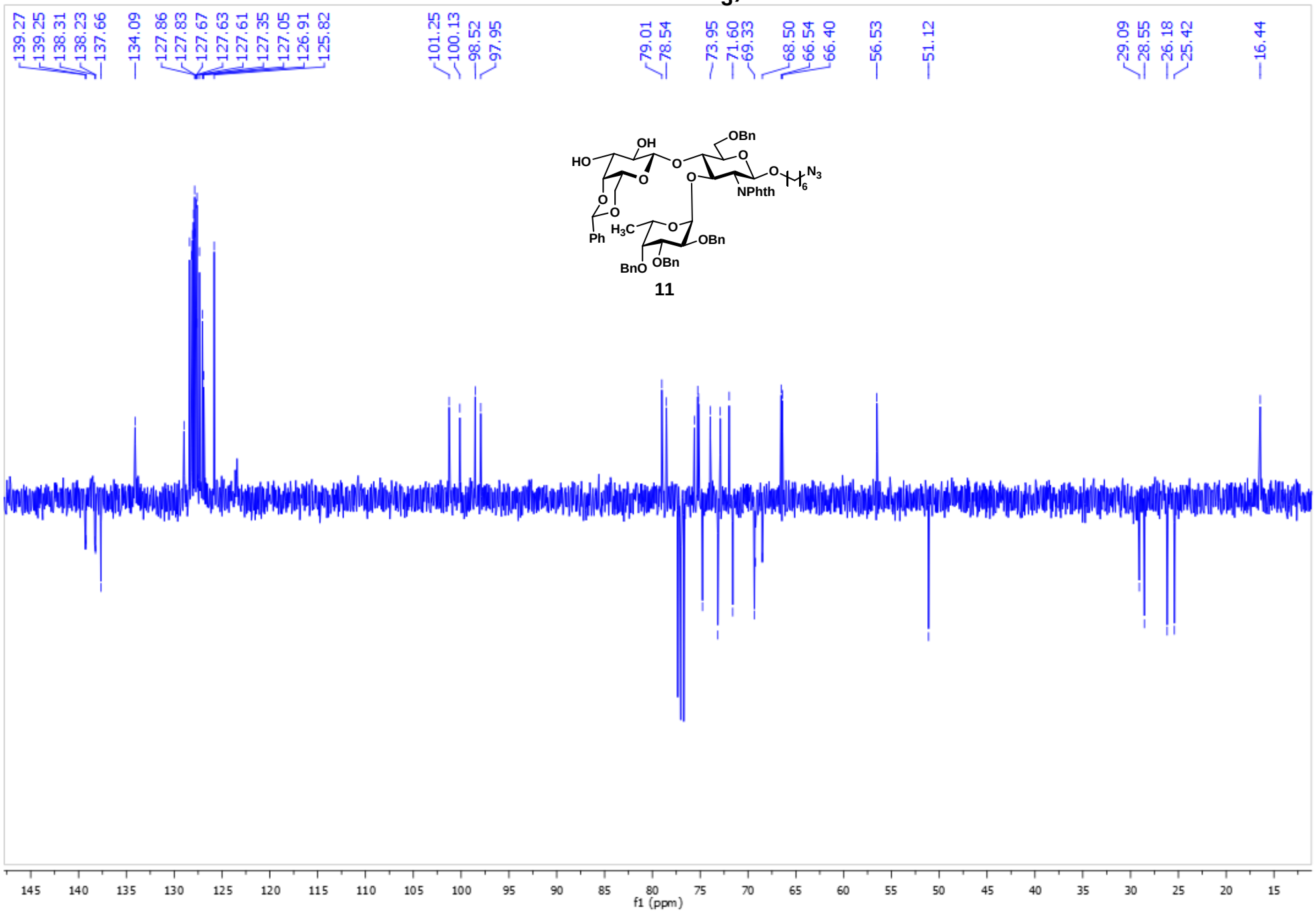
11

^{1}H-^{1}H COSY NMR in $CDCl_3$, 400 MHz



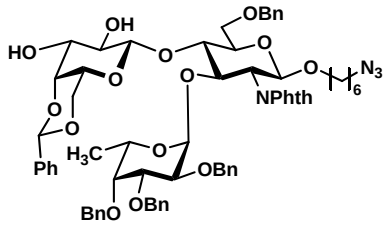
11

^{13}C NMR in $CDCl_3$, 100 MHz



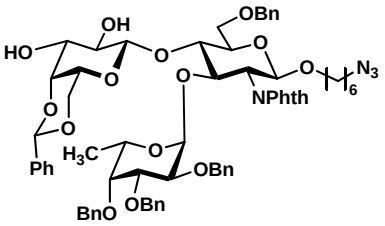
11

^{1}H-^{13}C GHSQC NMR in $CDCl_3$, 400 MHz

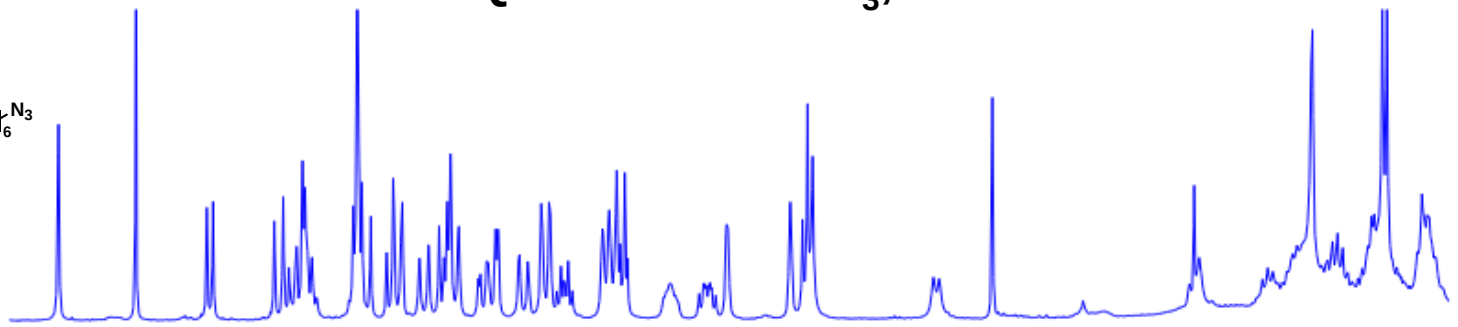


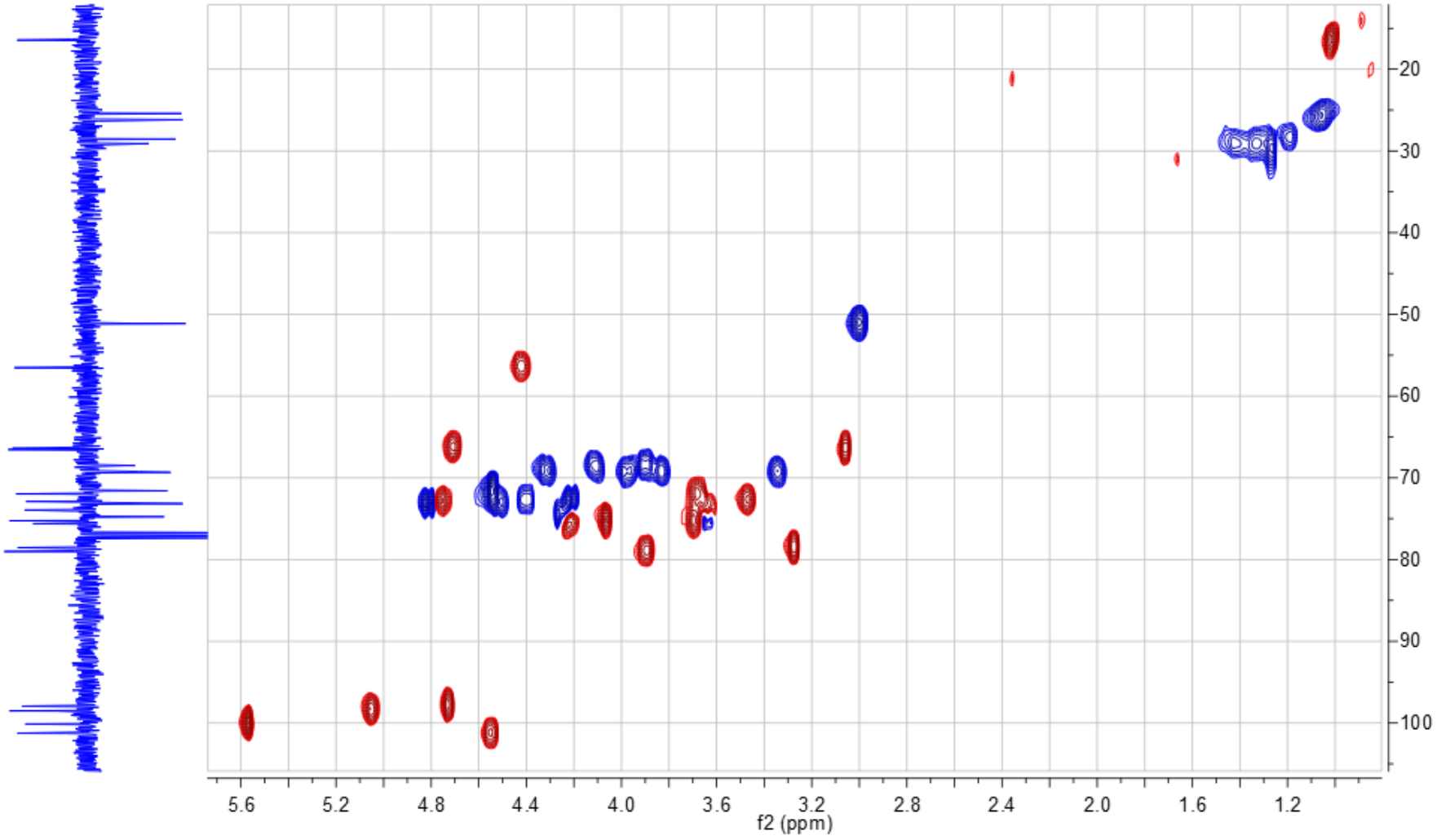
11

^{1}H-^{13}C GHSQC NMR in $CDCl_3$, 400 MHz

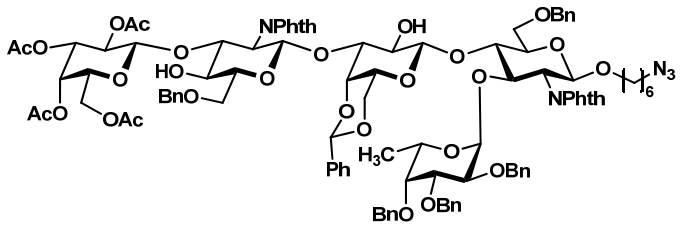


11



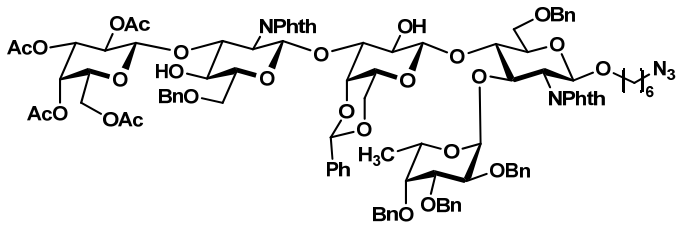


^{1}H NMR in $CDCl_3$, 400 MHz

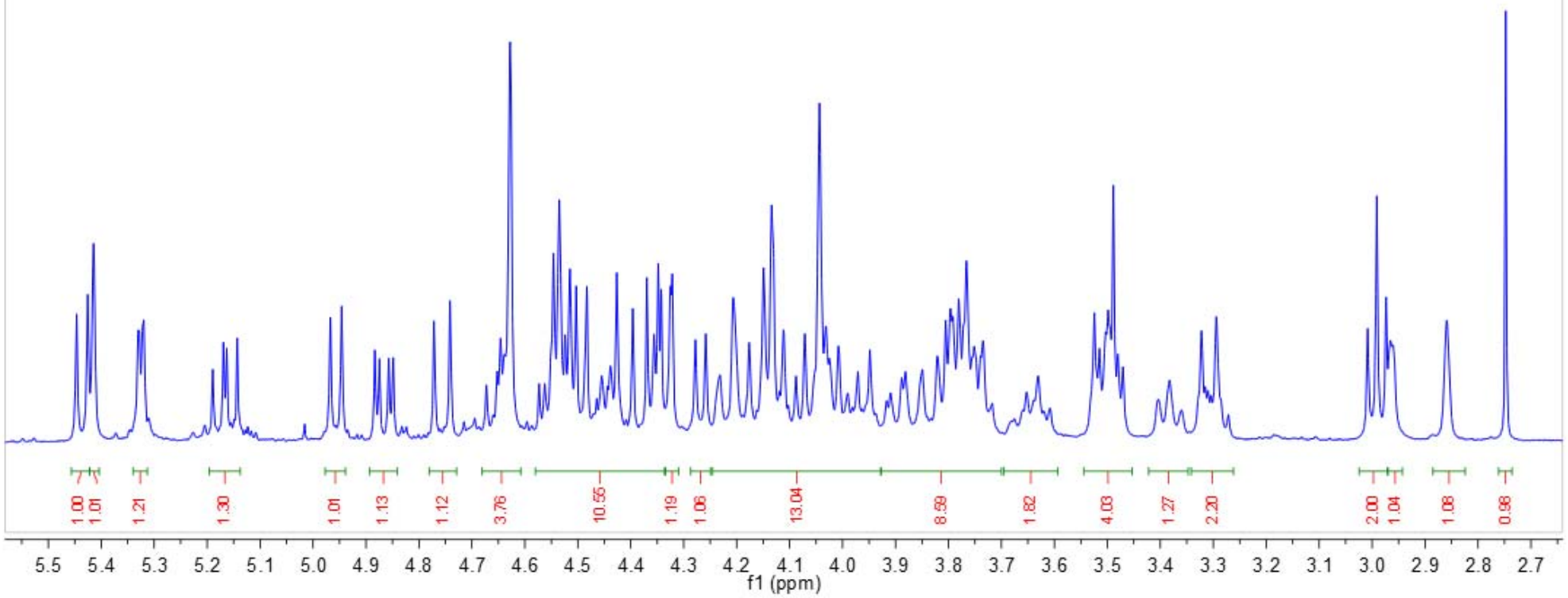


13

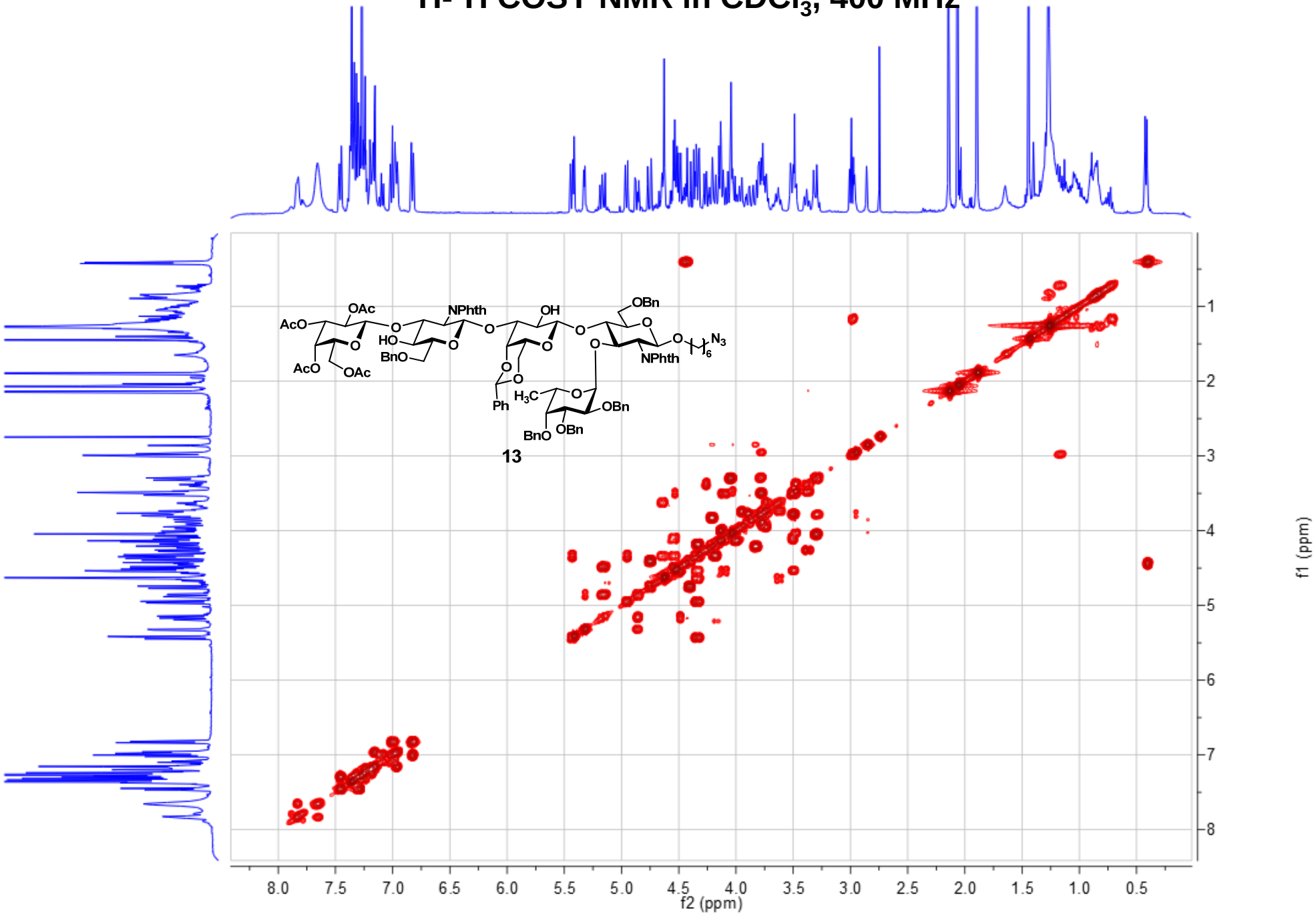
^{1}H NMR in $CDCl_3$, 400 MHz



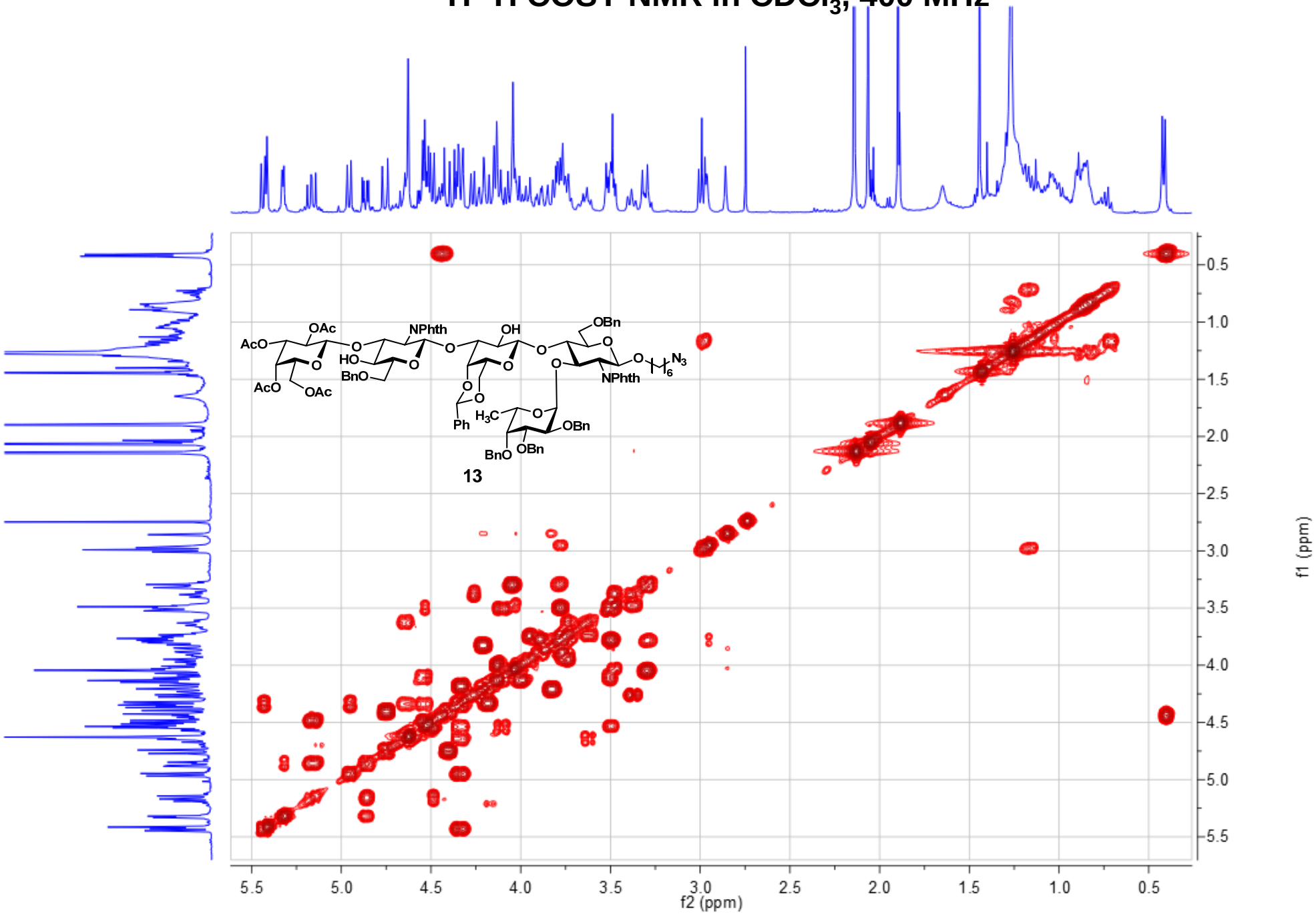
13



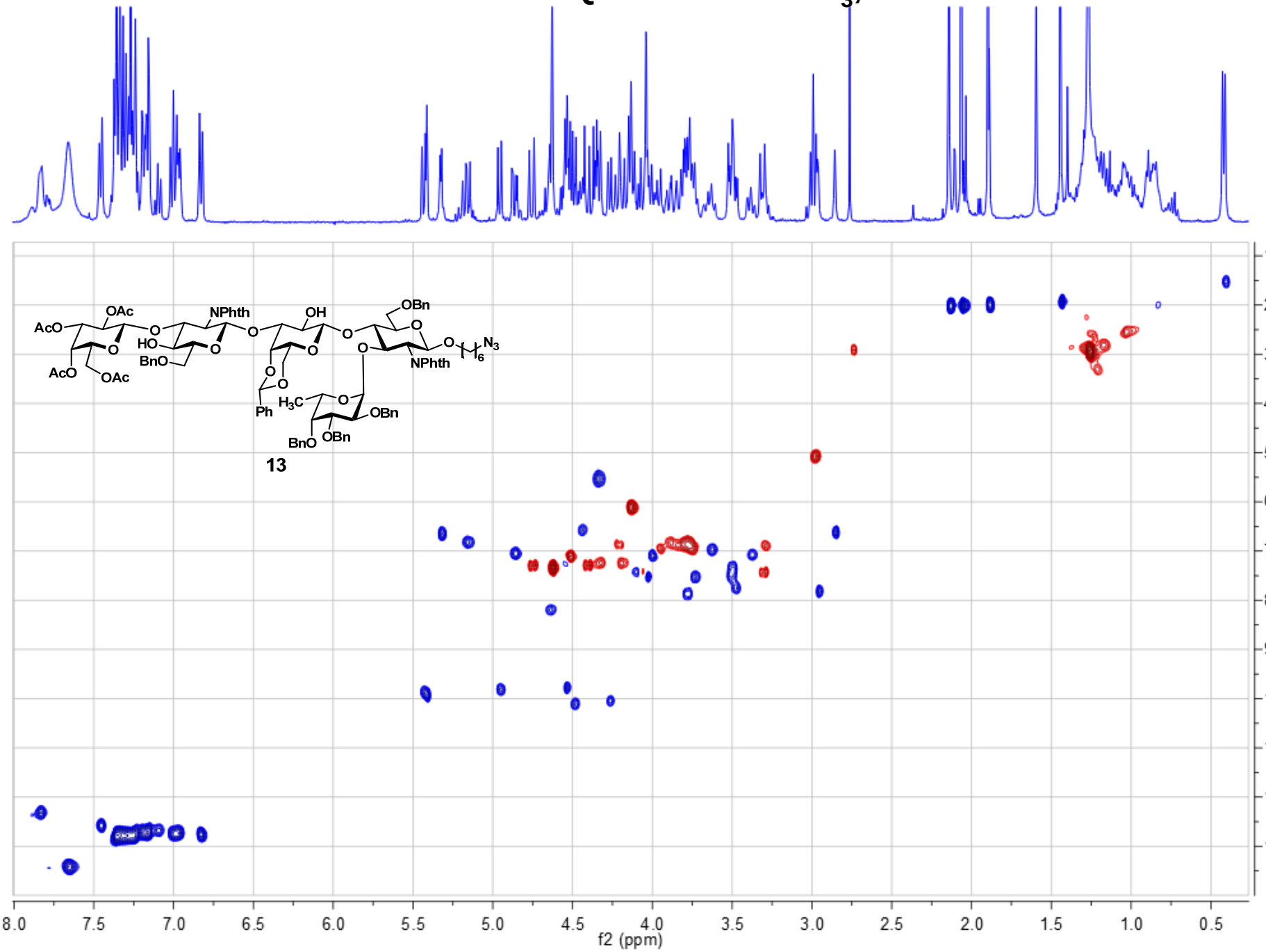
^{1}H-^{1}H COSY NMR in $CDCl_3$, 400 MHz



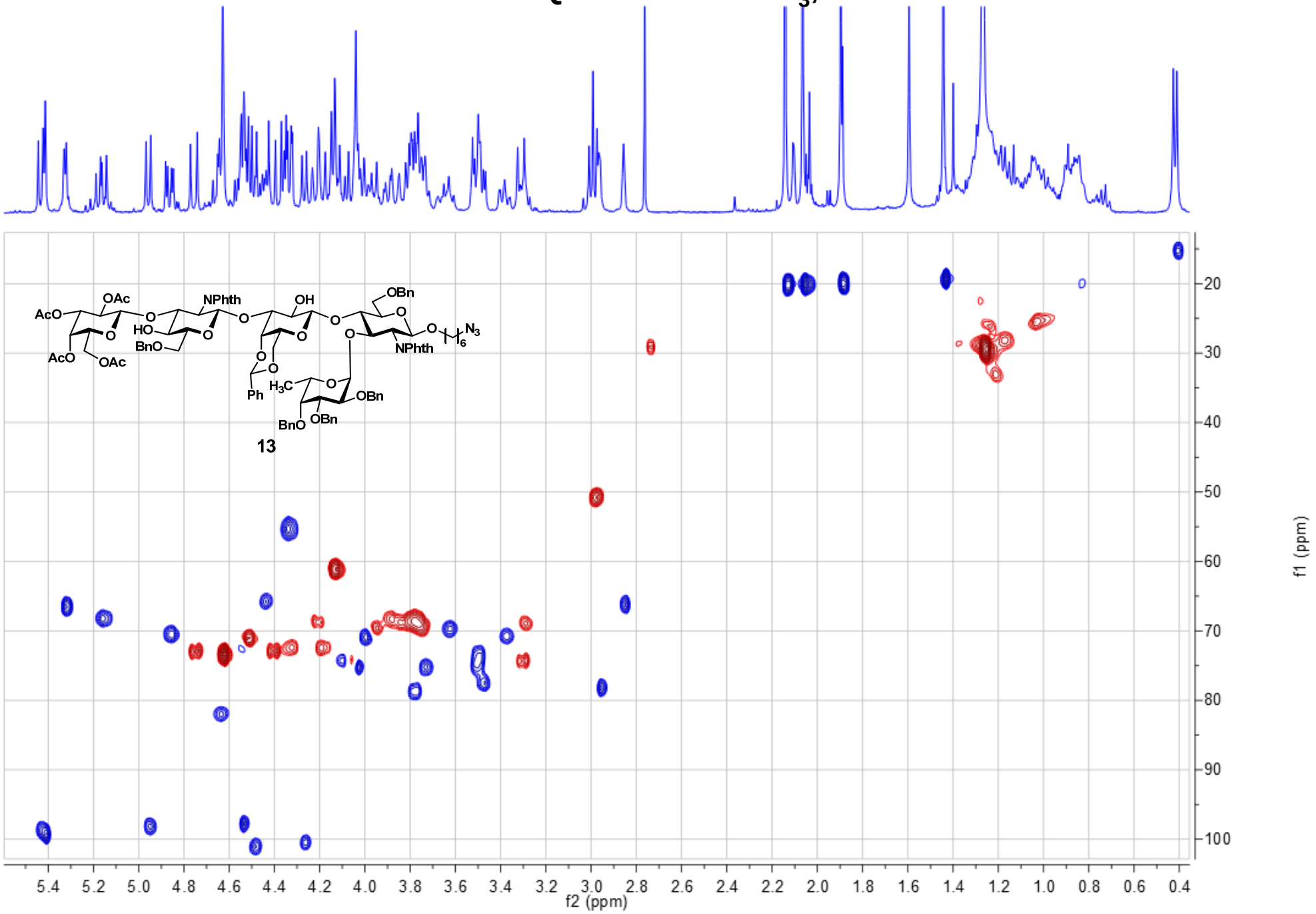
^{1}H-^{1}H COSY NMR in $CDCl_3$, 400 MHz



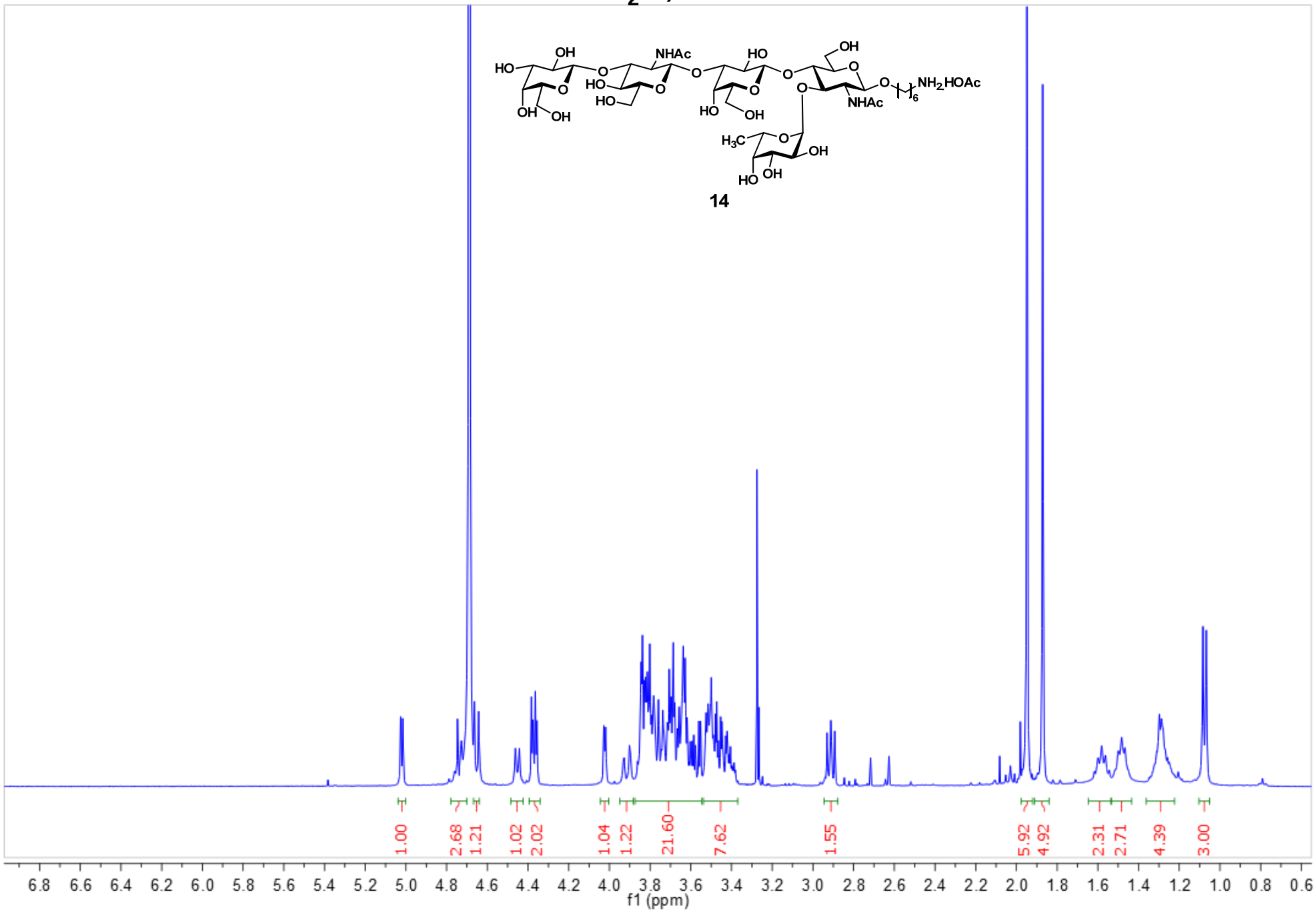
^{1}H-^{13}C GHSQC NMR in $CDCl_3$, 400 MHz

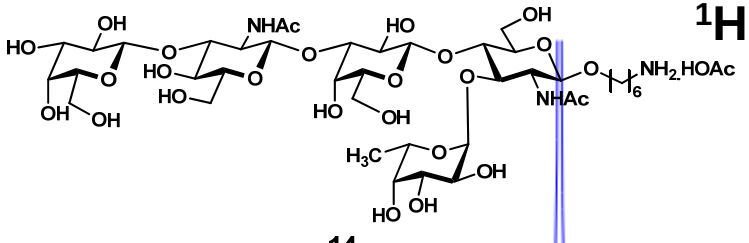


^{1}H-^{13}C GHSQC NMR in $CDCl_3$, 400 MHz



^{1}H NMR in D_2O, 400 MHz

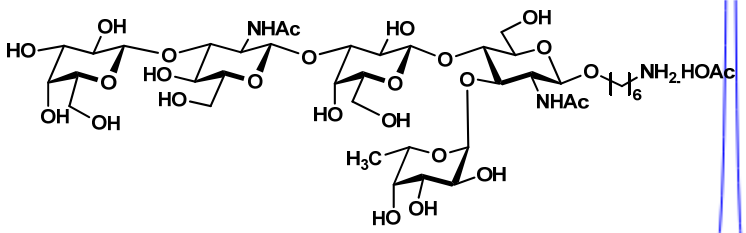




14

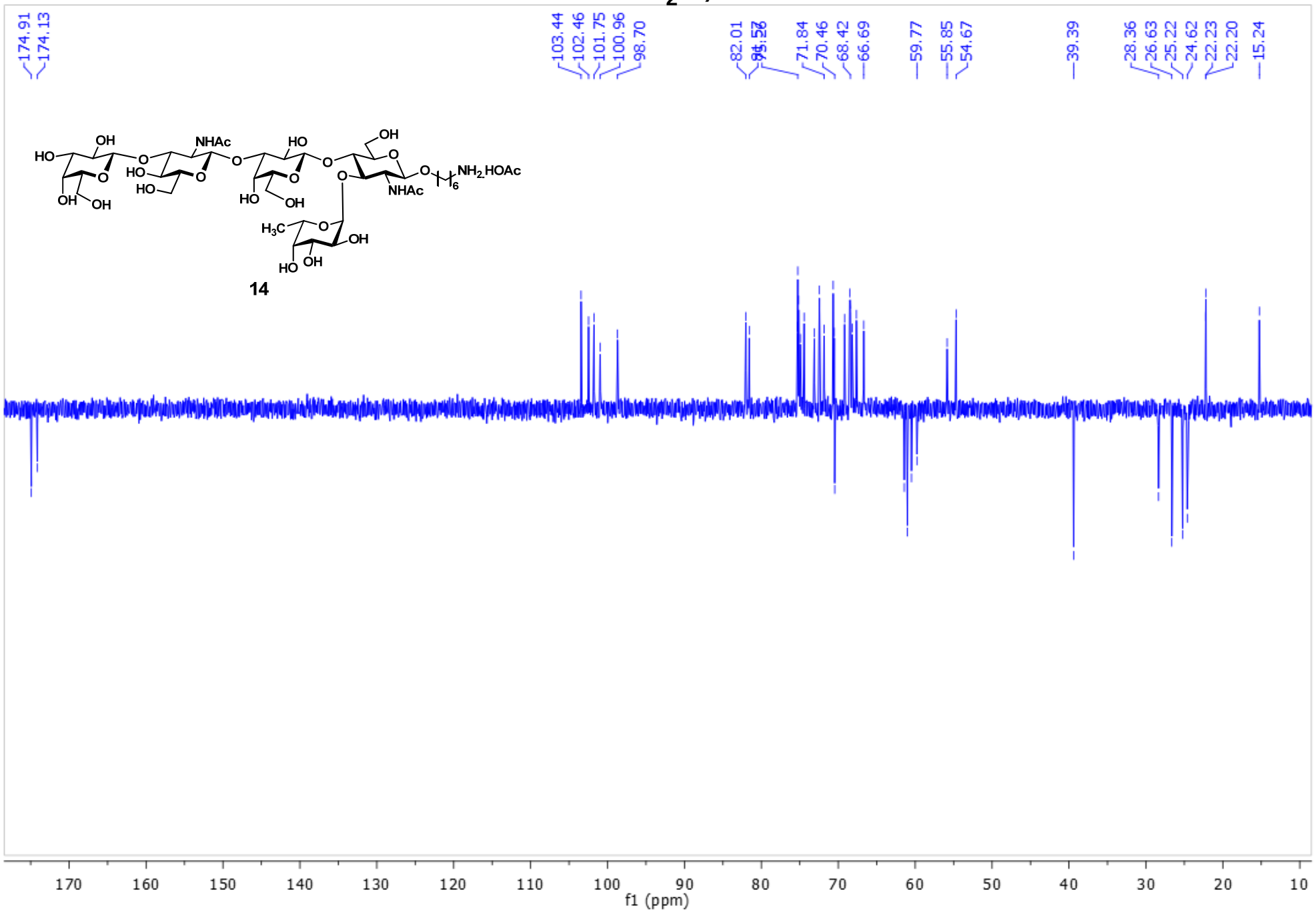
^{1}H-^{1}H COSY NMR in D_2O, 400 MHz

^{1}H-^{1}H COSY NMR in D_2O, 400 MHz



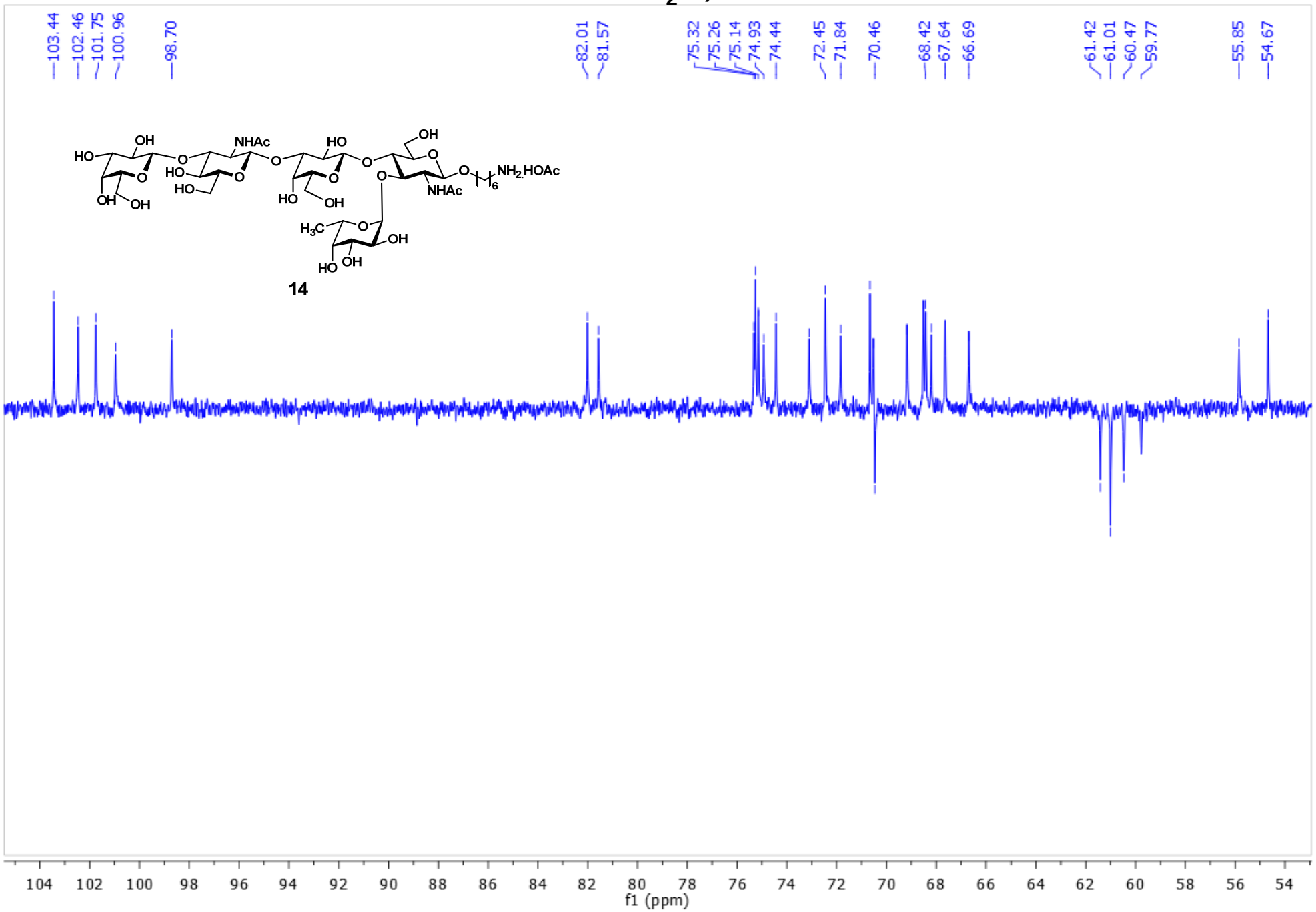
14

^{13}C NMR in D_2O, 100 MHz



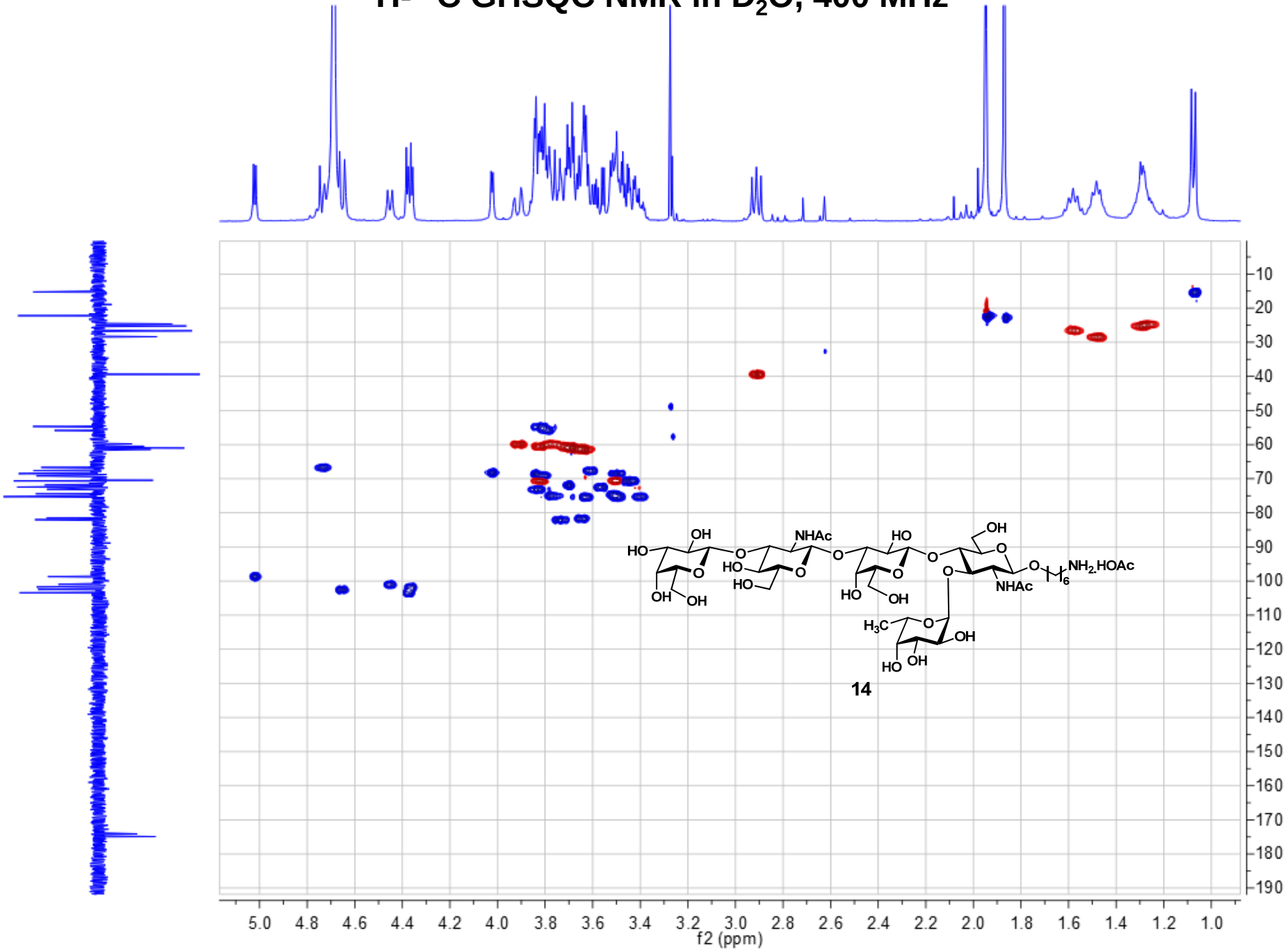
14

^{13}C NMR in D_2O, 100 MHz

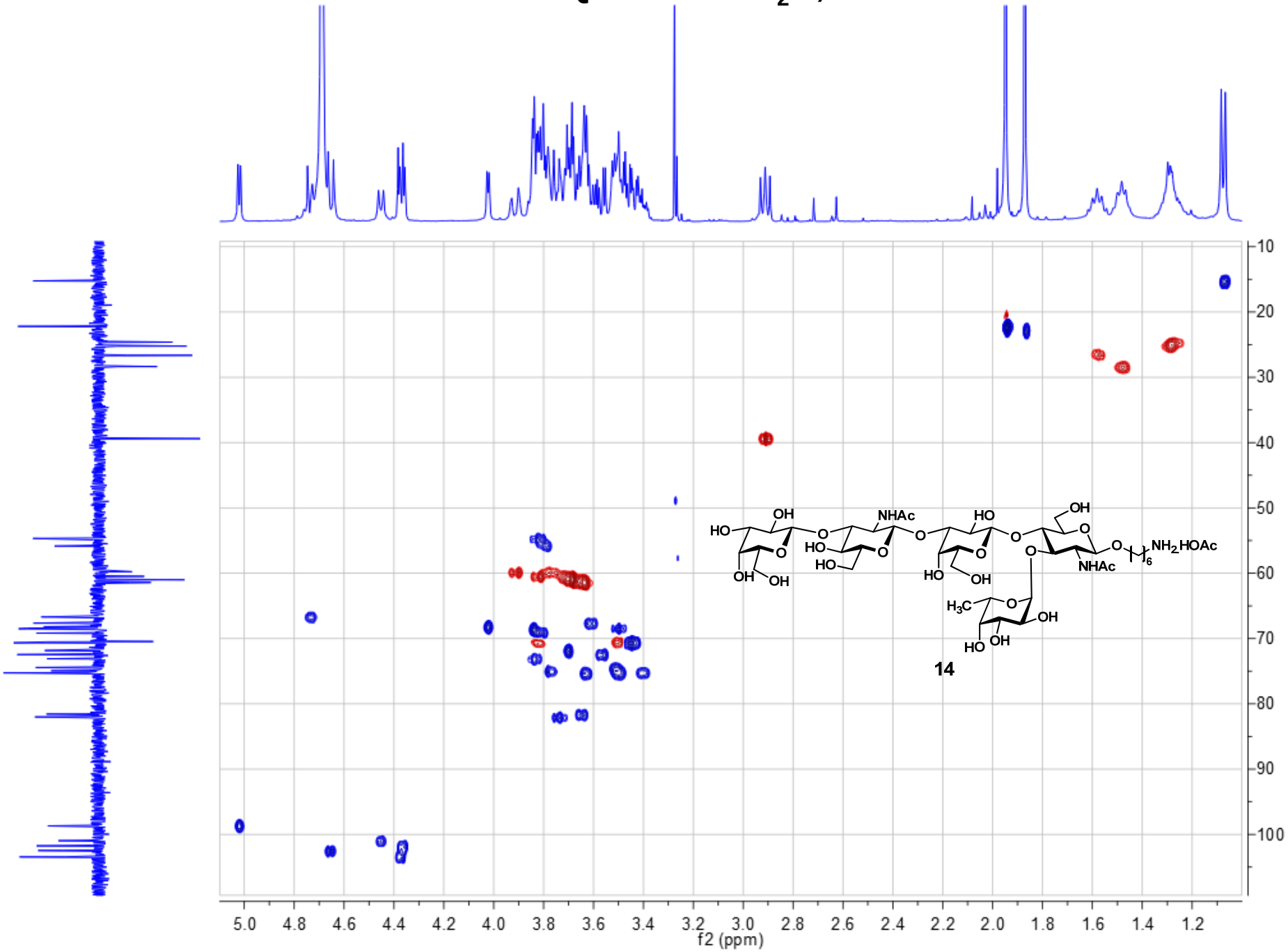


14

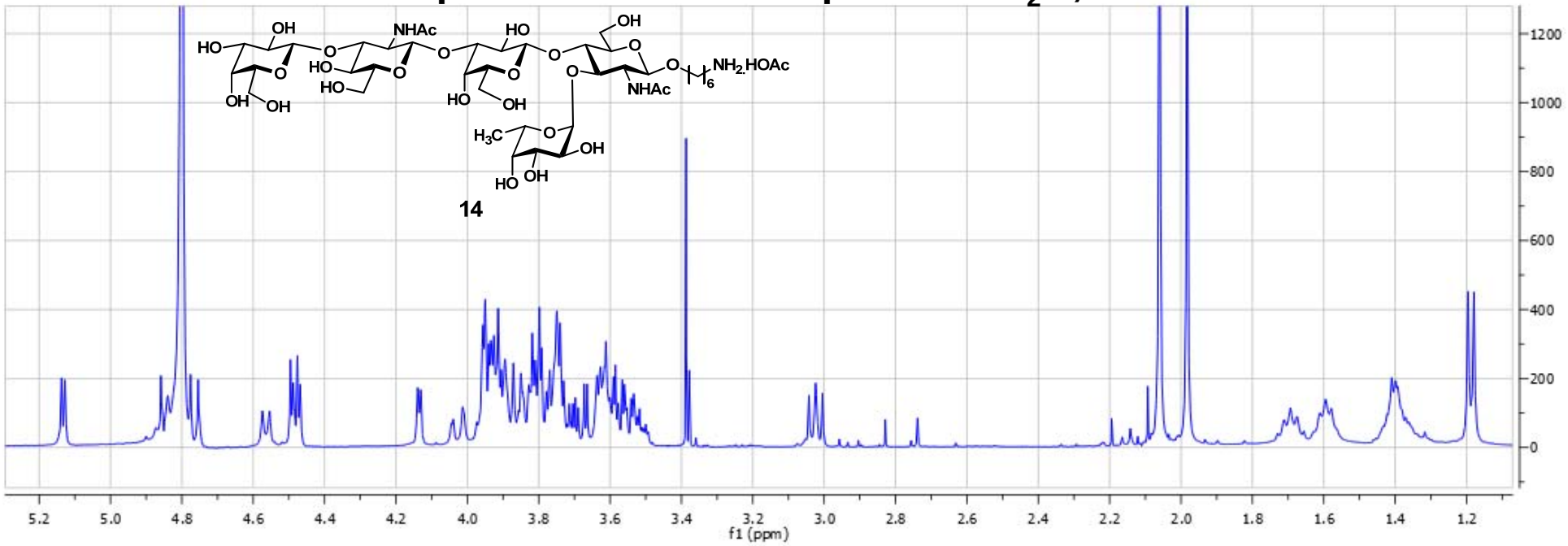
^{1}H-^{13}C GHSQC NMR in D_2O, 400 MHz

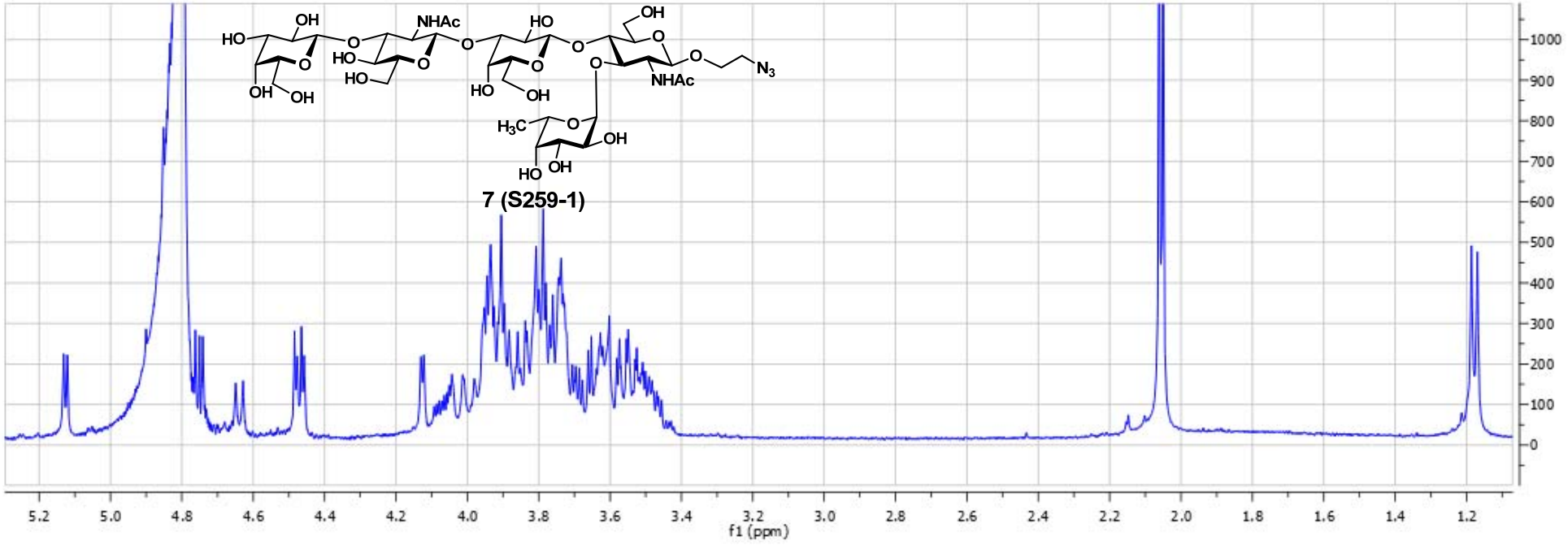


^{1}H-^{13}C GHSQC NMR in D_2O, 400 MHz

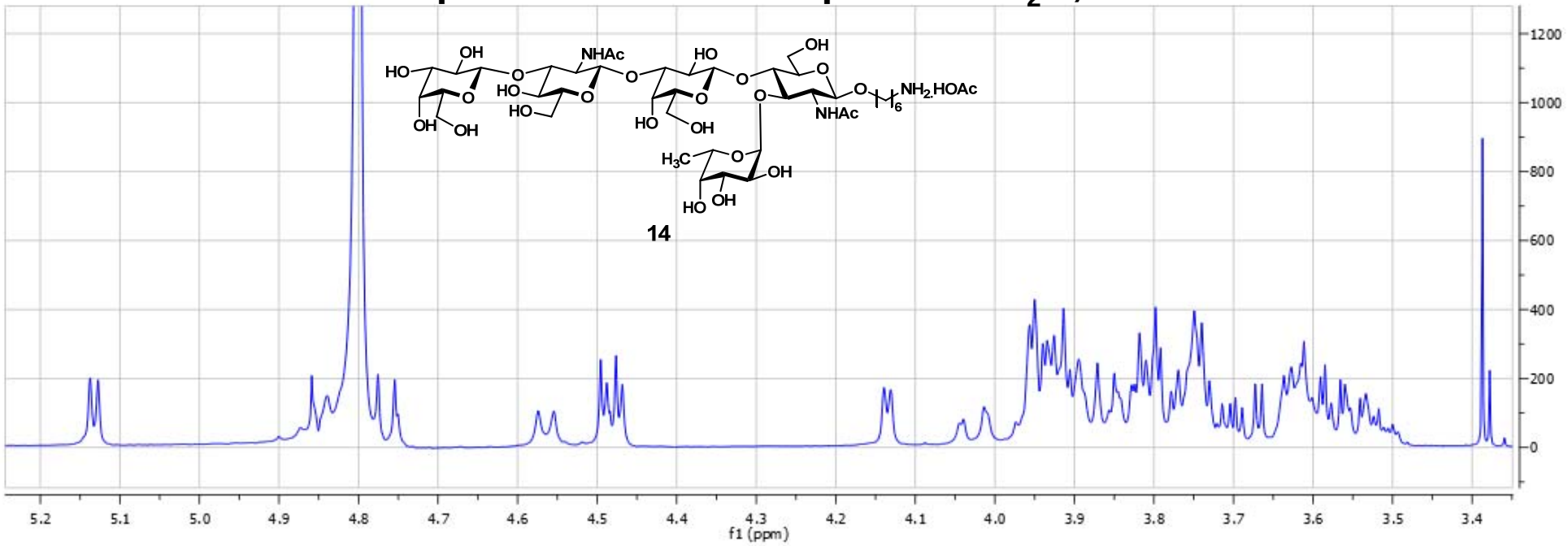


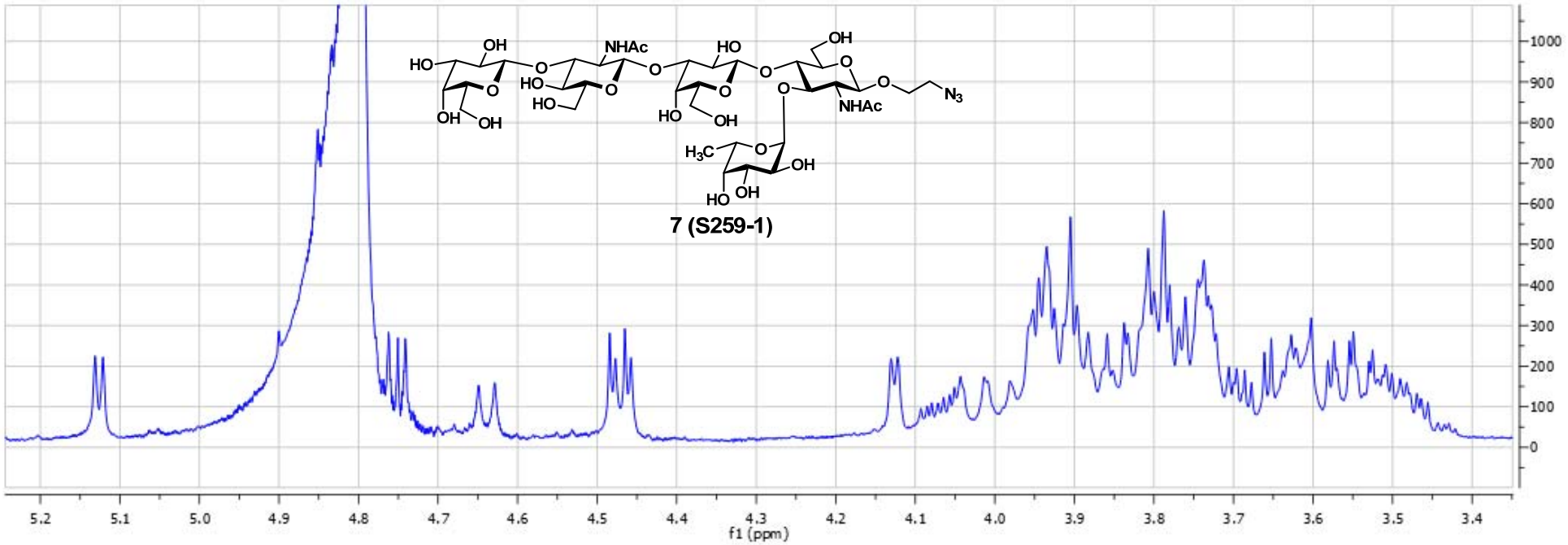
Comparison of ^1H NMR Spectra in D_2O, 400 MHz



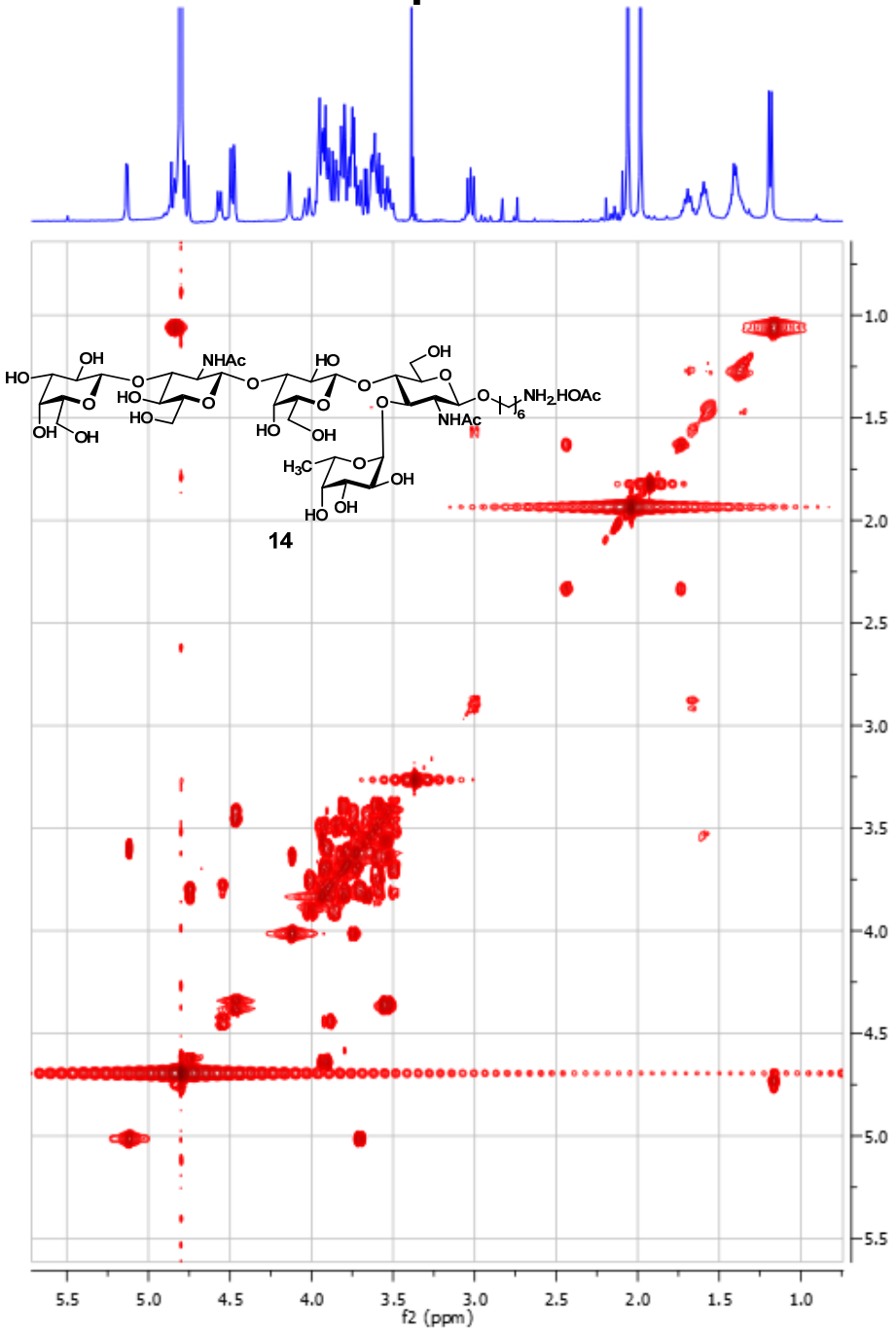


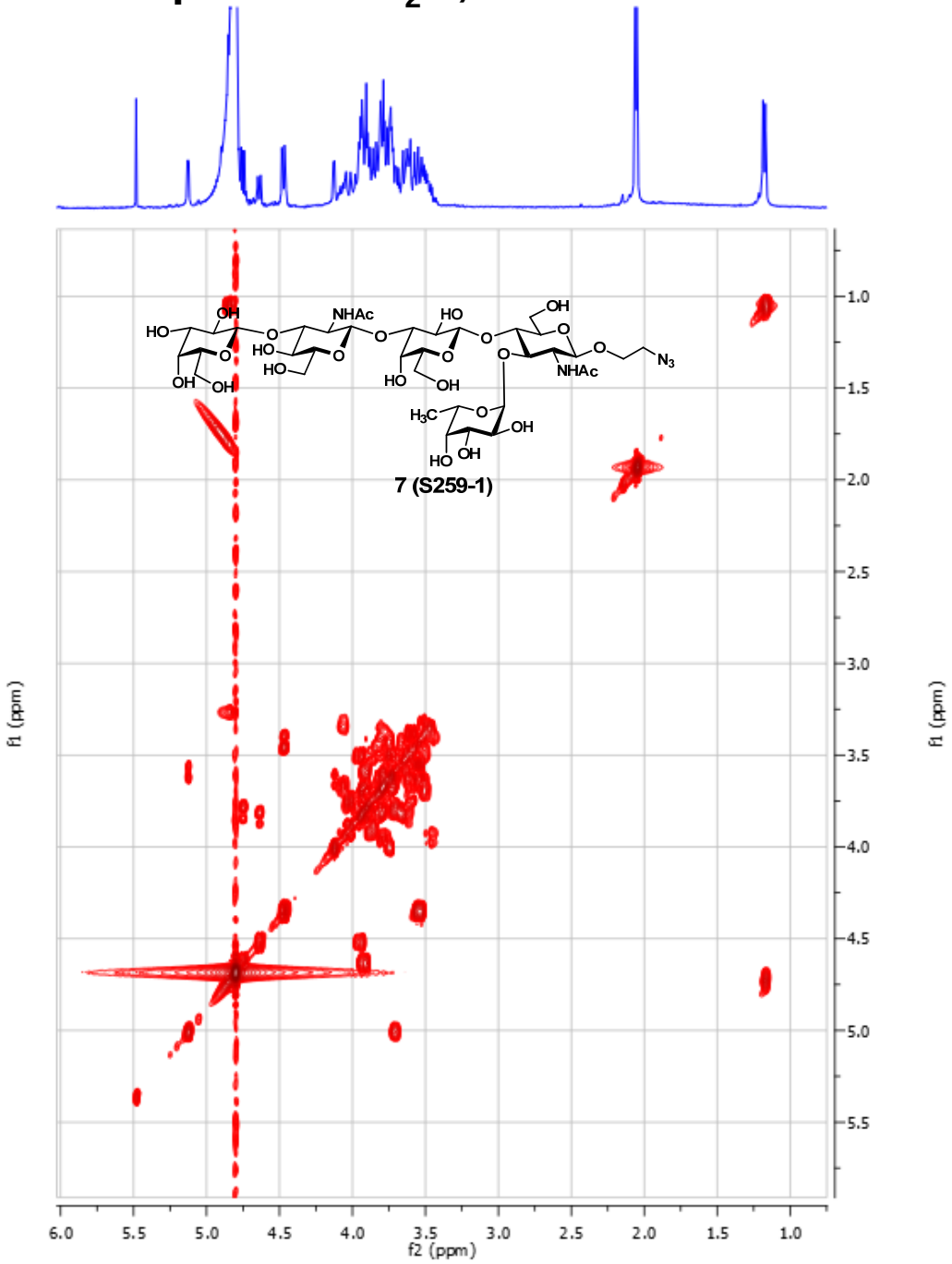
Comparison of ^{1}H NMR Spectra in D_2O, 400 MHz



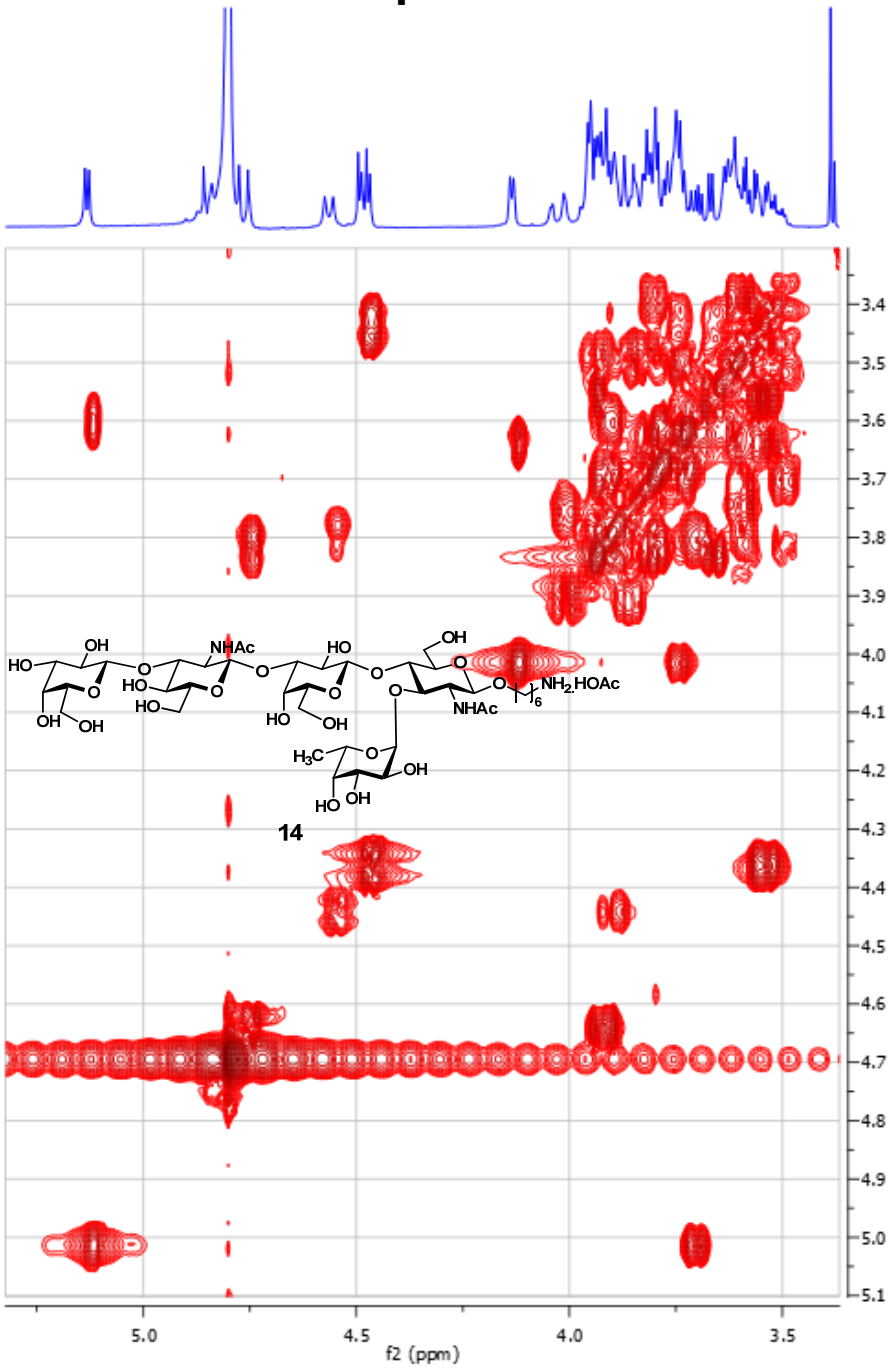


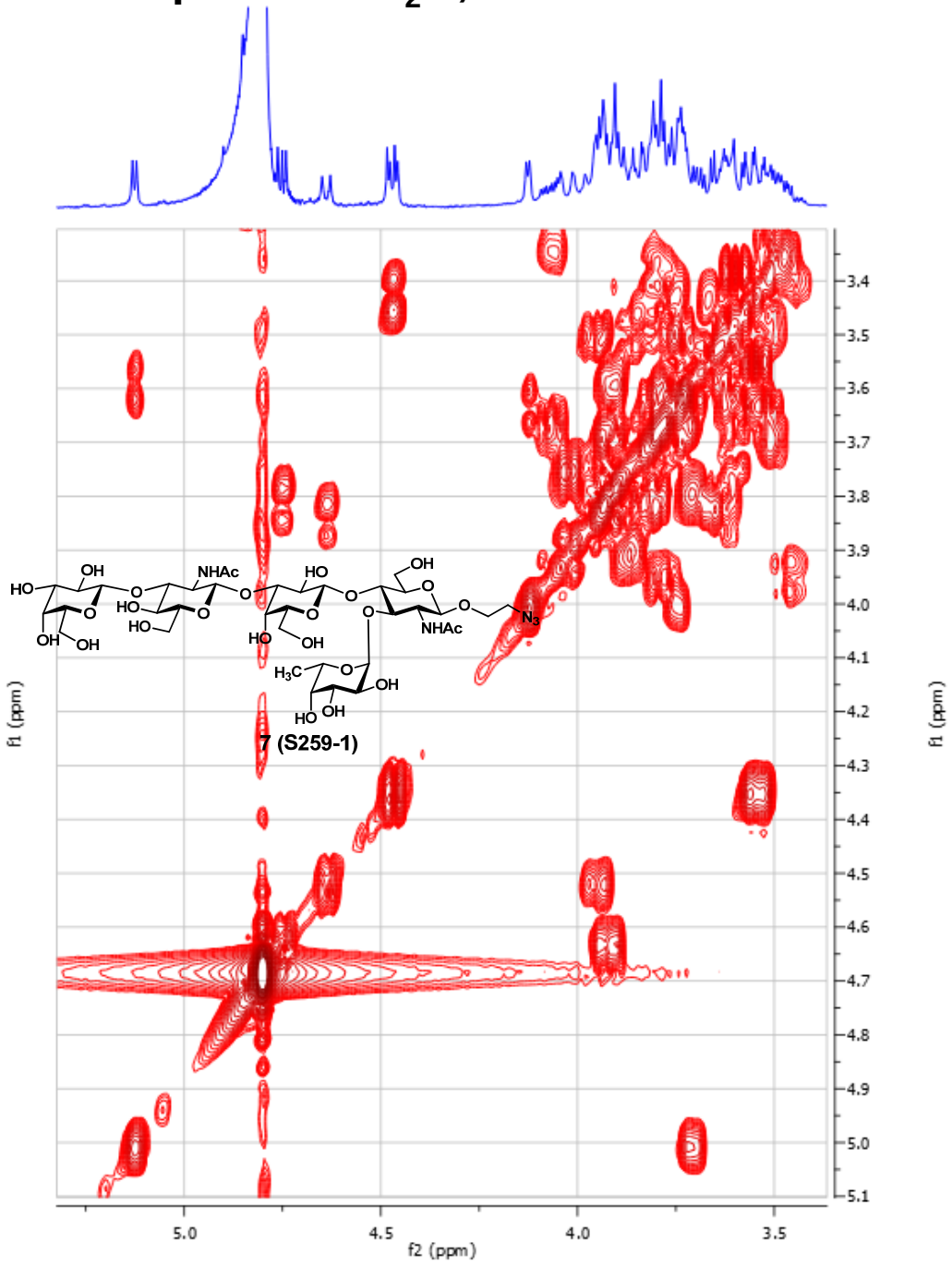
Comparison of ^{1}H-^{1}H GCOSY NMR Spectra in D_2O, 400 MHz



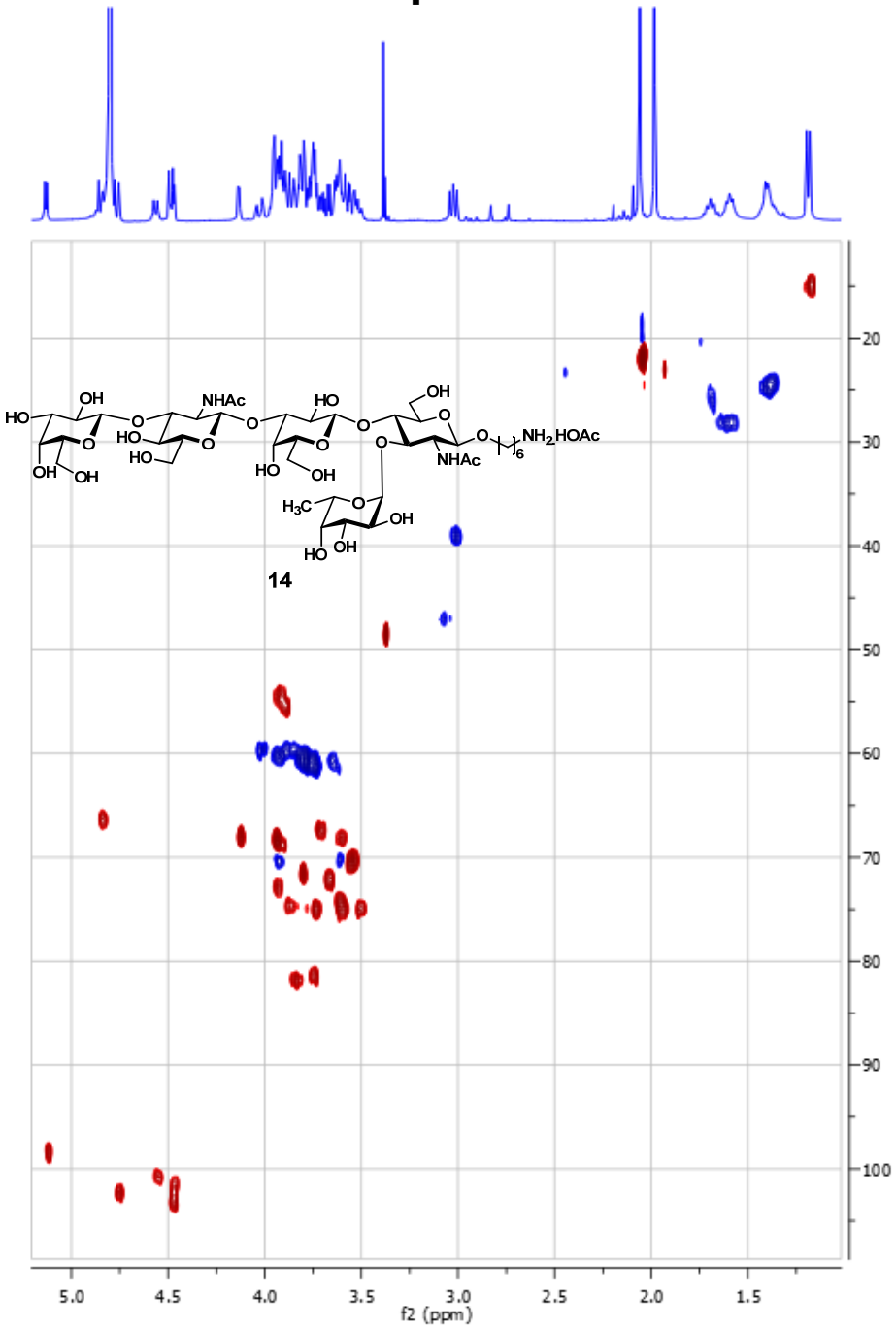


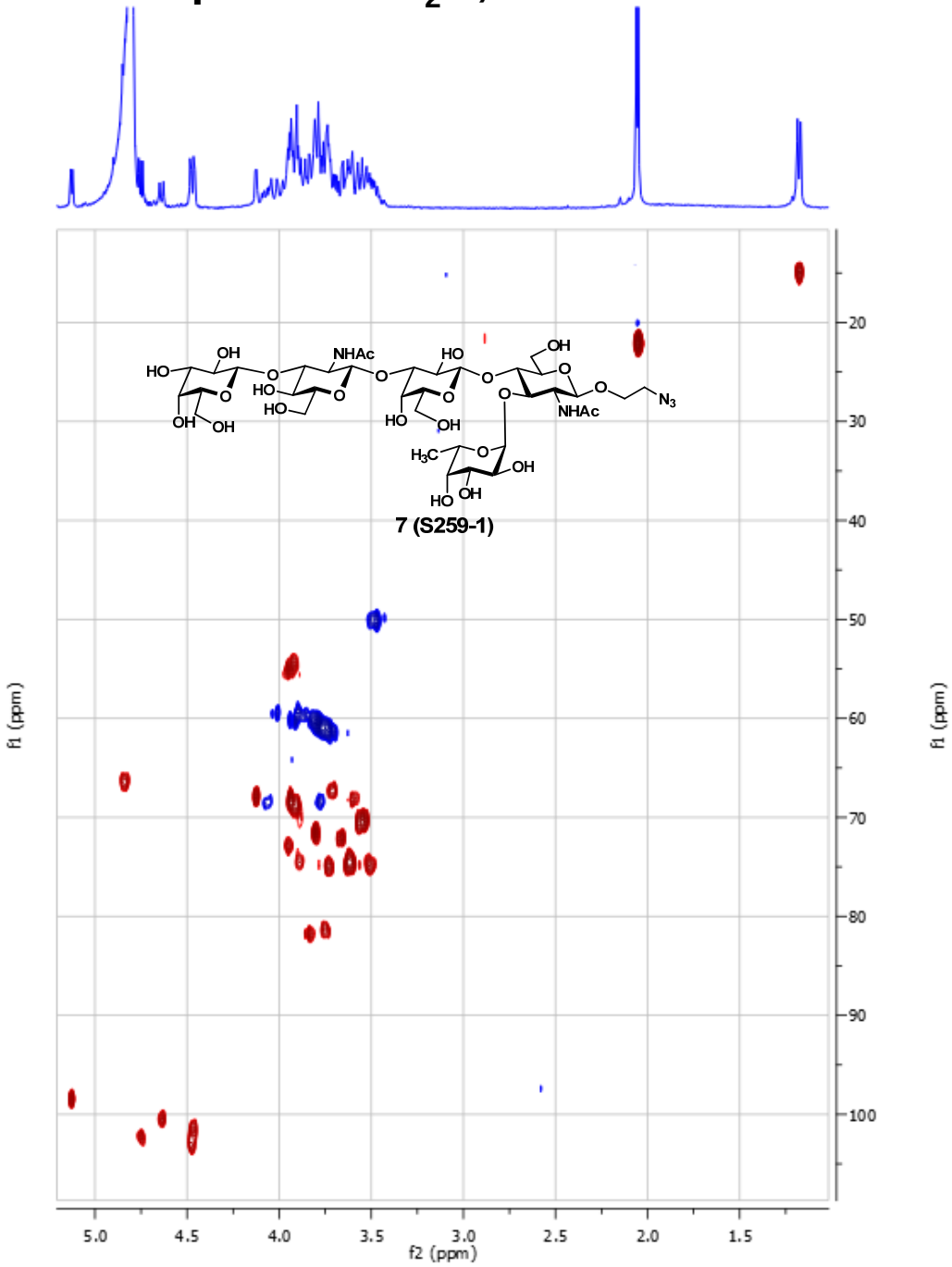
Comparison of ^{1}H-^{1}H GCOSY NMR Spectra in D_2O, 400 MHz





Comparison of ^{1}H-^{13}C GHSQC NMR Spectra in D_2O, 400 MHz





Comparison of ^{1}H-^{13}C GHSQC NMR Spectra in D_2O, 400 MHz