Electronic Supplementary Material (ESI) for RSC Medicinal Chemistry.
This journal is © The Royal Society of Chemistry 2021

Electronic Supplementary Information

Synthesis and immunological evaluation of *N*-acyl modified Globo H derivatives as anticancer vaccine candidates

Canjia Zhai[⊥], Xiu-Jing Zheng[⊥], Chengcheng Song and Xin-Shan Ye*

State Key Laboratory of Natural and Biomimetic Drugs, School of Pharmaceutical Sciences, Peking University, Xue Yuan Rd No. 38, Beijing 100191, China

*E-mail: xinshan@bjmu.edu.cn

Contents

1. Experimental section

General materials

Unless otherwise noted, all chemicals were purchased as reagent grade and used without further purification. Pyridine and dichloromethane (DCM) were distilled over calcium hydride (CaH_2). Et_2O was dried and distilled from sodium metal. MeOH was distilled from magnesium. *N*,*N*-Dimethylformamide (DMF) was dried with P_2O_5 and distilled under reduced pressure. Reactions were monitored through analytical thin-layer chromatography (TLC) on 0.25 mm silica gel 60-F_{254} precoated on aluminum plates (E. Merck), using UV light and/or staining with acidic ceric ammonium molybdate. Column chromatography was performed with 200-300 mesh silica gel or C18. Optical rotations were measured on a Rudolph Research Analytical Autopol IV automatic polarimeter. NMR (^{1}H, ^{13}C, HMBC) spectra were collected on Bruker AVANCE III-400 or III-600 spectrometer at 25 °C. Chemical shifts (ppm) were referenced to the solvent residual proton signal. High resolution mass spectrometry (HMRS) was performed on a Thermo Scientific LTQ Orbitrap Discovery or Waters Xevo G2 QTof spectrometer. Matrix-assisted laser desorption/ionization time of flight (MALDI-TOF) MS was performed on AB SCIEX 5800 spectrometer. For NMR spectra: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, dd = doublet of doublets, br. = broad.

General methods for the synthesis of compounds 4, 8, 10, 15α and 15β

The general procedure for glycosylation (synthesis of **4**, **8**, **10**, **15α** and **15β**) was performed according to the literature[1] with some modifications. Typically, Tf_2O (6.0 μL, 0.0357 mmol, 0.6 equiv.) was added via a micro-syringe to a stirred mixture of donor (0.0595 mmol, 1.0 equiv.), Ph_2SO (8.4 mg, 0.0417 mmol, 0.7 equiv.), and activated 4 Å molecular sieves (0.5 g, powder) in dry DCM or Et_2O (5 mL) at -72 °C (dry ice-ethanol bath in the Dewar flask) under argon atmosphere. After the donor disappeared as monitored by TLC (typically 15 min), a solution of acceptor (0.0595 mmol, 1.0 equiv.) in DCM or Et_2O (1 mL) was added dropwise. The dry ice was removed from the Dewar flask to allow the reaction mixture to warm to -20 °C slowly over about 2 h. Et_3N (0.2 mL) was then added to quench the reaction. The reaction solution was then filtered through celite and the filtrate was concentrated. The residue was purified through column chromatography on silica gel (petroleum ether/ethyl acetate, 4/1, v/v) to afford the corresponding product.

p-Tolyl 2-*O*-benzoyl-3,4,6-tri-*O*-benzyl-β-D-galactopyranosyl-(1→3)-2-azido-4,6-*O*-benzylidene-2-deoxy-1-thio-β-D-galactopyranoside (4):

Compound **4** was prepared following the general glycosylation procedure. White solid, 85% yield. ^{1}H NMR (400 MHz, $CDCl_3$): δ 2.24 (s, 3H), 3.32 (s, 1H), 3.54-3.66 (m, 6H), 3.84 (dd, 1H, J = 12.2 Hz, J = 1.1 Hz), 3.96 (d, 1H, J = 2.5 Hz), 4.24-4.28 (m, 2H), 4.30 (d, 1H, J = 9.1 Hz), 4.42-4.48 (m, 3H), 4.59-4.63 (m, 2H), 4.79 (d, 1H, J = 8.0 Hz), 4.98 (d, 1H, J = 11.7 Hz), 5.37 (s, 1H), 5.62 (dd, 1H, J = 9.8 Hz, J = 8.1 Hz), 6.87-6.89 (m, 2H), 7.11-7.20 (m, 5H), 7.27-7.41 (m, 17H), 7.47-7.49 (m, 2H), 7.52-7.56 (m, 1H), 7.99-8.01 (m, 2H). The ^{1}H NMR data coincide with those reported.[2]

p-Tolyl 3,4,6-tri-*O*-benzyl-β-D-galactopyranosyl-(1→3)-4,6-*O*-benzylidene-2-azido-2-deoxy-1-thio-β-D-galactopyranoside (5):

Compound **5** was prepared by the same procedure as described in the literature.[2] It was prepared from **4** (7.9 g, 8.44 mmol), affording **5** (5.8 g, 82% yield). ^{1}H NMR (400 MHz, $CDCl_3$): δ 2.31 (s, 3H), 2.46 (d, 1H, J = 2.4 Hz), 3.35 (s, 1H), 3.45 (dd, 1H, J = 9.7 Hz, J = 2.9 Hz), 3.50-3.61 (m, 4H), 3.78 (t, 1H, J = 10.0 Hz), 3.85 (d, 1H, J = 2.8 Hz), 3.88 (dd, 1H, J = 12.3 Hz, J = 1.3 Hz), 3.95 (dt, 1H, J = 9.8 Hz, J = 2.3 Hz), 4.21 (d, 1H, J = 2.9 Hz), 4.31 (dd, 1H, J = 12.3 Hz, J = 1.3 Hz), 4.34 (d, 1H, J = 9.8 Hz), 4.40-4.48 (m, 3H), 4.59 (d, 1H, J = 11.6 Hz), 4.71 (d, 1H, J = 12.0

Hz), 4.77 (d, 1H, *J* = 12.4 Hz), 4.90 (d, 1H, *J* = 11.6 Hz), 5.44 (s, 1H), 7.01-7.03 (m, 2H), 7.28-7.43 (m, 20H), 7.60-7.62 (m, 2H).

***p*-Tolyl 3,4,6-tri-*O*-benzyl-β-D-galactopyranosyl-(1→3)-4,6-*O*-benzylidene-2-amino-2-deoxy-1-thio-β-D-galactopyranoside (6):**

Compound **6** (84.4 mg, 86% yield) was synthesized according to the reported procedure[2] from compound **5** (101.7 mg, 0.123 mmol). ^{1}H NMR (400 MHz, $CDCl_3$): δ 2.30 (s, 4H), 3.25 (t, 1H, *J* = 9.8 Hz), 3.34-3.40 (m, 2H), 3.51-3.59 (m, 4H), 3.85-3.89 (m, 2H), 3.95 (dd, 1H, *J* = 9.3 Hz, *J* = 8.1 Hz), 4.16 (d, 1H, *J* = 2.3 Hz), 4.29 (d, 1H, *J* = 12.2 Hz), 4.35 (d, 1H, *J* = 9.5 Hz), 4.43-4.44 (m, 3H), 4.57 (d, 1H, *J* = 11.6 Hz), 4.64-4.71 (m, 2H), 4.86 (d, 1H, *J* = 11.6 Hz), 5.43 (s, 1H), 7.00-7.02 (m, 2H), 7.24-7.45 (m, 20H), 7.53-7.55 (m, 2H).

***p*-Tolyl 3,4,6-tri-*O*-benzyl-β-D-galactopyranosyl-(1→3)-4,6-*O*-benzylidene-2-(2,2,2-trichloroethoxycarbonylamino)-2-deoxy-1-thio-β-D-galactopyranoside (7):**

Compound **7** (63.0 mg, 64% yield) was synthesized according to the reported procedure[2] from compound **6** (81.5 mg, 0.101 mmol). ^{1}H NMR (400 MHz, $CDCl_3$): δ 2.32 (s, 3H), 2.48 (br. s, 1H), 3.31 (dd, 1H, *J* = 9.8 Hz, *J* = 2.4 Hz), 3.44 (s, 1H), 3.49-3.65 (m, 4H), 3.82 (d, 1H, *J* = 2.6 Hz), 3.87-3.91 (m, 2H), 4.24-4.33 (m, 4H), 4.44 (s, 1H), 4.57 (d, 1H, *J* = 11.6 Hz), 4.63-4.67 (m, 3H), 4.76 (d, 1H, *J* = 12.0 Hz), 4.87 (d, 1H, *J* = 11.6 Hz), 5.09 (d, 1H, *J* = 10.0 Hz), 5.31 (d, 1H, *J* = 7.2 Hz), 5.46 (s, 1H), 7.02-7.04 (m, 2H), 7.25-7.44 (m, 20H), 7.53-7.55 (m, 2H).

***p*-Tolyl 2,3,4-tri-*O*-benzyl-α-L-fucopyranosyl-(1→2)-3,4,6-tri-*O*-benzyl-β-D-galactopyranosyl-(1→3)-4,6-*O*-benzylidene-2-(2,2,2-trichloroethoxycarbonylamino)-2-deoxy-1-thio-β-D-galactopyranoside (8):**

Compound **8** was prepared following the general glycosylation procedure. White foam, 82% yield. $[\alpha]_D^{24}$ -60.3 (*c* 0.4, CH_2Cl_2); ^{1}H NMR (400 MHz, $CDCl_3$): δ 0.77 (d, 3H, *J* = 6.3 Hz), 2.27 (s, 3H), 3.22 (s, 1H), 3.38 (s, 1H), 3.50-3.53 (m, 3H), 3.56-3.66 (m, 2H), 3.81-3.84 (m, 2H), 3.90-3.94 (m, 2H), 4.10-4.19 (m, 3H), 4.28 (d, 1H, *J* = 12.4 Hz), 4.32 (dd, 1H, *J* = 10.6 Hz, *J* = 2.7 Hz), 4.40-4.51 (m, 6H),4.54-4.60 (m, 3H), 4.65 (d, 1H, *J* = 12.3 Hz), 4.69 (d, 2H, *J* = 11.3 Hz), 4.80 (m, 3H), 5.04 (d, 1H, *J* = 9.7 Hz), 5.45 (s, 1H), 5.58 (d, 1H, *J* = 3.5 Hz), 5.91 (d, 1H, *J* = 7.6 Hz), 7.01-7.03 (m, 2H), 7.06-7.15 (m, 5H), 7.21-7.36 (m, 28H), 7.40-7.42 (m, 2H), 7.49-7.51 (m, 2H); ^{13}C NMR (100 MHz, $CDCl_3$): δ 16.35, 21.34, 52.76, 66.86, 68.97, 69.35, 70.16, 72.19, 72.73, 72.82, 72.90, 73.39, 73.69, 73.87, 74.10, 74.51, 74.62, 75.01, 75.92, 76.19, 78.39, 79.58, 83.88, 84.99, 95.86, 97.13, 101.12, 102.39, 126.73, 126.87, 127.18, 127.33, 127.45, 127.49, 127.61, 127.84, 127.97, 128.05, 128.08, 128.15, 128.25, 128.37, 128.42, 128.48, 128.54, 128.64, 129.10, 129.71, 134.12, 137.94, 138.15, 138.19 (×2), 138.38, 138.71, 139.08, 139.23, 153.93; HRMS (ESI$^+$) Anal. Calcd for $C_{77}H_{84}Cl_3N_2O_{15}S$ $[M+NH_4]^+$ 1413.4653, found 1413.4680.

***p*-Tolyl 2,3,4-tri-*O*-benzyl-α-L-fucopyranosyl-(1→2)-3,4,6-tri-*O*-benzyl-β-D-galactopyranosyl-(1→3)-4,6-*O*-benzylidene-2-(2,2,2-trichloroethoxycarbonylamino)-2-deoxy-β-D-galactopyranosyl-(1→3)-2,4,6-tri-*O*-benzyl-1-thio-β-D-galactopyranoside (10):**

Compound **10** was prepared following the general glycosylation procedure. Light yellow foam, 80% yield. $[\alpha]_D^{24}$ -46.7 (*c* 0.1, CH_2Cl_2); ^{1}H NMR (400 MHz, $CDCl_3$): δ 0.73 (d, 3H, *J* = 6.3 Hz), 2.28 (s, 3H), 3.30 (s, 1H), 3.44-3.68 (m, 8H), 3.74 (d, 1H, *J* = 12.3 Hz), 3.83-3.88 (m, 6H), 3.93 (dd, 1H, *J* = 10.3 Hz, *J* = 3.2 Hz), 3.99 (d, 1H, *J* = 11.9 Hz), 4.14-4.19 (m, 4H), 4.23 (d, 1H, *J* = 12.1 Hz), 4.35-4.69 (m, 17H), 4.77-4.87 (m, 3H), 5.03 (d, 1H, *J* = 6.1 Hz), 5.16 (d, 1H, *J* = 11.9 Hz), 5.39 (br.s, 1H), 5.51 (s, 1H), 5.61 (d, 1H, *J* = 2.9 Hz), 6.96-6.98 (m, 2H), 7.07-7.12 (m, 5H), 7.16-7.35 (m, 41H), 7.39-7.44 (m, 6H); ^{13}C NMR (100 MHz, $CDCl_3$): δ 16.41, 21.24, 54.74, 66.63, 67.05, 68.83, 69.13, 69.31, 72.26, 72.58, 72.65, 72.89, 73.61, 73.66, 73.78, 73.98, 74.65, 74.71, 74.91, 75.68, 76.19, 76.37, 77.69, 77.90, 78.08, 79.43, 83.50, 83.65, 88.14, 95.82, 97.14, 101.36, 101.74, 103.15, 126.72, 126.78, 127.03, 127.17, 127.20, 127.28, 127.46, 127.65, 127.79, 127.83, 127.99, 128.10, 128.21, 128.30, 128.33, 128.37, 128.47, 128.50, 128.54, 128.65, 129.07, 129.74, 130.42, 131.99, 137.29, 137.94, 138.17 (×2), 138.27, 138.46, 138.73, 139.04, 139.26, 139.37 (×2), 154.06; HRMS (ESI$^+$) Anal. Calcd for $C_{104}H_{112}Cl_3N_2O_{20}S$ $[M+NH_4]^+$ 1845.6589, found 1845.6575.

Benzyl (5-pentyl) carbamate 4,6-*O*-benzylidene-β-D-galactopyranosyl-(1→4)-β-D-glucopyranoside (12):

Compound **12** was prepared as white solids according to the reported procedure[3] in 85% yield. $[\alpha]_D^{24}$ -11.1 (*c* 0.2, MeOH); ^{1}H NMR (400 MHz, CD_3OD): δ 1.37-1.67 (m, 6H), 3.11 (t, 2H, *J* = 6.8 Hz), 3.24 (t, 1H, *J* = 8.3 Hz), 3.40 (m, 1H), 3.52-3.71 (m, 6H), 3.84-3.94 (m, 3H), 4.14-4.22 (m, 3H), 4.28 (d, 1H, *J* = 7.8 Hz), 4.48 (d, 1H, *J* = 7.0 Hz), 5.06 (s, 2H), 5.63 (s, 1H), 7.28-7.38 (m, 8H), 7.53-7.55 (m, 2H); ^{13}C NMR (100 MHz, CD_3OD): δ 24.24, 30.35, 30.59, 41.71, 61.69, 67.29, 68.31, 70.18, 70.69, 71.75, 73.49, 74.82, 76.32, 76.43, 77.35, 79.98, 102.27, 104.25, 104.82, 127.48, 128.76, 128.93, 129.01, 129.45, 129.88, 138.50, 139.52, 158.93; HRMS (ESI^+) Anal. Calcd for $C_{32}H_{47}N_2O_{13}$ $[M+NH_4]^+$ 667.3073, found 667.3079.

Benzyl benzyl (5-pentyl) carbamate 2,3-di-*O*-benzyl-4,6-*O*-benzylidene-β-D-galactopyranosyl-(1→4)-2,3,6-tri-*O*-benzyl-β-D-glucopyranoside (13):

Compound **12** (166.4 mg, 0.256 mmol) and TBAI (18.9 mg, 0.0512 mmol) were dissolved in anhydrous DMF (2 mL). NaH (62.0 mg, 1.536 mmol, 60% in mineral oil) was added slowly into the solution under ice bath. BnBr (0.3 mL, 2.304 mmol) was then added into the mixture. The mixture was stirred at room temperature for 1 d. MeOH was added to quench the reaction. The solvent was then removed. The residue was diluted in EtOAc and washed with water and brine. The organic layer was dried over anhydrous Na_2SO_4 and evaporated. The residue was purified by column chromatography on silica gel (petroleum ether/ethyl acetate, 5/1→4/1, v/v) to give product **13** as colorless oil (224.1 mg, 73% yield). $[\alpha]_D^{24}$ +8.6 (*c* 0.3, CH_2Cl_2); ^{1}H NMR (400 MHz, $CDCl_3$): δ 1.28-1.57 (m, 6H), 2.92 (s, 1H), 3.15-3.22 (m, 2H), 3.33-3.45 (m, 4H), 3.61 (t, 1H, *J* = 9.1 Hz), 3.69-3.77 (m, 2H), 3.81-3.88 (m, 3H), 3.95 (t, 1H, *J* = 9.1 Hz), 4.01 (s, 1H), 4.19 (d, 1H, *J* = 12.4 Hz), 4.29-4.36 (m, 2H), 4.44-4.47 (m, 3H), 4.53 (d, 1H, *J* = 12.2 Hz), 4.68-4.78 (m, 5H), 4.82-4.89 (m, 2H), 5.14-5.18 (m, 3H), 5.45 (s, 1H), 7.16-7.25 (m, 15H), 7.27-7.38 (m, 21H), 7.44-7.46 (m, 2H), 7.51-7.52 (m, 2H); ^{13}C NMR (100 MHz, $CDCl_3$): δ 23.41, 27.58, 27.99, 29.73, 46.21, 47.20, 50.26, 50.54, 66.36, 67.16, 68.35, 68.98, 69.76, 69.83, 71.65, 72.99, 73.69, 74.96, 75.12, 75.30, 75.79, 77.69, 78.84, 79.68, 81.85, 83.07, 101.39, 102.89, 103.64, 126.59, 127.20, 127.29, 127.40, 127.47, 127.54, 127.63, 127.74, 127.77, 127.86, 127.95, 128.12, 128.14, 128.22, 128.26, 128.30, 128.39, 128.48, 128.55, 128.60, 128.88, 136.81, 136.92, 137.95, 138.12, 138.44, 138.57, 138.74, 138.89, 139.00, 156.19, 156.73; HRMS (ESI^+) Anal. Calcd for $C_{74}H_{83}N_2O_{13}$ $[M+NH_4]^+$ 1207.5890, found 1207.5870.

Benzyl benzyl (5-pentyl) carbamate 2,3,6-tri-*O*-benzyl-β-D-galactopyranosyl-(1→4)-2,3,6-tri-*O*-benzyl-β-D-glucopyranoside (14):

Compound **13** (8.6 g, 7.19 mmol) and $NaBH_3CN$ (1.47 g, 23.3 mmol) were dissolved in anhydrous THF (150 mL). Hydrochloric acid diethyl ether solution (2 M, 39 mL, 77.8 mmol) was added into the mixture under ice bath. The mixture was stirred at room temperature for 4.5 h. The solvent was then evaporated. The residue was diluted in DCM and washed with 10% HCl solution, water, and brine. The organic layer was dried over anhydrous Na_2SO_4 and evaporated. The residue was purified by column chromatography on silica gel (petroleum ether/ethyl acetate, 5/1→2/1, v/v) to give product **14** as colorless oil (7.0 g, 82%). $[\alpha]_D^{24}$ +13.1 (*c* 0.3, CH_2Cl_2); ^{1}H NMR (400 MHz, $CDCl_3$): δ 1.34-1.63 (m, 6H), 2.51 (s, 1H), 3.15 (m, 1H), 3.23 (m, 1H), 3.29-4.00 (m, 14H), 4.32-4.47 (m, 7H), 4.54 (d, 1H, *J* = 12.1 Hz), 4.61-4.80 (m, 6H), 4.86 (m, 1H), 4.99 (d, 1H, *J* = 10.8 Hz), 5.15 (d, 2H, *J* = 9.6 Hz), 7.12-7.40 (m, 40H); ^{13}C NMR (100 MHz, $CDCl_3$): δ 23.35, 27.51, 27.92, 29.40, 46.17, 47.14, 50.21, 50.51, 66.10, 67.11, 68.27, 68.46, 69.96, 71.93, 72.80, 73.08, 73.46, 74.85, 75.08, 75.20, 75.31, 76.60, 79.38, 81.09, 81.76, 82.88, 102.52, 103.57, 127.14, 127.21, 127.45, 127.50, 127.57, 127.59, 127.62, 127.73, 127.76, 127.81, 127.89, 128.03, 128.08, 128.24, 128.34, 128.43, 128.50, 136.77, 136.84, 137.93, 138.21, 138.31, 138.64, 138.68, 139.13, 156.12, 156.66; HRMS (ESI^+) Anal. Calcd for $C_{74}H_{85}N_2O_{13}$ $[M+NH_4]^+$ 1209.6046, found 1209.6041.

Benzyl benzyl (5-pentyl) carbamate 2,3,4-tri-*O*-benzyl-α-L-fucopyranosyl-(1→2)-3,4,6-tri-*O*-benzyl-β-D-galactopyranosyl-(1→3)-4,6-*O*-benzylidene-2-(2,2,2-trichloroethoxycarbonylamino)-2-deoxy-β-D-galactopyranosyl-(1→3)-2,4,6-tri-*O*-benzyl-α-

D-galactopyranosyl-(1→4)-2,3,6-tri-*O*-benzyl-β-D-galactopyranosyl-(1→4)-2,3,6-tri-*O*-benzyl-β-D-glucopyranoside (15α):

Compound **15α** was prepared following the general glycosylation procedure. Colorless oil, 73% yield. $[\alpha]_D^{24}$ -240.0 (*c* 0.01, CH_2Cl_2); ^{1}H NMR (600 MHz, $CDCl_3$): δ 0.43 (d, 3H, *J* = 6.1 Hz), 1.28-1.60 (m, 6H), 2.79 (s, 1H), 3.13 (m, 1H), 3.18-3.24 (m, 2H), 3.26-3.32 (m, 4H), 3.39-3.52 (m, 6H), 3.54-3.59 (m, 3H), 3.62-3.65 (m, 2H), 3.68 (d, 1H, *J* = 10.3 Hz), 3.77 (dd, 1H, *J* = 10.9 Hz, *J* = 4.3 Hz), 3.81-3.87 (m, 4H), 3.92-3.96 (m, 2H), 3.99 (dd, 2H, *J* = 9.6 Hz, *J* = 4.3 Hz), 4.02 (dd, 1H, *J* = 10.1 Hz, *J* = 3.2 Hz), 4.07-4.14 (m, 7H), 4.17-4.25 (m, 4H), 4.27 (m, 1H), 4.31-4.37 (m, 5H), 4.40-4.51 (m, 12H), 4.56-4.58 (m, 2H), 4.61-4.86 (m, 14H), 4.95 (d, 1H, *J* = 3.2 Hz), 5.04 (d, 1H, *J* = 11.2 Hz), 5.15 (d, 2H, *J* = 10.1 Hz), 5.22 (d, 1H, *J* = 11.4 Hz), 5.42 (s, 1H), 5.54 (d, 1H, *J* = 3.5 Hz), 7.00-7.06 (m, 6H), 7.08-7.14 (m, 10H), 7.15-7.34 (m, 68H), 7.36-7.37 (m, 4H), 7.43-7.44(m, 2H); ^{13}C NMR (150 MHz, $CDCl_3$): δ 16.08, 23.52, 27.67, 28.09, 29.58, 46.33, 47.31, 50.37, 50.68, 54.36, 66.21, 66.66, 67.27, 67.75, 68.09, 68.45, 68.91, 69.16, 69.38, 69.87, 71.99, 72.12, 72.49, 72.65, 72.91, 72.95, 73.09, 73.16, 73.20, 73.28, 73.49, 73.60, 73.73, 73.90, 74.54, 74.74, 74.90, 74.93, 75.16, 75.20, 75.38, 75.41, 76.20, 76.40, 77.57, 78.62, 79.16, 79.65, 80.21, 81.55, 81.68, 82.01, 84.08, 95.72, 97.45, 100.78, 101.76, 102.51, 102.71, 102.78, 103.65, 126.62, 126.98, 127.00, 127.18, 127.21, 127.25, 127.31, 127.36, 127.47, 127.55, 127.59, 127.61, 127.63, 127.71, 127.78, 127.86, 127.97, 128.05, 128.11, 128.14, 128.28, 128.31, 128.36, 128.38, 128.39, 128.44, 128.50, 128.57, 128.59, 128.66, 128.70, 129.04, 129.75, 136.93, 137.03, 137.91, 138.08, 138.15, 138.25, 138.29, 138.44, 138.47, 138.50, 138.63, 138.91, 138.93, 138.96, 139.03, 139.39, 139.66, 139.81, 153.91, 156.28, 156.84; HRMS (ESI^+) Anal. Calcd for $C_{171}H_{189}Cl_3N_4O_{33}$ $[M+2NH_4]^{2+}$ 1465.6144, found 1465.6103; HMBC: $^1J_{C1'''',\,H1''''}$ = 172.3 Hz, $^1J_{C1'',\,H1''}$ = 170.3 Hz, other four $^1J_{C1,\,H1}$ = 161.2, 161.2, 161.2, 161.7 Hz.

Benzyl benzyl (5-pentyl) carbamate 2,3,4-tri-*O*-benzyl-α-L- fucopyranosyl-(1→2)-3,4,6-tri-*O*-benzyl-β-D-galactopyranosyl-(1→3)-4,6-*O*-benzylidene-2-(2,2,2-trichloroethoxycarbonylamino)-2-deoxy-β-D-galactopyranosyl-(1→3)-2,4,6-tri-*O*-benzyl-β-D galactopyranosyl-(1→4)-2,3,6-tri-*O*-benzyl-β-D-galactopyranosyl-(1→4)-2,3,6-tri-*O*-benzyl-β-D-glucopyranoside (15β):

Compound **15β** was prepared following the general glycosylation procedure. Colorless oil, 20%. $[\alpha]_D^{24}$ -80.0 (*c* 0.02, CH_2Cl_2); ^{1}H NMR (600 MHz, $CDCl_3$): δ 0.62 (d, 3H, *J* = 6.0 Hz), 1.28-1.54 (m, 6H), 3.15 (m, 1H), 3.23-3.26 (m, 3H), 3.29 (t, 1H, *J* = 8.3 Hz), 3.34-3.54 (m, 13H), 3.60-3.78 (m, 9H), 3.86-3.92 (m, 4H), 3.98 (d, 1H, *J* = 11.3 Hz), 4.10-4.37 (m, 17H), 4.40-4.52 (m, 11H), 4.55-4.60 (m, 4H), 4.63-4.70 (m, 5H), 4.76-4.83 (m, 3H), 4.86 (d, 1H, *J* = 9.5 Hz), 4.94 (m, 2H), 5.09-5.16 (m, 4H), 5.22 (d, 1H, *J* = 11.6 Hz), 5.49 (s, 1H), 5.55 (d, 1H, *J* = 2.9 Hz), 6.94 (t, 1H, *J* = 7.3 Hz), 7.06-7.35 (m, 83H), 7.41-7.42 (m, 6H); ^{13}C NMR (150 MHz, $CDCl_3$): δ 16.28, 23.51, 27.67, 28.09, 29.55, 46.35, 47.31, 50.37, 50.68, 54.65, 66.56, 66.95, 67.27, 68.30, 68.75, 68.95, 69.14, 69.18, 69.76, 69.99, 72.29, 72.60, 72.69, 73.01, 73.23, 73.57, 73.61, 73.69, 73.88, 74.66, 74.91, 75.03, 75.15, 75.32, 75.73, 76.09, 76.28, 76.58, 76.67, 78.38, 79.27, 79.87, 80.51, 81.30, 81.81, 82.53, 83.13, 83.66, 95.89, 97.32, 101.49, 101.81, 102.74, 102.96 (×2), 103.70, 126.70, 126.87, 127.04, 127.16, 127.26, 127.29, 127.38, 127.40, 127.45, 127.55, 127.58, 127.60, 127.65, 127.70, 127.79, 127.82, 127.90, 127.96, 128.01, 128.17, 128.24, 128.29, 128.34, 128.38, 128.40, 128.42, 128.49, 128.55, 128.62, 128.65, 129.07, 136.92, 138.00, 138.06, 138.18, 138.24, 138.35, 138.46, 138.51, 138.78, 138.89, 138.93, 139.00, 139.18, 139.44, 139.67, 139.93, 153.99, 156.30, 156.84; HRMS (ESI) Anal. Calcd for $C_{171}H_{189}Cl_3N_4O_{33}$ $[M+2NH_4]^{2+}$ 1465.6144, found 1465.6160; HMBC: $^1J_{C1'''',\,H1''''}$ = 169.9 Hz, other five $^1J_{C1,\,H1}$ = 161.4, 162.4, 160.2, 160.2, 163.2 Hz.

5-Aminopentyl α-L-fucopyranosyl-(1→2)-β-D-galactopyranosyl-(1→3)-2-acetamido-2-deoxy-β-D-galactopyranosyl-(1→3)-α-D-galactopyranosyl-(1→4)-β-D-galactopyranosyl-(1→4)-β-D-glucopyranoside (G0):

A mixture of **15α** (29.0 mg, 0.01 mmol) and 1 M NaOH (0.2 mL, 0.2 mmol) in THF (1.5 mL) was stirred at 50 °C overnight. The solvent was then removed. The residue was diluted in DCM, and the organic phase was then washed with water and brine, and then dried over Na_2SO_4, filtered, and concentrated. The resulting residue was then dissolved in pyridine (0.7 mL), a catalytic amount of

DMAP was added. Acetic anhydride (0.18 mL, 1.9 mmol) was added slowly to the solution under ice bath. The mixture was stirred at room temperature overnight. After which, a few drops of water were added to the mixture to quench the reaction. The solution was then diluted with EtOAc. The organic phase was washed with saturated $NaHCO_3$ solution and brine, and dried over Na_2SO_4, filtered, and concentrated. The resulting residue was purified by column chromatography on silica gel (petroleum ether/ethyl acetate, 3/1, v/v) to afford the *N*-acetyl product. The *N*-acetyl product was dissolved in a mixed solvent of MeOH/HOAc/H_2O (3 mL/1 mL/1 mL). $Pd(OH)_2$ (58 mg, 20% on carbon) was then added. The mixture was stirred under hydrogen atmosphere for 2 d, and then filtered through celite and concentrated. The residue was purified by C18 reversed-phase column chromatography ($H_2O \rightarrow H_2O$/MeOH, 8/1, v/v) to give product **G0** as white solids (7.0 mg, 64% yield). ^{1}H NMR (600 MHz, D_2O): δ 1.24 (d, 3H, J = 6.6 Hz), 1.48 (m, 2H), 1.71 (m, 4H), 2.07 (s, 3H), 3.03 (t, 2H, J = 7.4 Hz), 3.33 (t, 1H, J = 8.6 Hz), 3.60-4.04 (m, 31H), 4.06 (d, 1H, J = 3.1 Hz), 4.13 (d, 1H, J = 2.2 Hz), 4.24-4.27 (m, 2H), 4.41 (t, 1H, J = 6.5 Hz), 4.51 (d, 1H, J = 8.0 Hz), 4.54 (d, 1H, J = 7.8 Hz), 4.57 (d, 1H, J = 7.7 Hz), 4.64 (d, 1H, J = 7.7 Hz), 4.92 (d, 1H, J = 4.0 Hz), 5.26 (d, 1H, J = 4.0 Hz). The ^{1}H NMR data coincide with those reported previously.[4]

5-Aminopentyl α-L-fucopyranosyl-(1→2)-β-D-galactopyranosyl-(1→3)-2-fluoroacetamido-2-deoxy-β-D-galactopyranosyl-(1→3)-α-D-galactopyranosyl-(1→4)-β-D-galactopyranosyl-(1→4)-β-D-glucopyranoside (G1):

A mixture of **15α** (60.3 mg, 0.0208 mmol) and 1 M NaOH (0.42 mL, 0.42 mmol) in THF (5 mL) was stirred at 50 °C overnight. The solvent was then removed. The residue was diluted in DCM, and the organic phase was then washed with water and brine, and then dried over Na_2SO_4, filtered, and concentrated. The resulting residue was then dissolved in pyridine (5 mL), a catalytic amount of DMAP was added. Fluoroacetyl chloride (14.5 μL, 0.208 mmol) was added slowly to the solution under ice bath. The mixture was stirred at room temperature for 2 h. After which, a few drops of water were added to the mixture to quench the reaction. The solution was then diluted with EtOAc. The organic phase was washed with saturated $NaHCO_3$ solution and brine, and dried over Na_2SO_4, filtered, and concentrated. The residue was purified by column chromatography on silica gel (petroleum ether/ethyl acetate, 3/1, v/v) to afford the *N*-acetyl product. The *N*-acetyl product was dissolved in a mixed solvent of MeOH/HOAc/H_2O (6 mL/2 mL/2 mL). $Pd(OH)_2$ (70 mg, 20% on carbon) was then added. The mixture was stirred under hydrogen atmosphere for 3 d, and then filtered through celite and concentrated. The residue was purified by C18 reversed-phase column chromatography ($H_2O \rightarrow H_2O$/MeOH, 8/1, v/v) to give product **G1** as white solids (10.9 mg, 47% yield). $[\alpha]_D^{24}$ -43.6 (*c* 0.03, H_2O); ^{1}H NMR (600 MHz, D_2O): δ 1.26 (d, 3H, J = 6.6 Hz), 1.48 (m, 2H), 1.71 (m, 4H), 3.03 (t, 2H, J = 7.6 Hz), 3.33 (t, 1H, J = 8.5 Hz), 3.59-3.74 (m, 12H), 3.75-3.90 (m, 12H), 3.91-3.99 (m, 4H), 4.02 (dd, 1H, J = 12.4 Hz, J = 2.0 Hz), 4.06 (d, 1H, J = 3.1 Hz), 4.11-4.16 (m, 3H), 4.25-4.29 (m, 2H), 4.41 (t, 1H, J = 6.6 Hz), 4.51 (d, 1H, J = 8.0 Hz), 4.54 (d, 1H, J = 7.8 Hz), 4.64 (d, 2H, J = 7.7 Hz), 4.919 (dd, 1H, J = 37.9 Hz, J = 14.5 Hz), 4.922 (d, 1H, J = 4.1 Hz), 5.00 (dd, 1H, J = 37.9 Hz, J = 14.5 Hz), 5.26 (d, 1H, J = 4.1 Hz); ^{13}C NMR (150 MHz, D_2O): δ 16.08, 22.86, 27.19, 28.94, 40.13, 52.37, 60.85, 61.14, 61.71, 61.75, 62.28, 67.57, 68.48, 68.80, 69.18, 69.87, 69.90, 70.26, 70.86, 70.95, 71.65, 72.60, 72.90, 73.73, 74.33, 75.32, 75.50, 75.58, 75.86, 76.26, 76.54, 76.88, 77.93, 79.51, 79.62, 80.68 (d, J = 182.11 Hz), 100.06, 101.19, 102.62, 102.74, 104.10, 104.37, 172.03 (d, J = 18.37 Hz); HRMS (ESI^+) Anal. Calcd for $C_{43}H_{76}FN_2O_{30}$ $[M+H]^+$ 1119.4461, found 1119.4431; HMBC: $^1J_{C1'''',''''',H1''''}$ = 171.8 Hz, $^1J_{C1'',H1''}$ = 170.1 Hz, other four $^1J_{C1,H1}$ = 160.6, 162.1, 162.1, 159.0 Hz.

5-Aminopentyl α-L-fucopyranosyl-(1→2)-β-D-galactopyranosyl-(1→3)-2-difluoroacetamido-2-deoxy-β-D-galactopyranosyl-(1→3)-α-D-galactopyranosyl-(1→4)-β-D-galactopyranosyl-(1→4)-β-D-glucopyranoside (G2):

A mixture of **15α** (124.5 mg, 0.043 mmol) and 1 M NaOH (0.86 mL, 0.86 mmol) in THF (8.5 mL) was stirred at 50 °C overnight. The solvent was then removed. The residue was diluted in DCM, and the organic phase was then washed with water and brine, and then dried over Na_2SO_4, filtered, and concentrated. The resulting residue was dissolved in pyridine (5 mL), a catalytic amount of

DMAP was added. Difluoroacetic anhydride (101.3 μL, 0.86 mmol) was added slowly to the solution under ice bath. The mixture was stirred at room temperature for 3 h. After which, a few drops of water were added to the mixture to quench the reaction. The solution was then diluted with EtOAc. The organic phase was washed with saturated $NaHCO_3$ solution and brine, and dried over Na_2SO_4, filtered, and concentrated. The resulting residue was purified by column chromatography on silica gel (petroleum ether/ethyl acetate, 3/1, v/v) to afford the *N*-acetyl product. The *N*-acetyl product was dissolved in a mixed solvent of MeOH/HOAc/H_2O (6 mL/2 mL/2 mL). $Pd(OH)_2$ (120 mg, 20% on carbon) was then added. The mixture was stirred under hydrogen atmosphere for 3 d, and then filtered through celite and concentrated. The residue was purified by C18 reversed-phase column chromatography ($H_2O \rightarrow H_2O$/MeOH, 8/1, v/v) to give product **G2** as colorless gel-like solids (30.0 mg, 61% yield). $[\alpha]_D^{24}$ -293.3 (*c* 0.01, H_2O); ^{1}H NMR (600 MHz, D_2O): δ 1.26 (d, 3H, *J* = 6.6 Hz), 1.48 (m, 2H), 1.71 (m, 4H), 3.03 (t, 2H, *J* = 7.6 Hz), 3.33 (t, 1H, *J* = 8.6 Hz), 3.59-3.74 (m, 13H), 3.76-3.88 (m, 11H), 3.89-3.93 (m, 2H), 3.94-3.99 (m, 2H), 4.02 (dd, 1H, *J* = 12.4 Hz, *J* = 2.0 Hz), 4.06 (d, 1H, *J* = 3.1 Hz), 4.11-4.17 (m, 3H), 4.20 (dd, 1H, *J* = 13.3 Hz, *J* = 6.5 Hz), 4.28 (d, 1H, *J* = 2.6 Hz), 4.41 (t, 1H, *J* = 6.5 Hz), 4.51 (d, 1H, *J* = 8.0 Hz), 4.54 (d, 1H, *J* = 7.8 Hz), 4.64 (d, 1H, *J* = 7.7 Hz), 4.66 (d, 1H, *J* = 7.7 Hz), 4.91 (d, 1H, *J* = 4.0 Hz), 5.22 (d, 1H, *J* = 4.1 Hz), 6.22 (t, 1H, *J* = 53.6 Hz); ^{13}C NMR (150 MHz, D_2O): δ 16.12, 22.86, 27.19, 28.94, 40.13, 52.81, 60.85, 61.12, 61.67, 61.77, 62.28, 67.57, 68.50, 68.85, 69.19, 69.84, 69.86, 70.00, 70.86, 70.96, 71.65, 72.53, 72.89, 73.73, 74.23, 75.32, 75.53, 75.58, 75.92, 76.21, 76.24, 77.46, 77.96, 79.54, 79.61, 100.37, 101.19, 102.48, 102.74, 104.09, 104.18, 109.10 (t, *J* = 248.78 Hz), 166.40 (t, *J* = 25.56 Hz); HRMS (ESI$^+$) Anal. Calcd for $C_{43}H_{75}F_2N_2O_{30}$ $[M+H]^+$ 1137.4367, found 1137.4402; HMBC: $^1J_{C1'''', H1''''}$ = 171.8 Hz, $^1J_{C1'', H1''}$ = 169.3 Hz, other four $^1J_{C1, H1}$ = 162.5, 158.4, 162.4, 158.4 Hz.

5-Aminopentyl α-L-fucopyranosyl-(1→2)-β-D-galactopyranosyl-(1→3)-2-trifluoroacetamido-2-deoxy-β-D-galactopyranosyl-(1→3)-α-D-galactopyranosyl-(1→4)-β-D-galactopyranosyl-(1→4)-β-D-glucopyranoside (G3):

A mixture of **15α** (70.2 mg, 0.0242 mmol) and 1 M NaOH (0.48 mL, 0.48 mmol) in THF (5 mL) was stirred at 50 °C overnight. The solvent was then removed. The residue was diluted in DCM, and the organic phase was then washed with water and brine, and then dried over Na_2SO_4, filtered, and concentrated. The resulting residue was dissolved in pyridine (5 mL), a catalytic amount of DMAP was added. Trifluoroacetic anhydride (68.2 μL, 0.484 mmol) was added slowly to the solution under ice bath. The reaction mixture was stirred at room temperature for 5 h. After which, a few drops of water were added to the mixture to quench the reaction. The solution was then diluted with EtOAc. The organic phase was washed with saturated $NaHCO_3$ solution and brine, and dried over Na_2SO_4, filtered, and concentrated. The residue was purified by column chromatography on silica gel (petroleum ether/ethyl acetate, 3/1, v/v) to afford the *N*-acetyl product. The *N*-acetyl product was dissolved in a mixed solvent of MeOH/HOAc/H_2O (6 mL/2 mL/2 mL). $Pd(OH)_2$ (70 mg, 20% on carbon) was then added. The mixture was stirred under hydrogen atmosphere for 3 d, and then filtered through celite and concentrated. The residue was purified by C18 reversed-phase column chromatography ($H_2O \rightarrow H_2O$/MeOH, 8/1, v/v) to give product **G3** as colorless gel-like solids (18.5 mg, 66%). $[\alpha]_D^{24}$ -50.9 (*c* 0.04, H_2O); ^{1}H NMR (600 MHz, D_2O): δ 1.26 (d, 3H, *J* = 6.5 Hz), 1.48 (m, 2H), 1.71 (m, 4H), 3.03 (t, 2H, *J* = 7.3 Hz), 3.33 (t, 1H, *J* = 8.6 Hz), 3.59-3.70 (m, 8H), 3.71-3.74 (m, 3H), 3.76-3.86 (m, 11H), 3.88-3.92 (m, 3H), 3.94-3.97 (m, 1H), 3.99 (dd, 1H, *J* = 10.6 Hz, *J* = 3.1 Hz), 4.02 (dd, 1H, *J* = 12.3 Hz, *J* = 1.9 Hz), 4.06 (d, 1H, *J* = 3.1 Hz), 4.12-4.17 (m, 4H), 4.29 (d, 1H, *J* = 2.8 Hz), 4.41 (t, 1H, *J* = 6.5 Hz), 4.51 (d, 1H, *J* = 8.0 Hz), 4.54 (d, 1H, *J* = 7.8 Hz), 4.62 (d, 1H, *J* = 7.7 Hz), 4.69 (d, 1H, *J* = 7.7 Hz), 4.91 (d, 1H, *J* = 4.0 Hz), 5.22 (d, 1H, *J* = 4.0 Hz); ^{13}C NMR (150 MHz, D_2O): δ 16.09, 22.86, 27.18, 28.94, 40.13, 53.33, 60.86, 61.11, 61.64, 61.77, 67.57,68.47, 68.83, 69.14, 69.81, 69.86, 70.02, 70.86, 70.97, 71.66, 72.52, 72.91, 73.71, 74.19, 75.33, 75.54, 75.57, 75.85, 75.97, 76.23, 77.54, 78.06, 79.55, 79.70, 100.44, 101.27, 102.37, 102.73, 103.94, 104.12, 116.65 (q, *J* = 286.31 Hz), 160.39 (q, *J* = 37.32 Hz); HRMS (ESI$^+$) Anal. Calcd for $C_{43}H_{74}F_3N_2O_{30}$ $[M+H]^+$ 1155.4273,

found 1155.4273; HMBC: $^{1}J_{C1'''', H1''''}$ = 170.4 Hz, $^{1}J_{C1'', H1''}$ = 170.4 Hz, other four $^{1}J_{C1, H1}$ = 161.7, 161.7, 162.1, 162.1 Hz.

General procedure for the preparation of glycoconjugates

A solution of each of the Globo H derivatives (**G0**, **G1**, **G2**, **G3** (~ 5.00 mg, 4.6 μmol, 1 equiv.)) in DMF (150 μL) was added to a second solution containing homobifunctional linker **16** (17.0 mg, 43.8 μmol, 10 equiv.) and Et_3N (8.3 μL) in DMF (150 μL). Small amount of DMF (150 μL) was used to rinse the tube containing the Globo H derivatives for the transfer. The reaction was shaken at room temperature and monitored through TLC ($MeOH/NH_4OH/H_2O$, 4/2/1, v/v/v, reaction time 2-6 h). After completion, the solvent of each reaction was then removed through rotavoporation. Each of the reaction residual was then dissolved in H_2O and washed with DCM. After rotavaping the aqueous layer solution, the residual was then dissolved in 0.01 M PBS solution (960 μL, pH 7.6). A solution of protein CRM 197 or BSA (285 μL, 14 mg/mL) in 0.01 M PBS was then added to the reaction mixture. The reaction was shaken at room temperature overnight and the solution was centrifuged (4 °C, 12,000 g) for 5 min. The supernatant was then collected and dialyzed against PBS at 4 °C (molecular weight cutoff value 14,000 Da, 6 changes of PBS). The carbohydrate loading percentage of each glycoconjugate was determined by MALDI-TOF MS analysis (Table S1, ESI) and calculated according to the following equation, and the protein content was determined by BCA assay.[5] The glycoconjugate samples were stored at -80 °C for use.

$$\text{Loading (\%)} = \frac{M\,(glycoconjugate) - M\,(protein)}{M\,(glycoconjugate)} \times 100\%$$

Immunological studies

Compounds and reagents: Carrier protein CRM197 was purchased from Sinovac® (China). C34 was generously provided by Dr. Chung-Yi Wu (Academia Sinica, Taipei, Taiwan).

Immunization of mice: Group of six mice (female pathogen-free Balb/c, age six to eight weeks, from Department of Laboratory Animal Science, Peking University Health Science Center) were vaccinated intramuscularly four times at biweekly intervals with **G0**-CRM197 or its modified analogues-CRM197 (**G1**-CRM197, **G2**-CRM197, **G3**-CRM197) (each contains 2.0 μg of carbohydrate in PBS) and glycolipid adjuvant C34 (2.0 μg). All experiments were carried out according to the International Association for the Study of Pain ethical guidelines (Zimmermann, 1983) and approved by the Institutional Animal Care and Use Committee of Peking University. Mice were bled prior to the initial vaccination, 13 days after the 3rd vaccination, and 14 days after the 4th vaccination. Blood was clotted to obtain sera, which were stored at -80 °C.

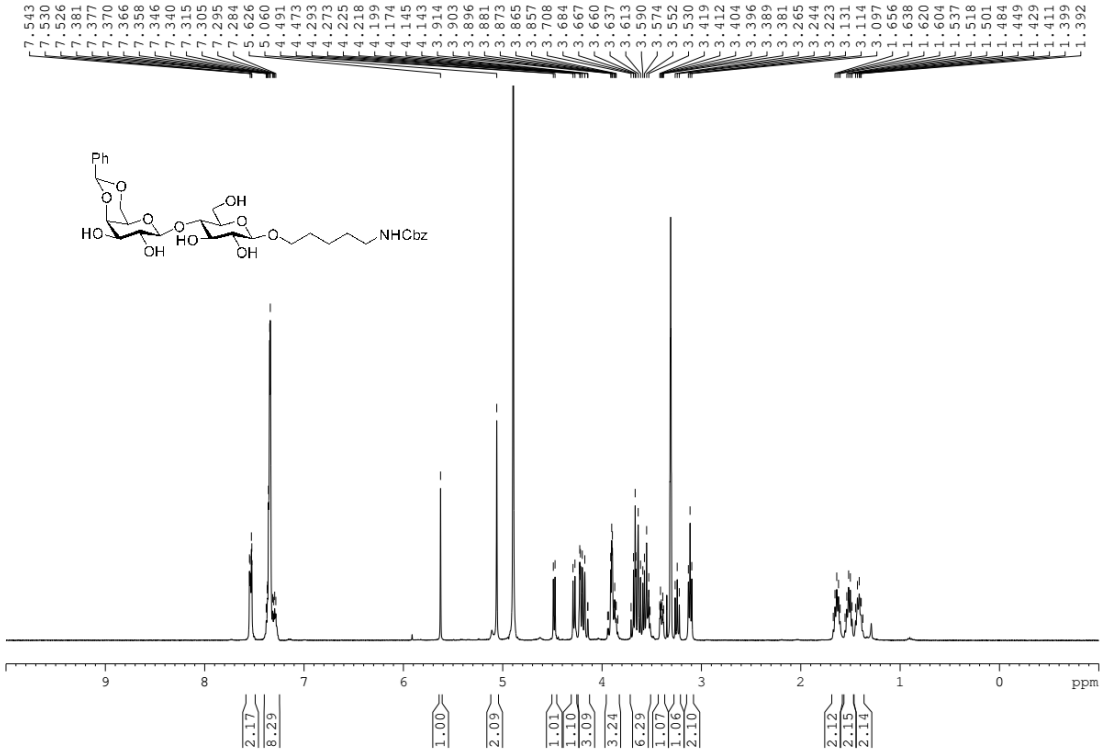
Serological assays: The antibody titers of the pooled sera and individual mouse were measured by ELISA. The plates were coated with **G0**-BSA (100 μL/well containing 0.02 μg of **G0**) or modified-**G0**-BSA analogues (**G1**-BSA, **G2**-BSA, **G3**-BSA) and incubated at 4 °C overnight (0.1 M bicarbonate buffer, pH = 9.6). After washing three times with PBST (0.05% Tween-20 in PBS), 3% BSA in PBS was added to block the microwells at 37 °C for 1 h. After washing three times, serially diluted sera were added to the microwells (100 μL/well) and then incubated at 37 °C for 1 h. The plate was then washed and incubated with a 1:5000 dilution of horseradish peroxidase conjugated goat anti-mouse IgG (γ-chain specific) (Southern Biotechnology Associates, Inc., Buckingham, AL) incubated at 37 °C for 1 h. After which, the plate was washed and developed with the *O*-phenylenediamine (OPD) substrate in the dark for 15 min, terminated with 2 M H_2SO_4, and then read at 490 nm. The antibody titer was defined as the highest dilution showing an absorbance of 0.1, after subtracting background.

Flow cytometry experiment: Globo H-expressing MCF-7 human breast cancer cells were obtained. Single-cell suspensions of 5×10^5 cells/tube were washed with 3% fetal calf serum (FCS) in PBS and incubated with 25 μL of 1:20 diluted test sera (sera from preimmunization and 14 days after the fourth vaccination) for 30 min on ice. After two washes with 3% FCS in PBS, 25 μL of 1:15-diluted goat anti-mouse IgG (γ-chain specific) labeled with FITC (Southern Biotechnology Associates, Inc., Birmingham, AL) was added, and the mixture was incubated for 30 min on ice. After a final wash, fixed with 1% formaldehyde and assayed using a FAC Scan (Becton Dickinson).

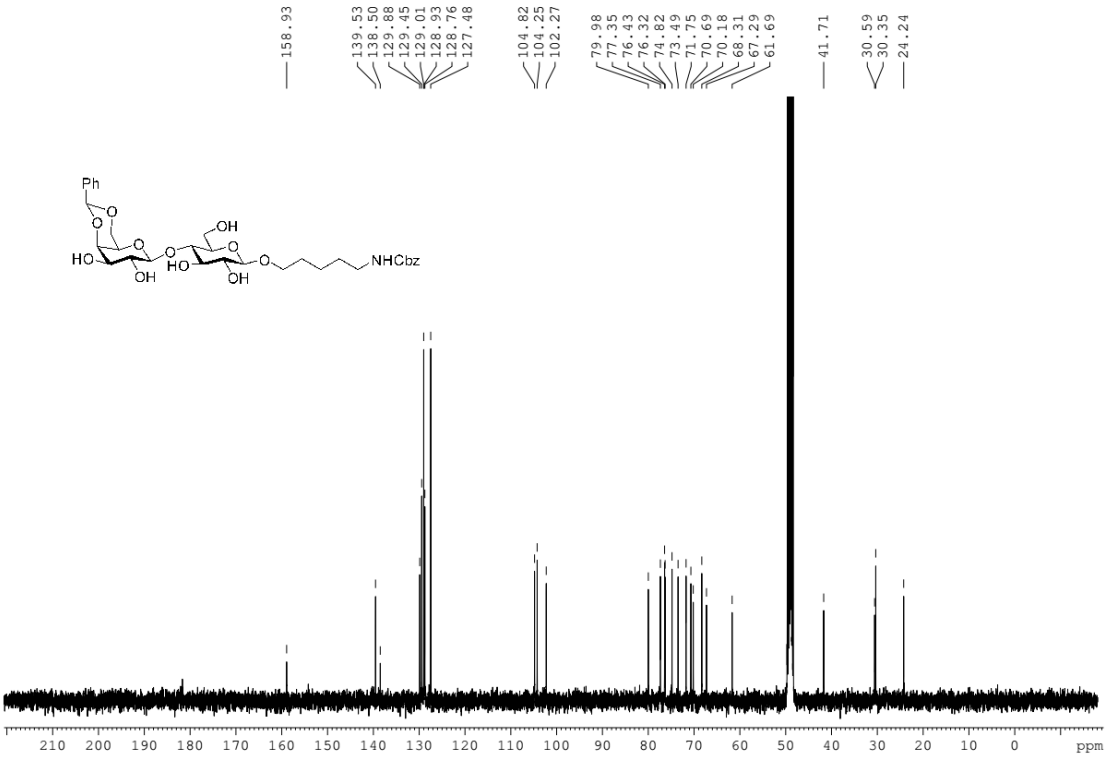
Complement-dependent cytotoxicity (CDC) assay: CDC was examined on Globo H antigen-positive MCF-7 cancer cells with a non-radioactive cytotoxicity assay kit (Promega). MCF-7 cells were seeded into 96-well round-bottom microtiter plates (Corning) at 5×10^3 cells/well. Diluted pooled immunological sera from each group were added and incubated at 37 °C for 1 h (the final concentration of sera was 1:10). After washing cells twice, rabbit complement (1:20) was added to the cells and incubated at 37 °C for another 4 h. The content of lactate dehydrogenase (LDH) of supernatants was tested according to specifications of the manufacturer. The assay was performed in triplicate. The percentage of cell lysis was calculated by the following equation:

$$\text{Lysis (\%)} = \frac{(exptl.\ release)-(spontaneous\ release)}{(max.\ release)-(spontaneous\ release)} \times 100\%$$

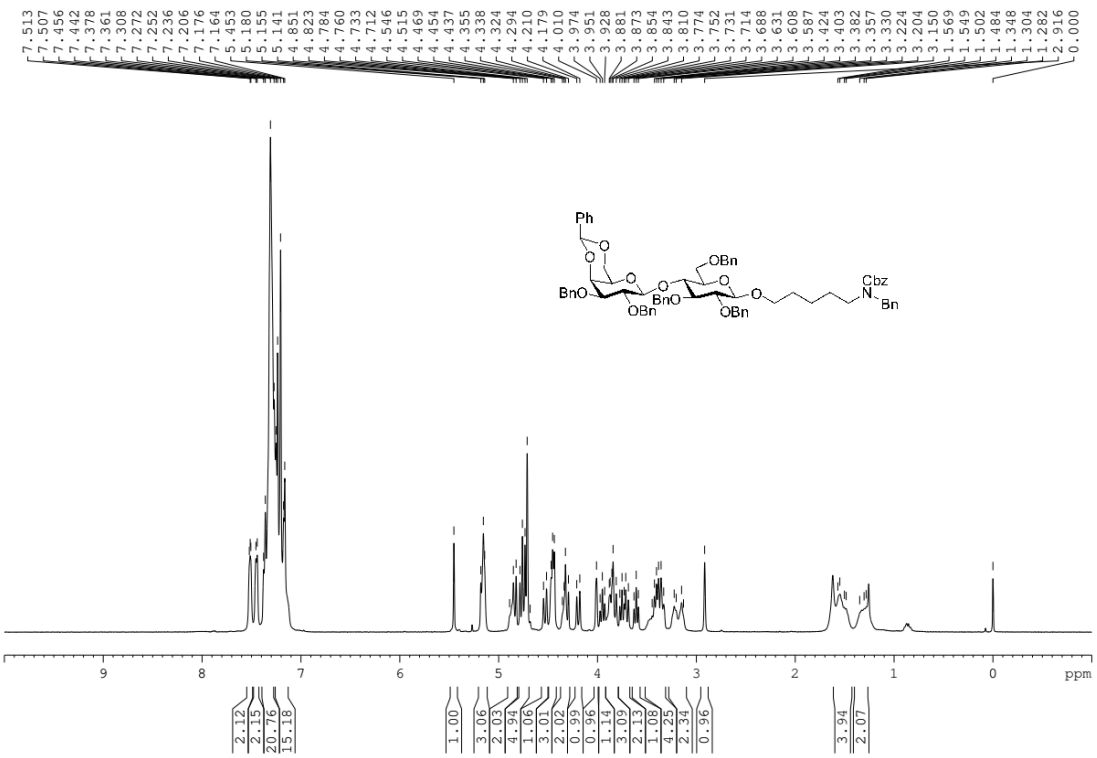
2. NMR (^{1}H, ^{13}C, HMBC) spectra of compounds



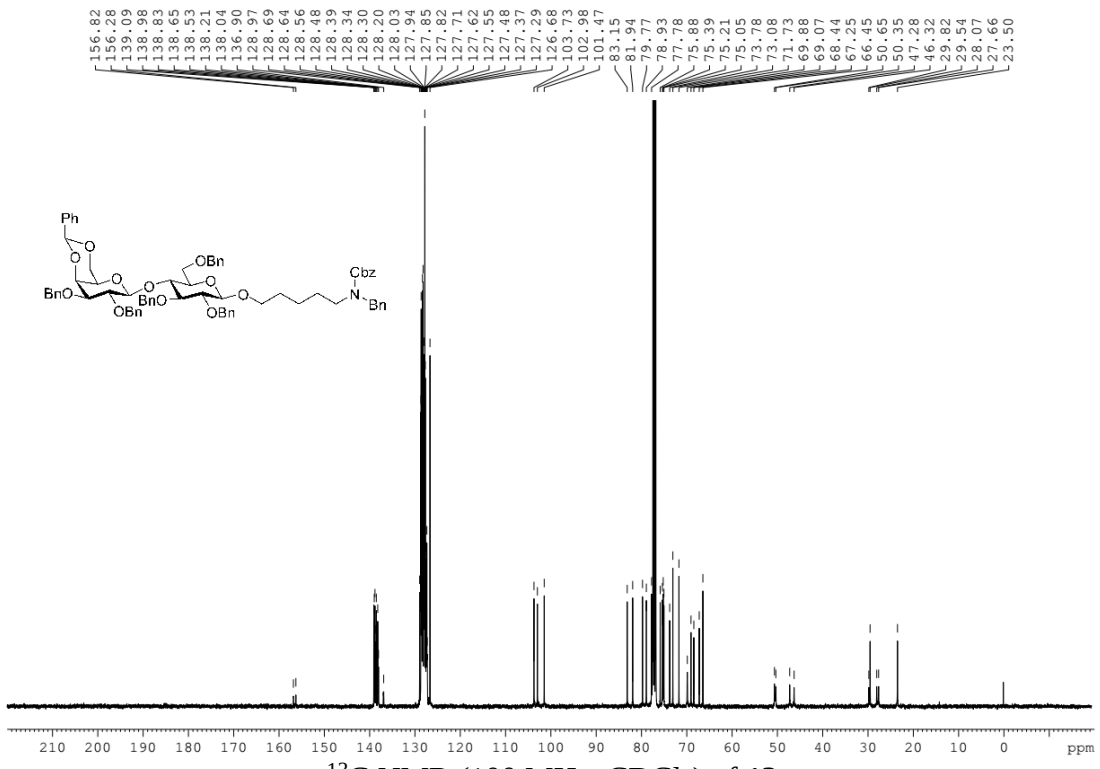
^{1}H NMR (400 MHz, CD_3OD) of **12**



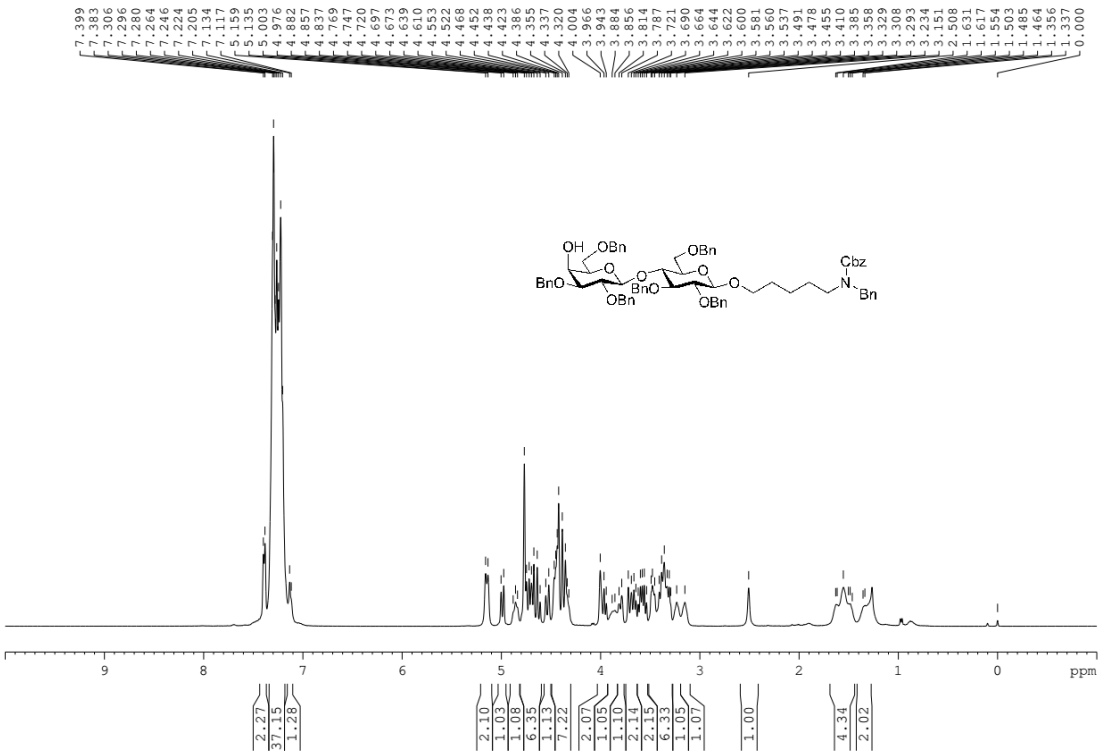
^{13}C NMR (100 MHz, CD_3OD) of **12**



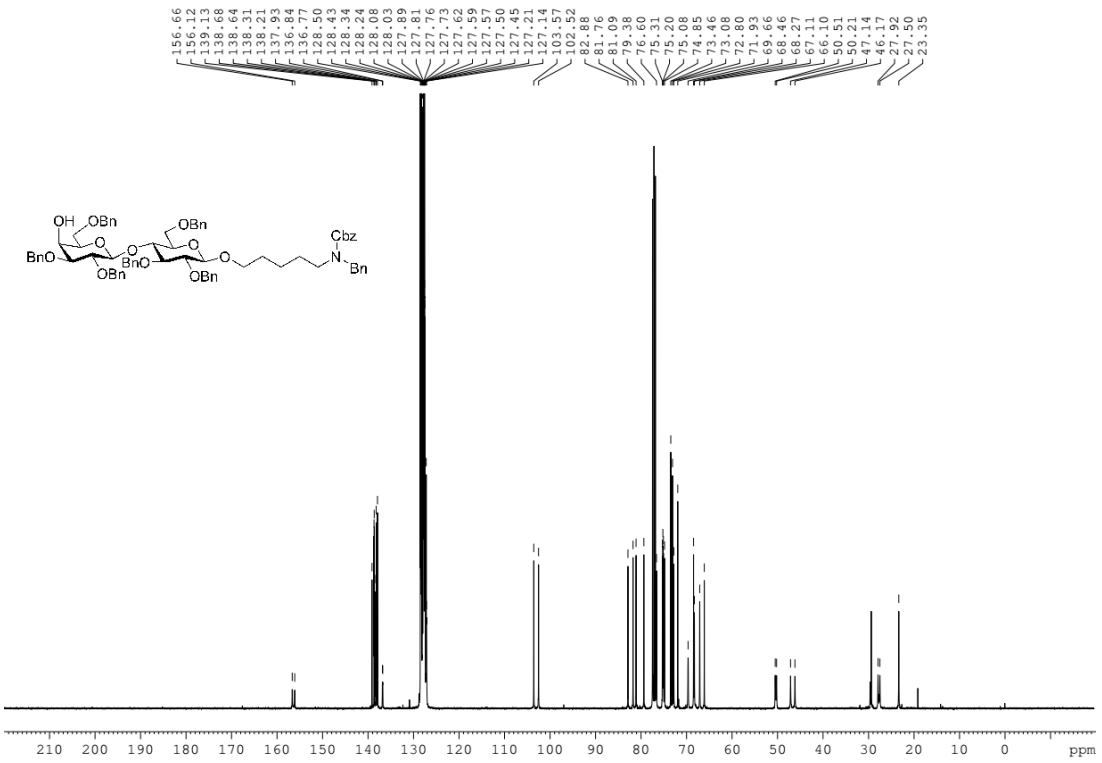
^{1}H NMR (400 MHz, $CDCl_3$) of **13**



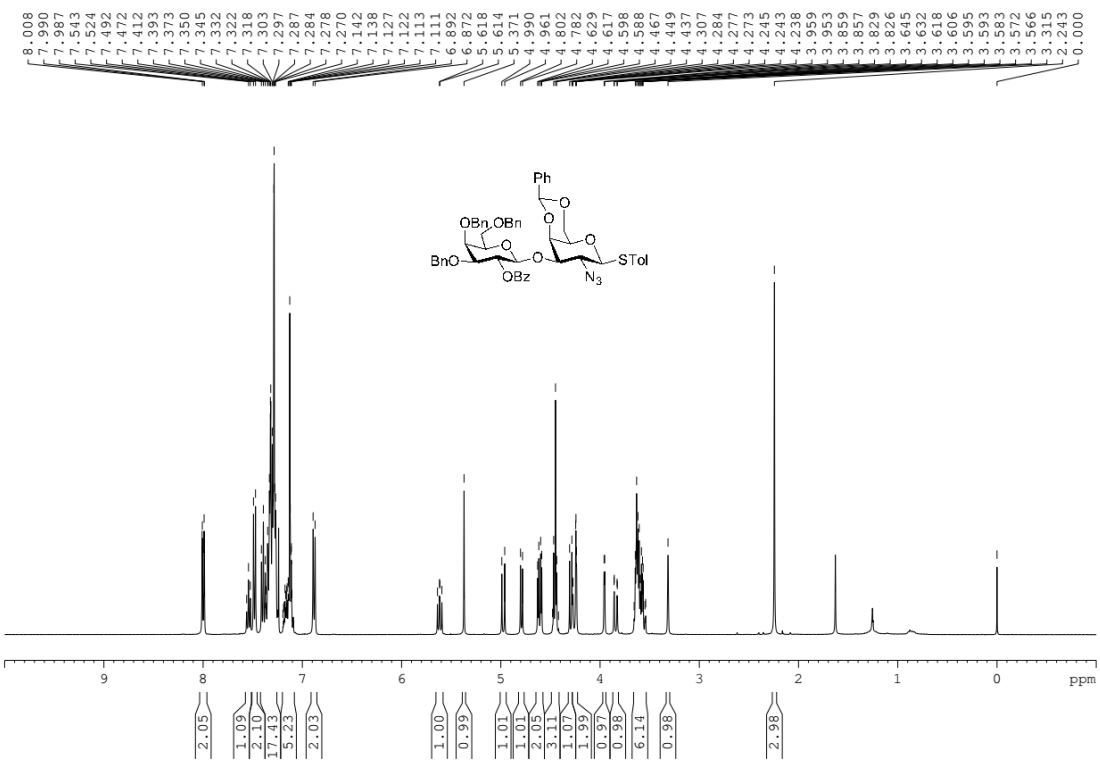
^{13}C NMR (100 MHz, $CDCl_3$) of **13**



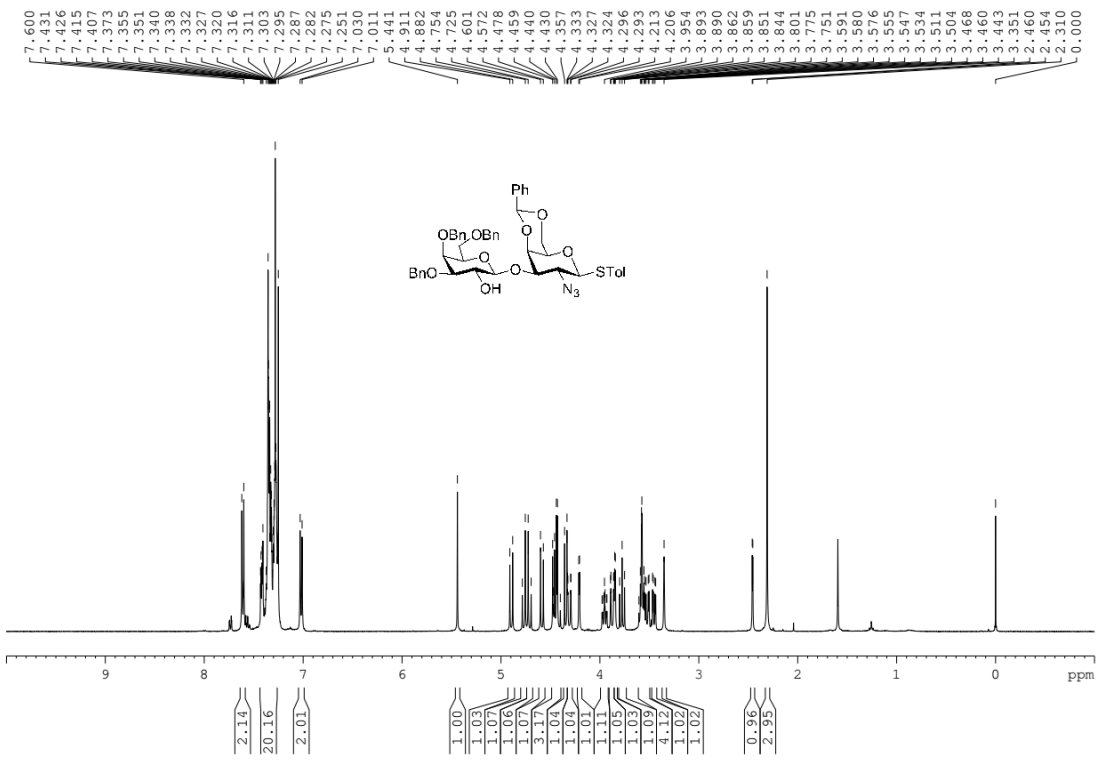
^{1}H NMR (400 MHz, $CDCl_3$) of **14**



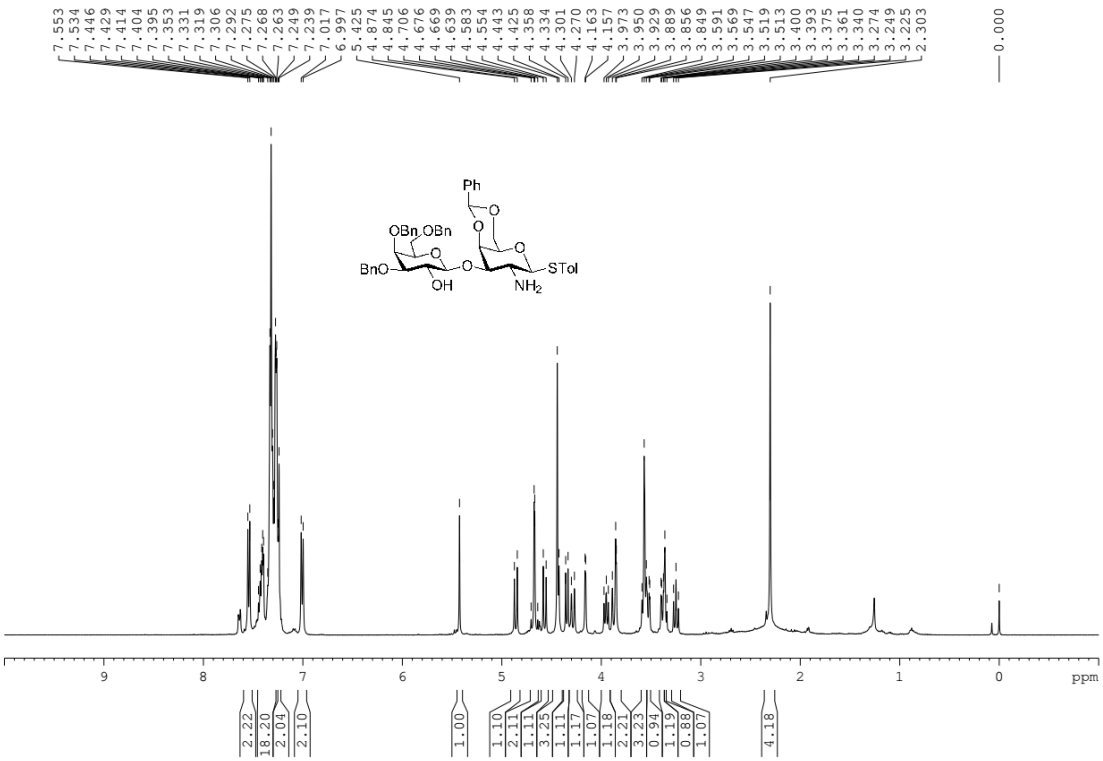
^{13}C NMR (100 MHz, $CDCl_3$) of **14**



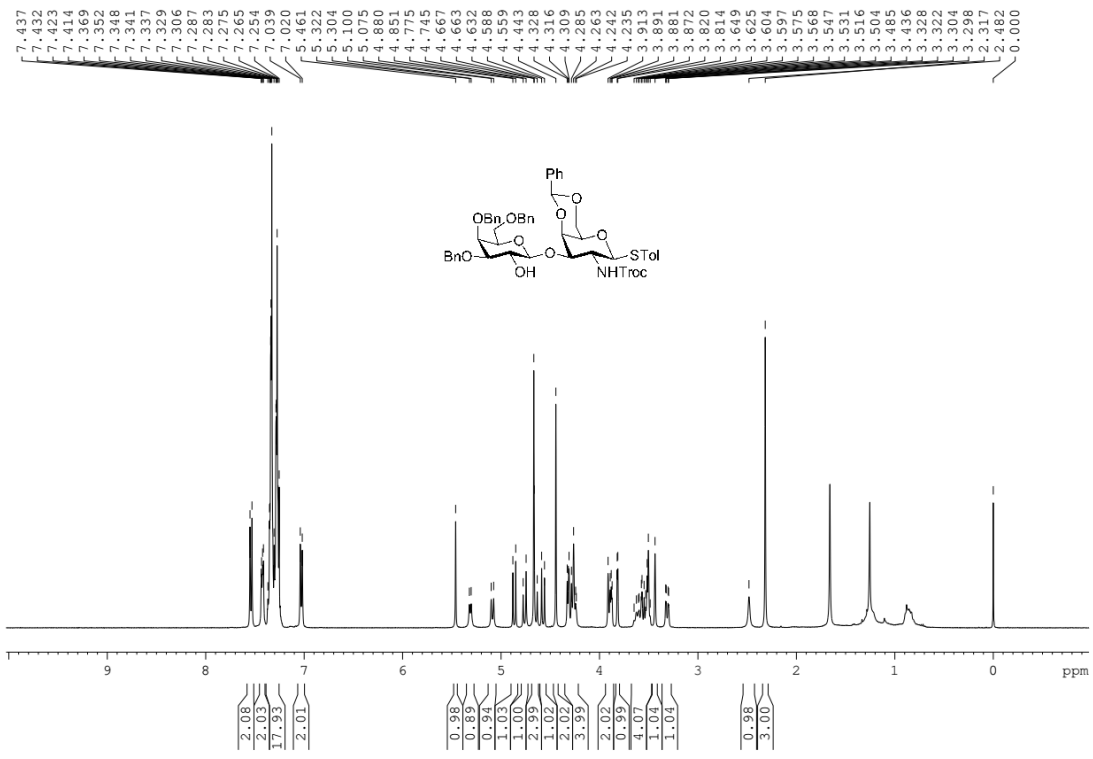
^{1}H NMR (400 MHz, $CDCl_3$) of **4**



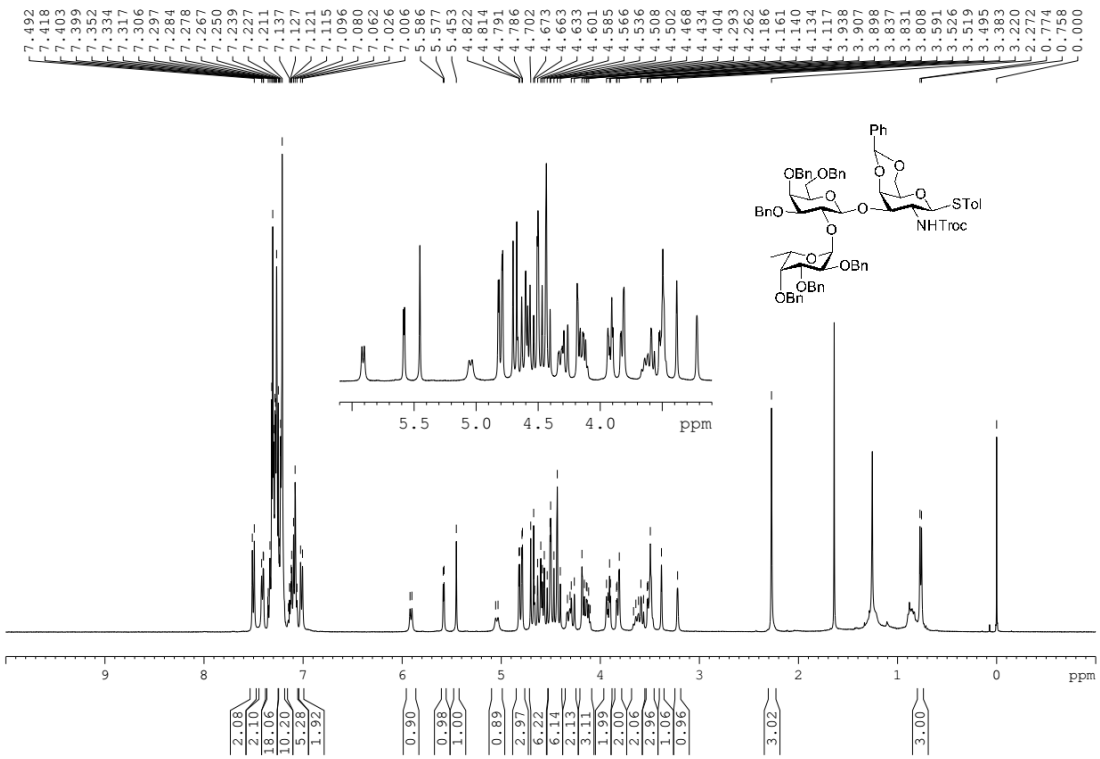
^{1}H NMR (400 MHz, $CDCl_3$) of **5**



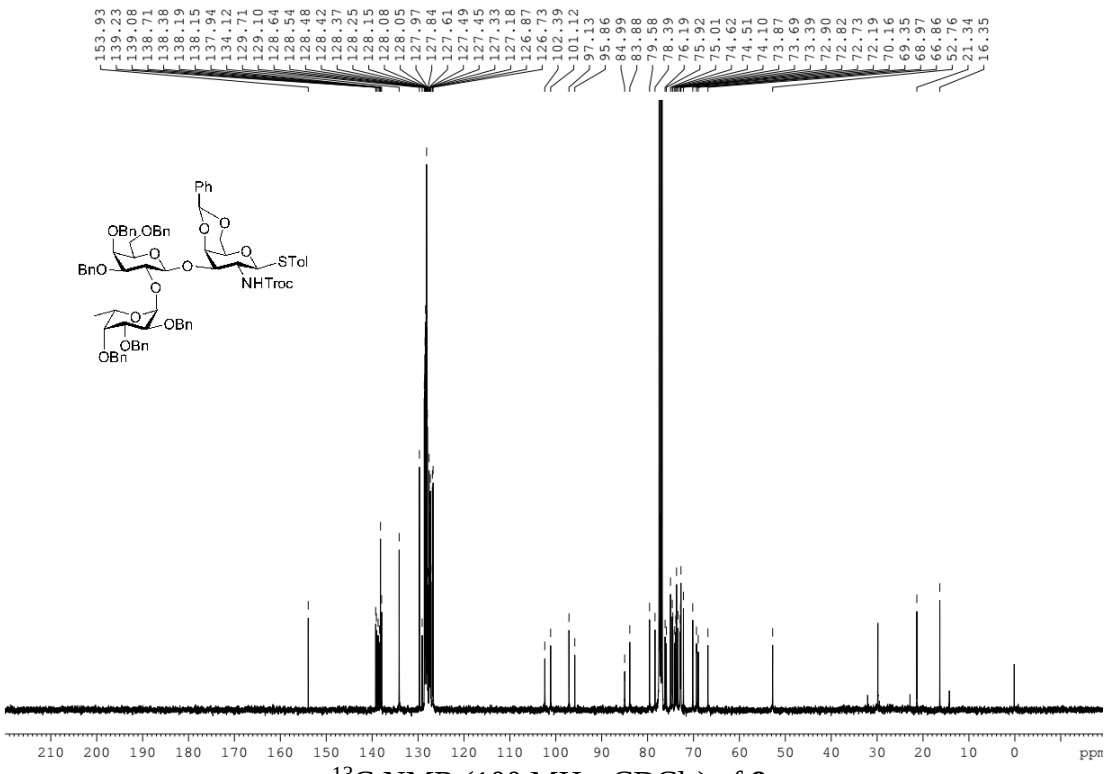
^{1}H NMR (400 MHz, $CDCl_3$) of **6**



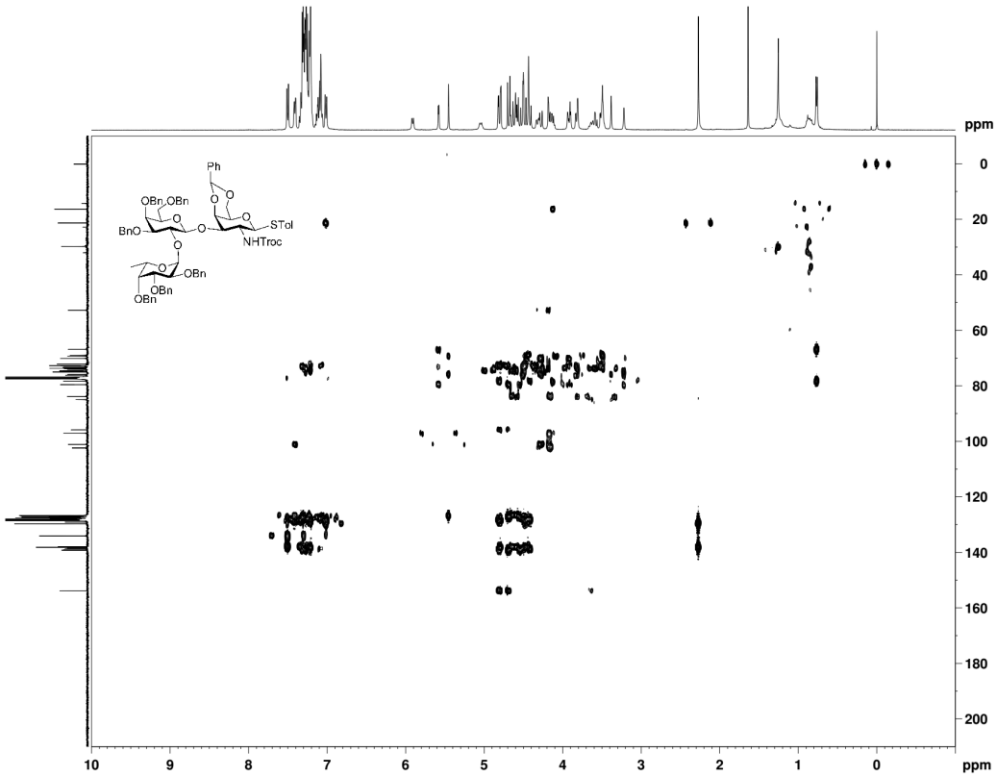
^{1}H NMR (400 MHz, $CDCl_3$) of **7**



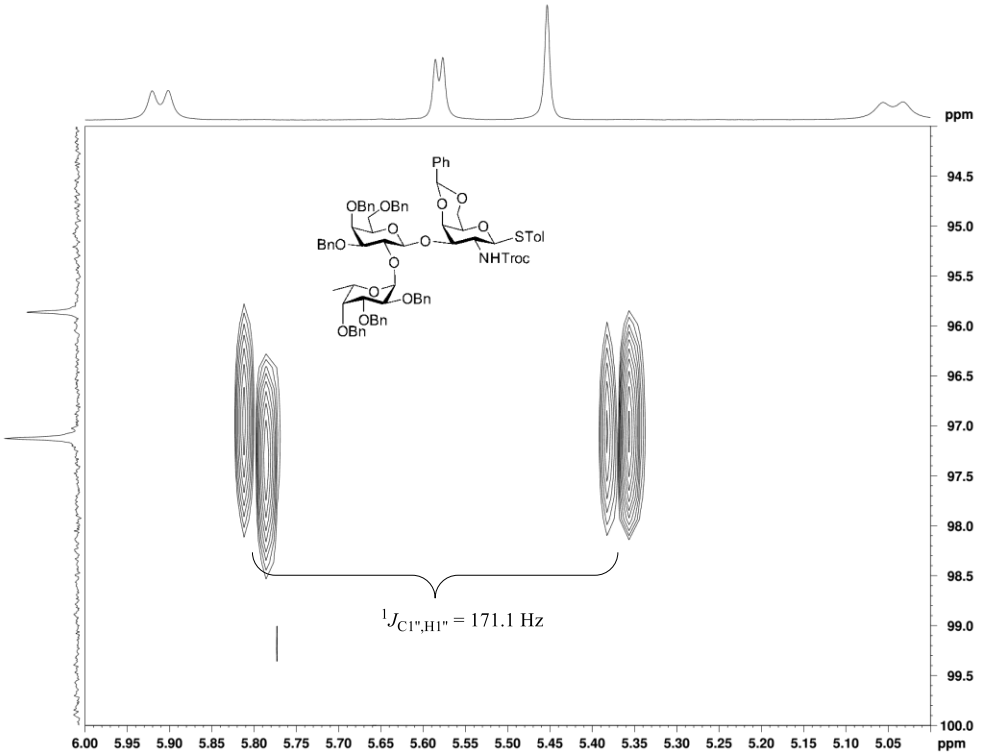
^{1}H NMR (400 MHz, $CDCl_3$) of **8**



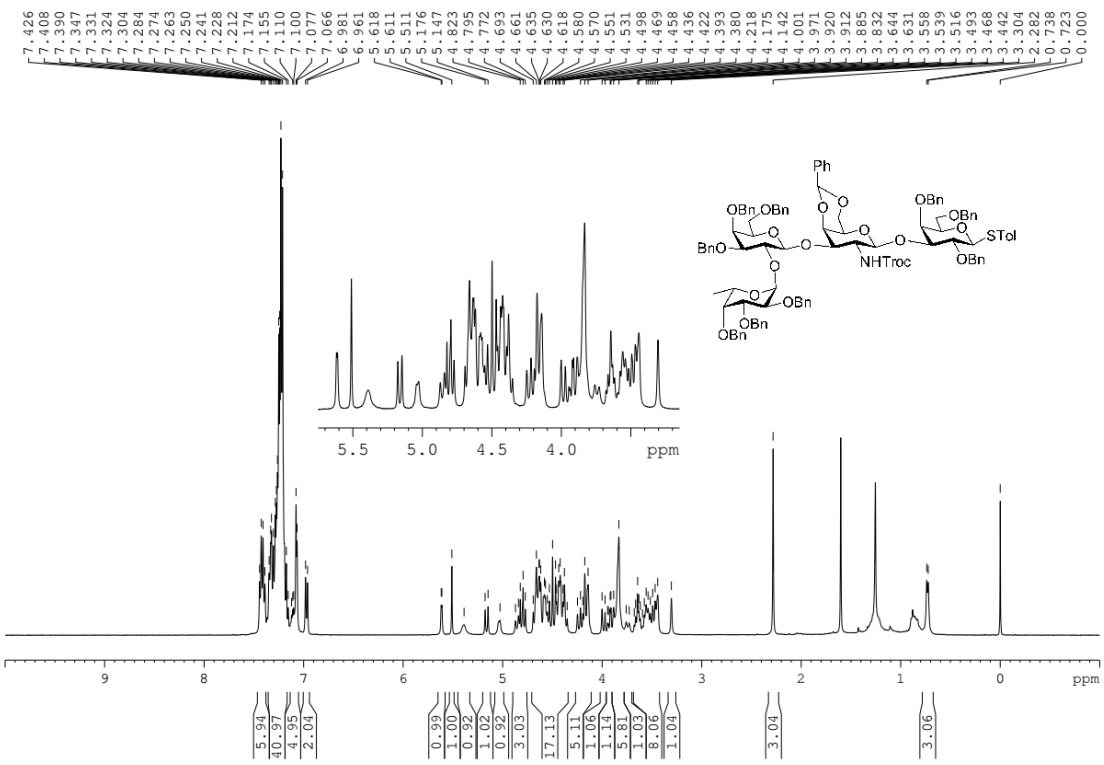
^{13}C NMR (100 MHz, $CDCl_3$) of **8**



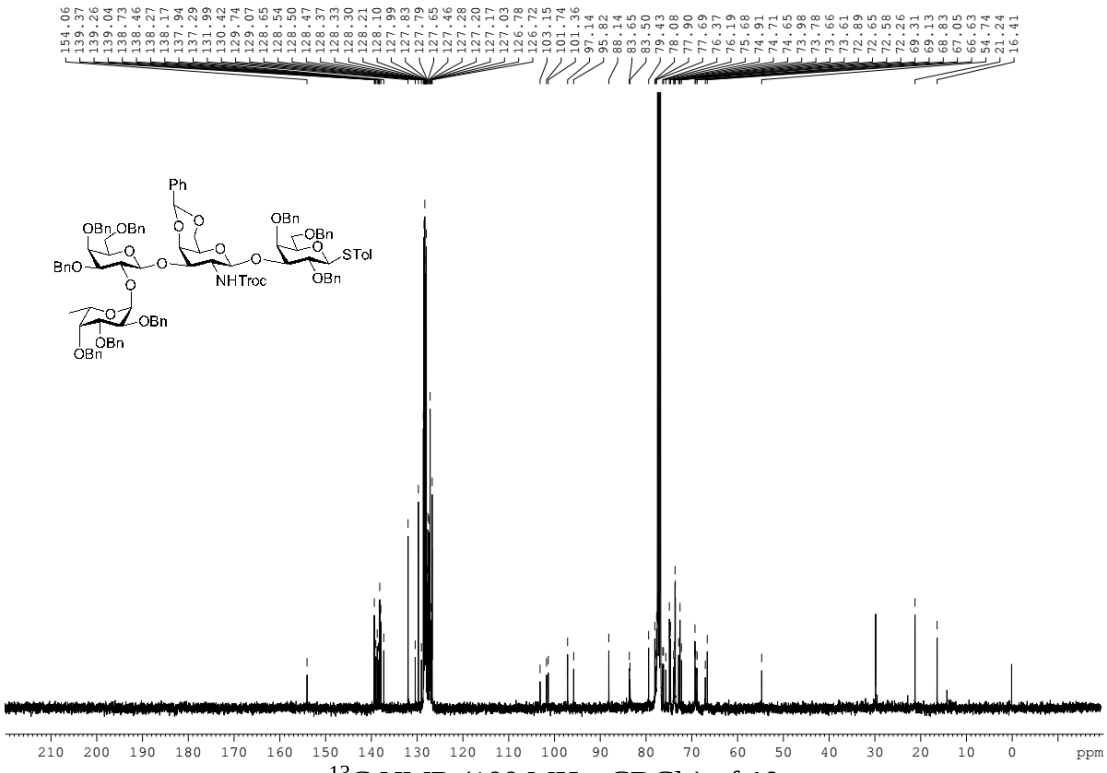
HMBC NMR of **8**



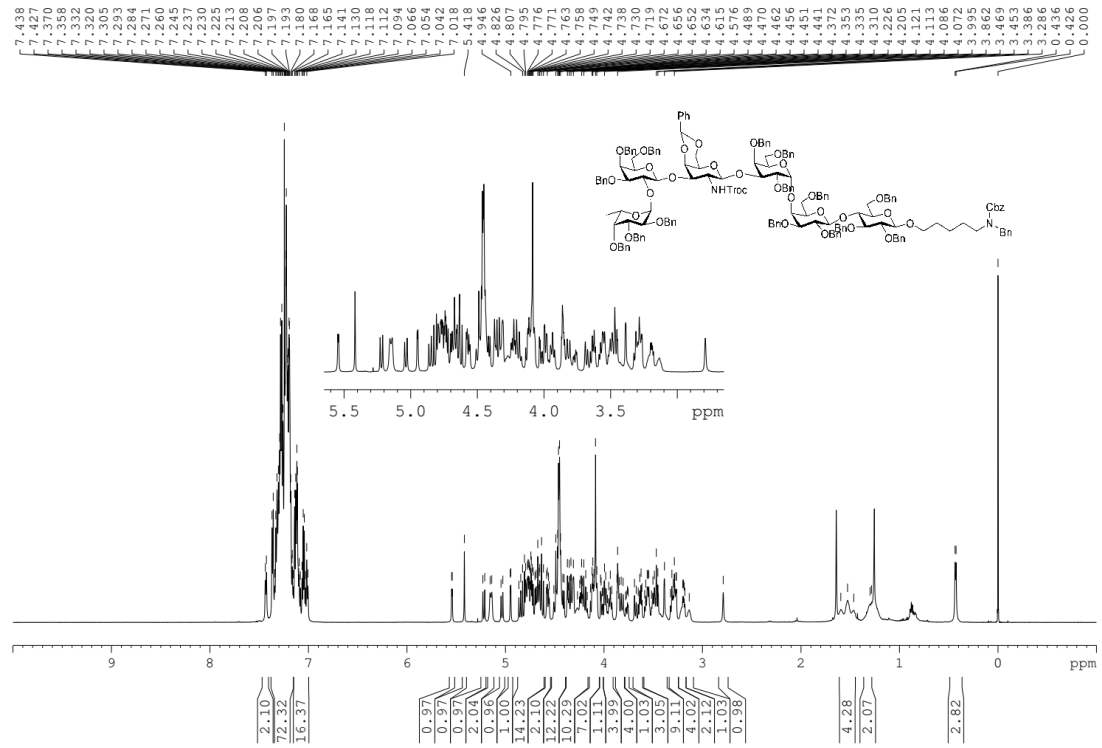
Partial HMBC NMR of **8**



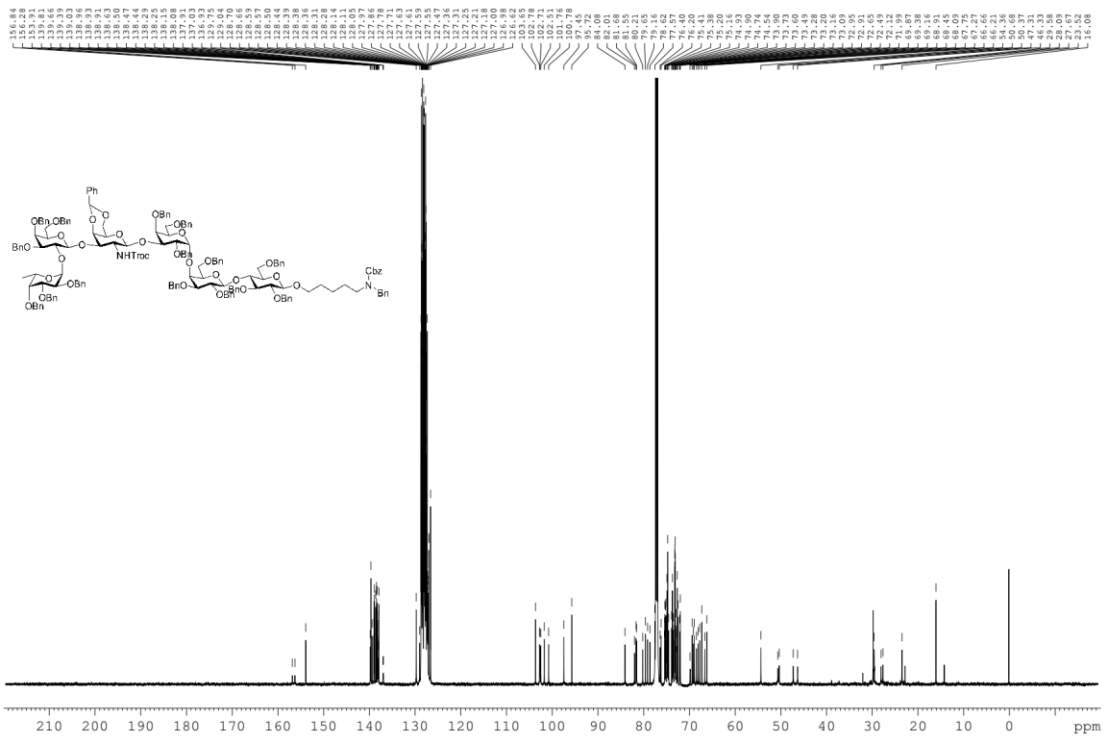
^{1}H NMR (400 MHz, $CDCl_3$) of **10**



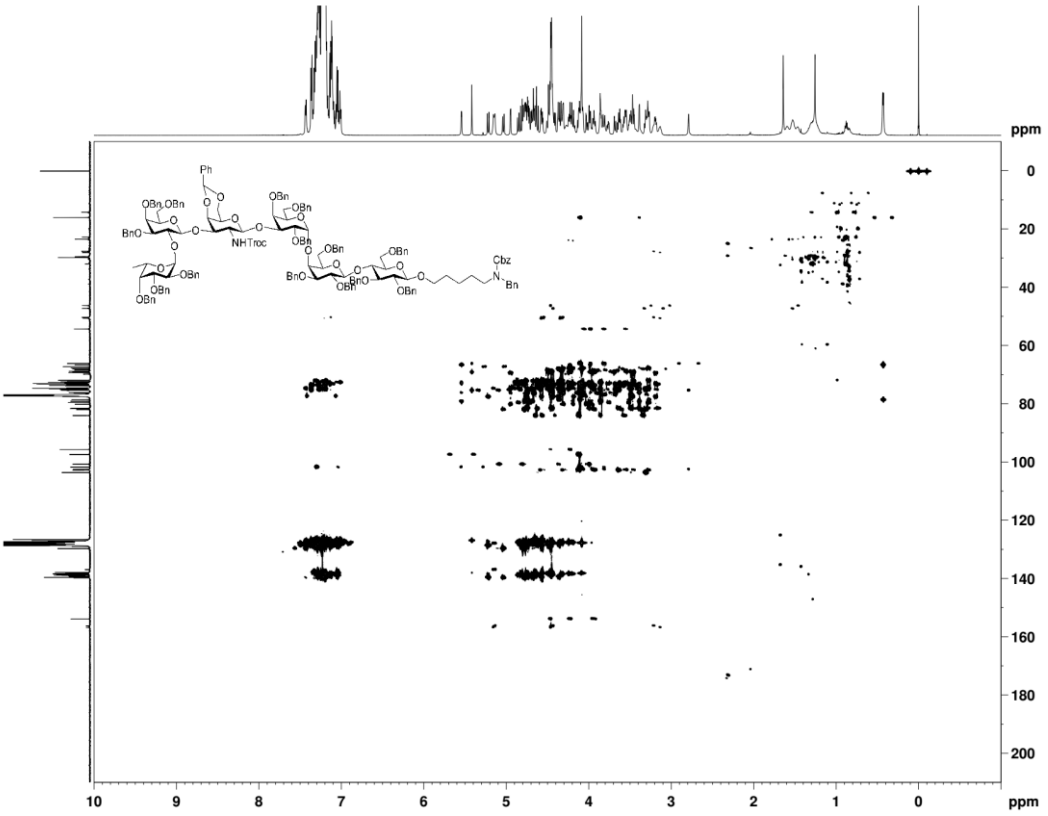
^{13}C NMR (100 MHz, $CDCl_3$) of **10**



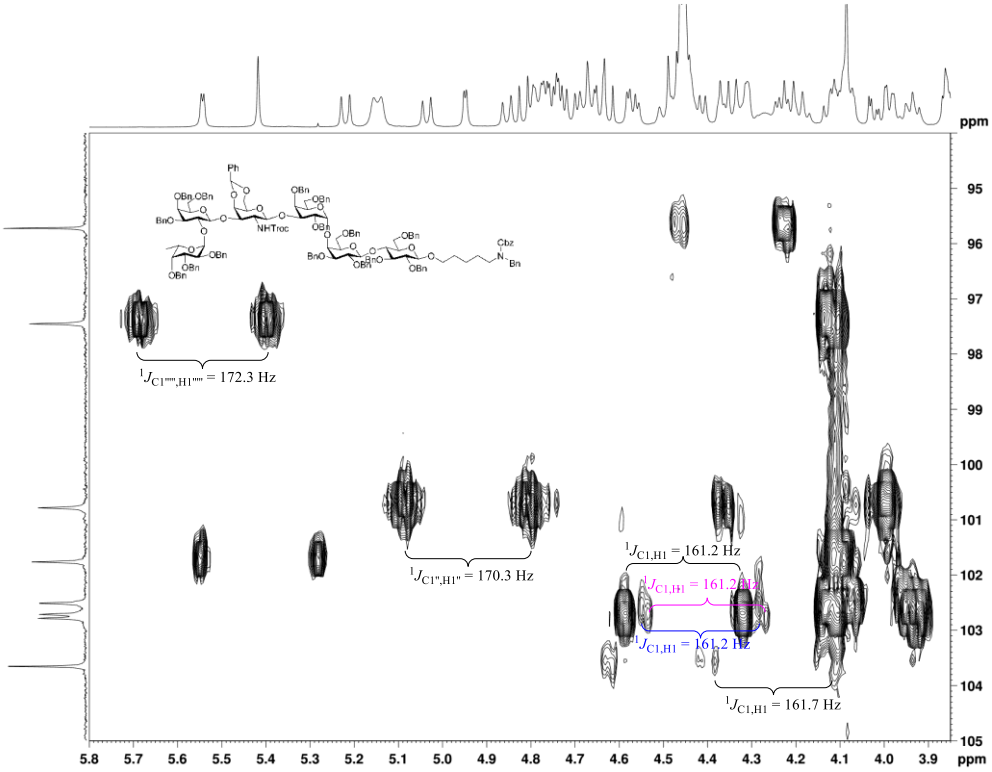
^{1}H NMR (600 MHz, $CDCl_3$) of **15α**



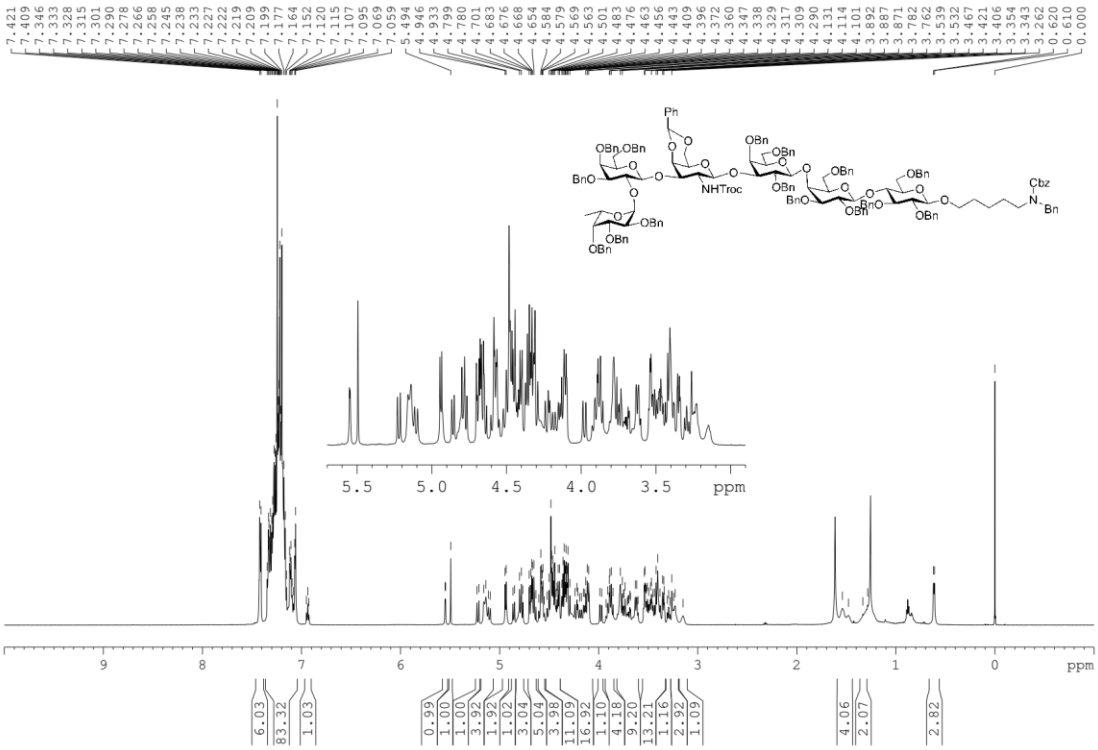
^{13}C NMR (150 MHz, $CDCl_3$) of **15α**



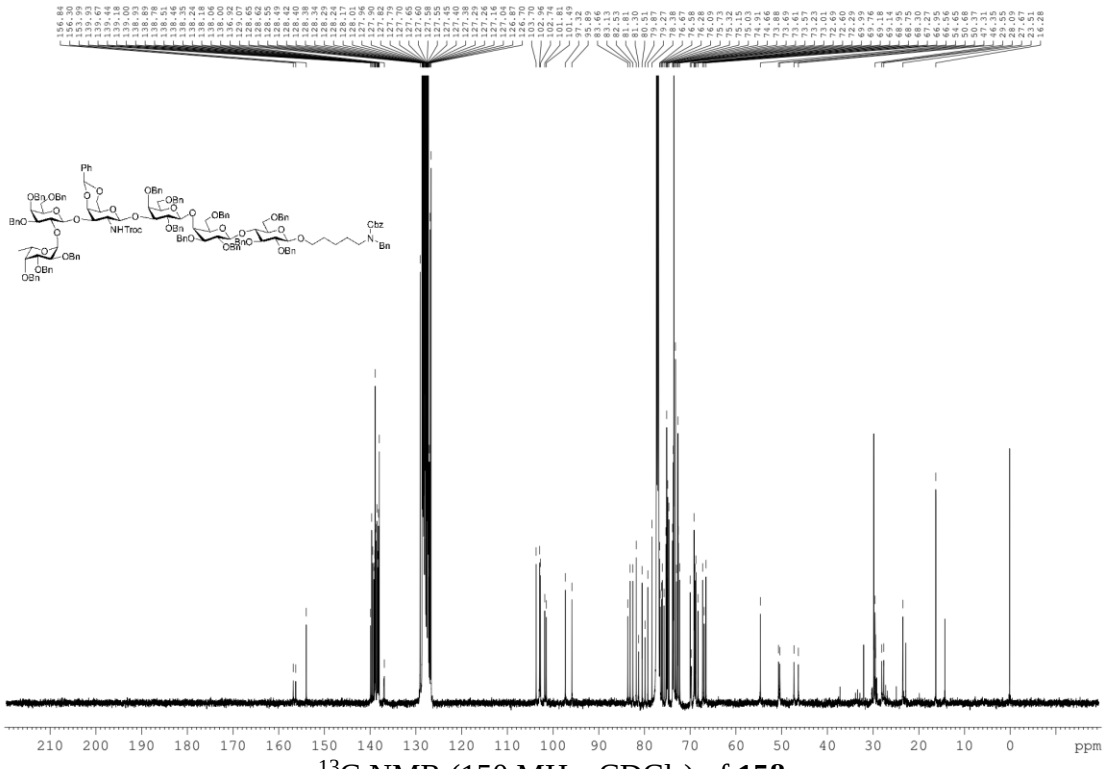
HMBC of **15α**



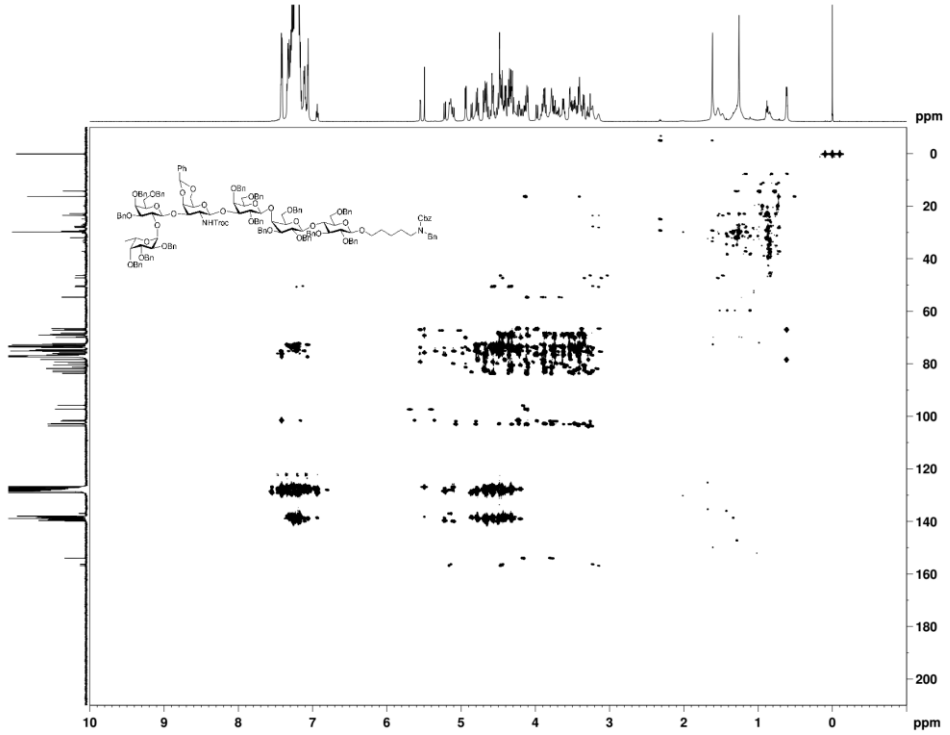
Partial HMBC of **15α**



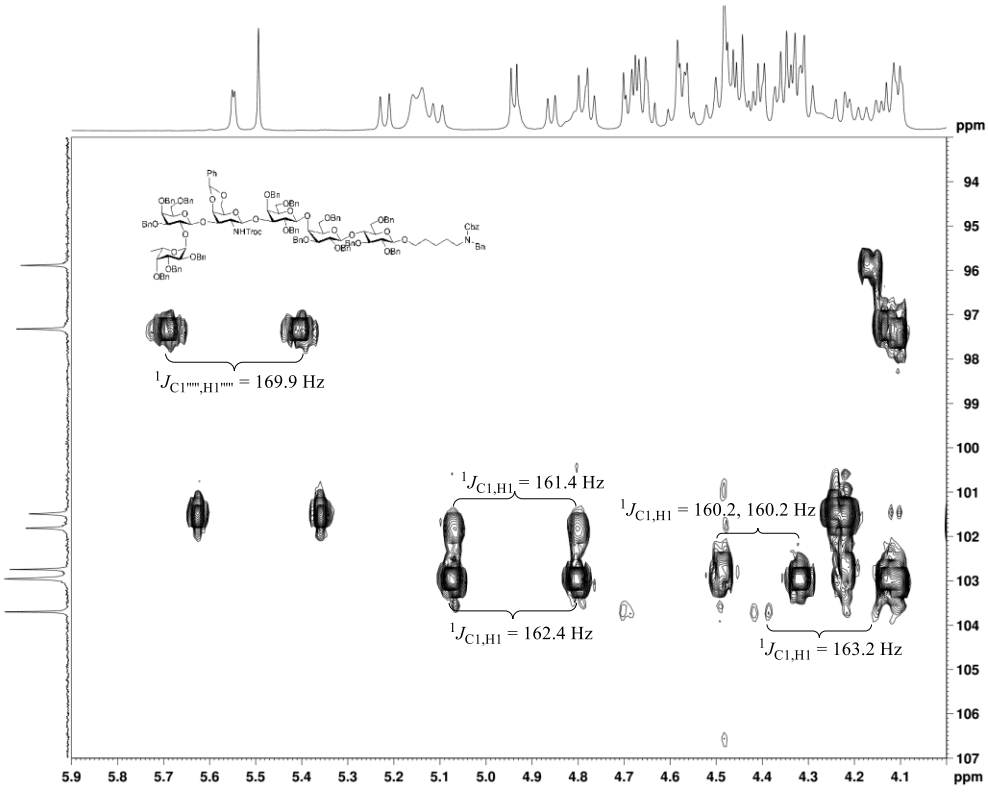
^{1}H NMR (600 MHz, $CDCl_3$) of **15β**



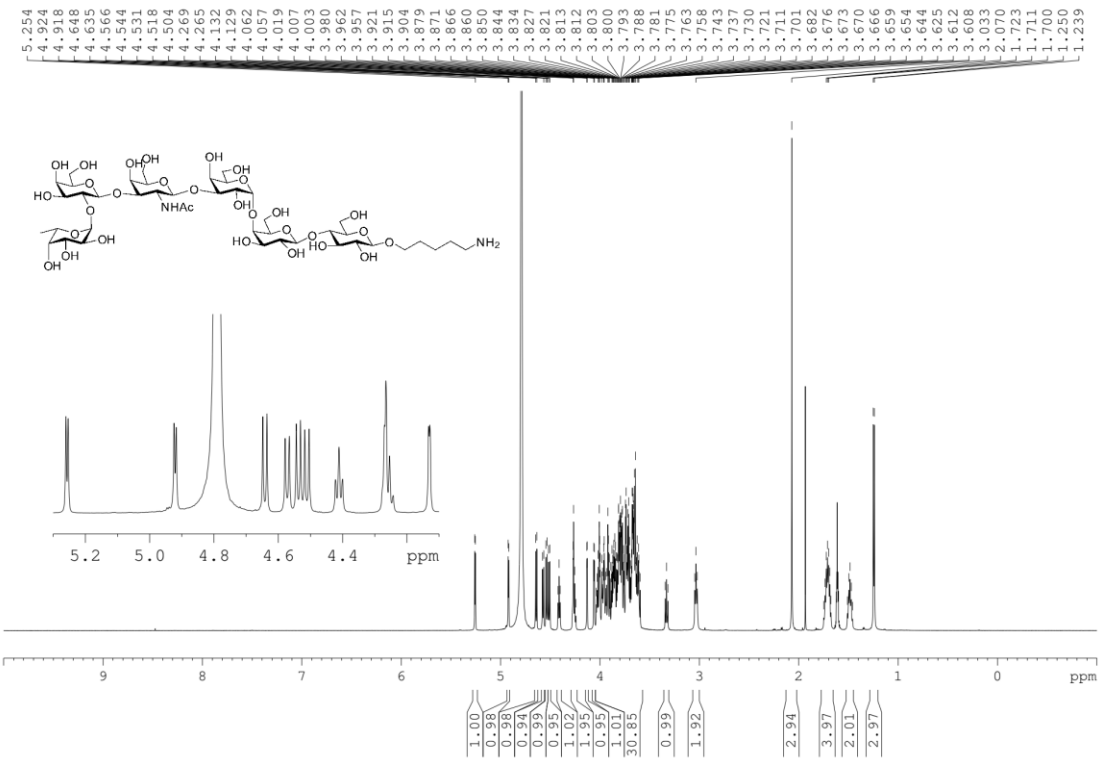
^{13}C NMR (150 MHz, $CDCl_3$) of **15β**



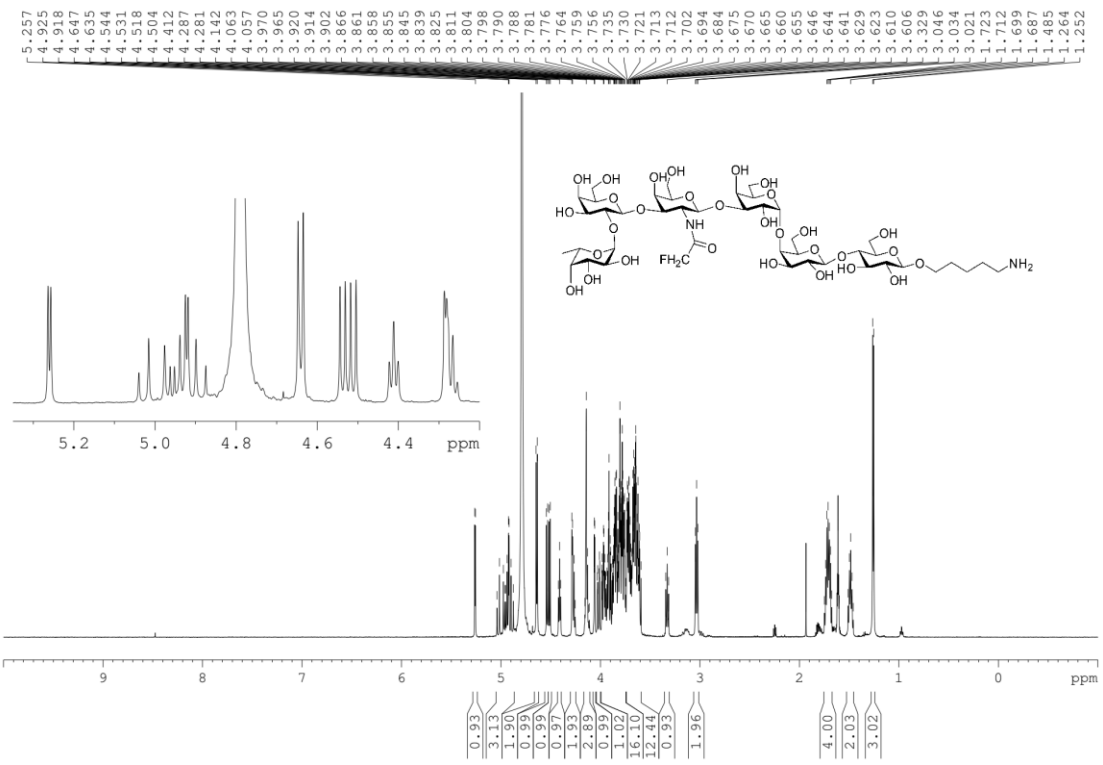
HMBC of **15β**



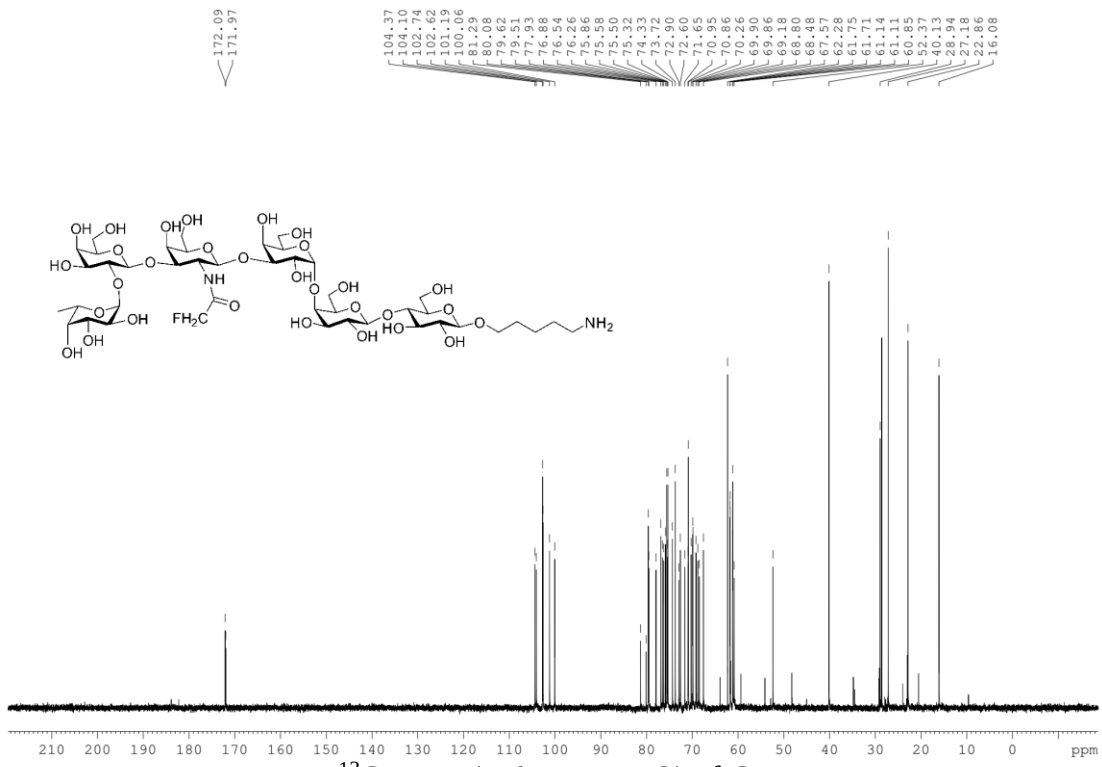
Partial HMBC of **15β**



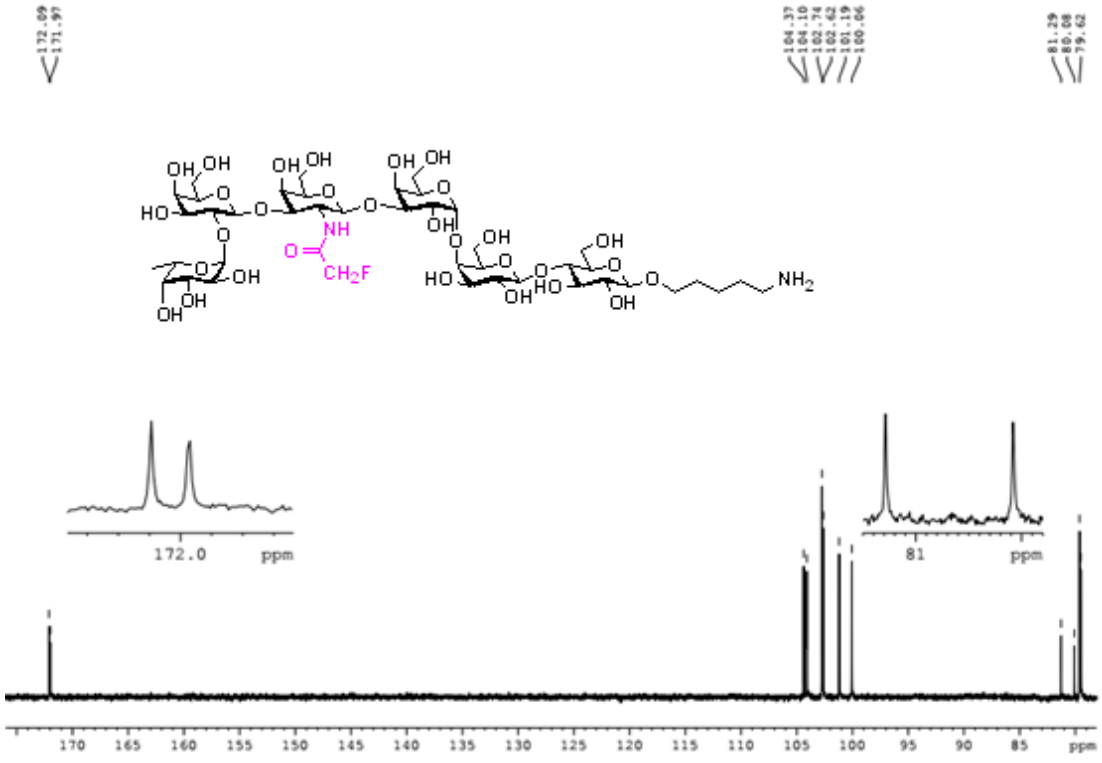
^{1}H NMR (600 MHz, D_2O) of **G0**



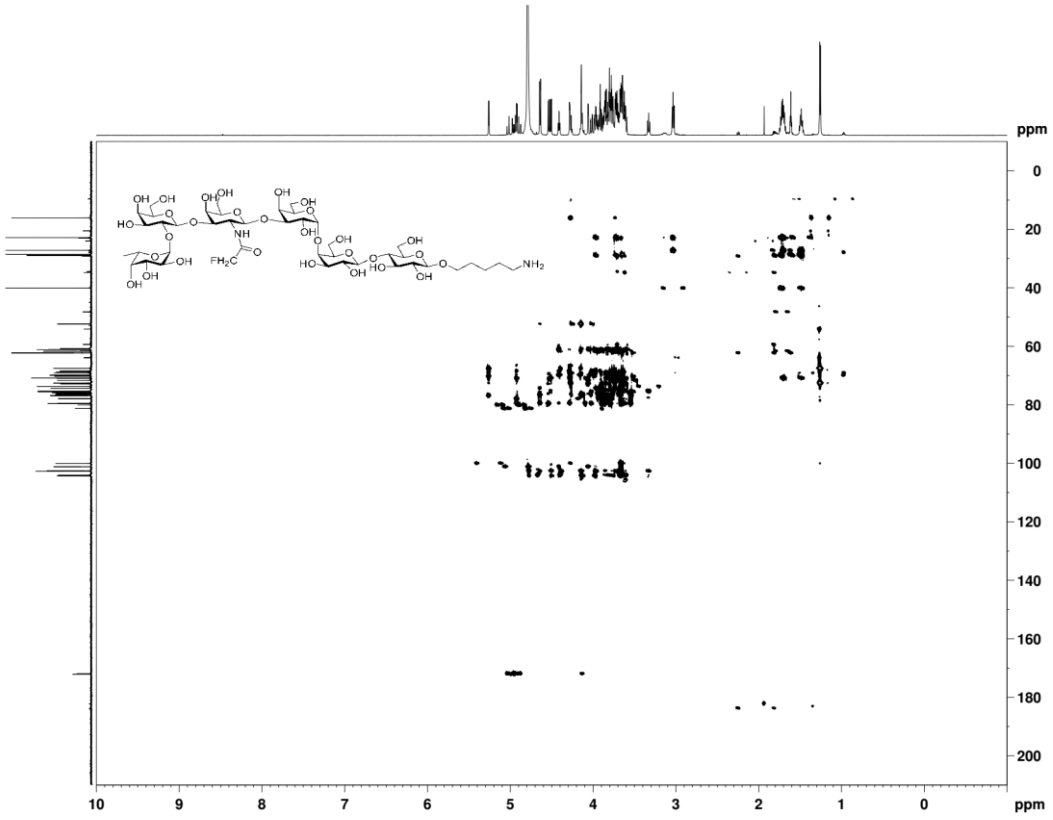
^{1}H NMR (600 MHz, D_2O) of **G1**



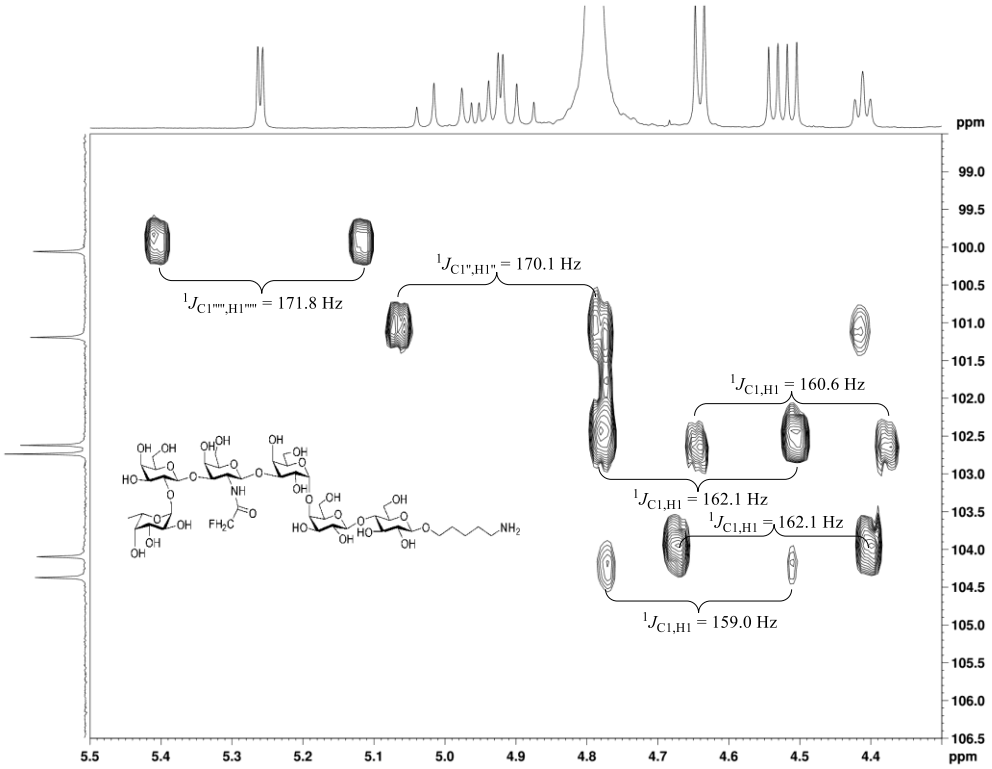
^{13}C NMR (150 MHz, D_2O) of **G1**



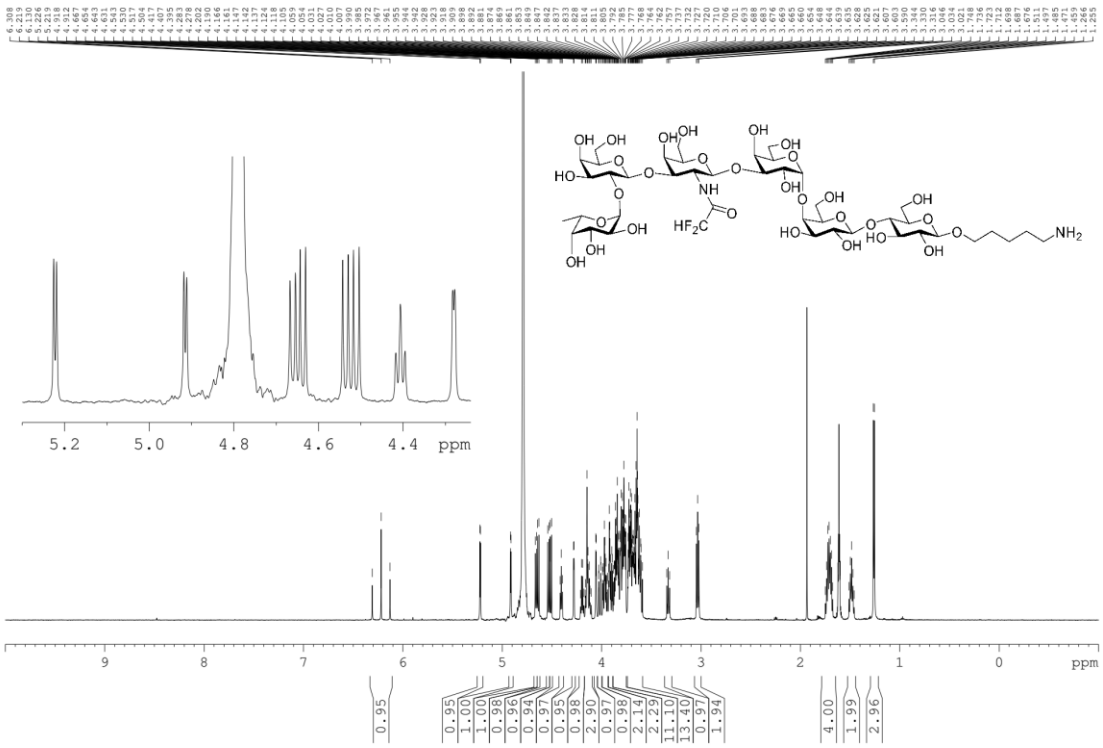
Partial ^{13}C NMR (150 MHz, D_2O) of **G1**



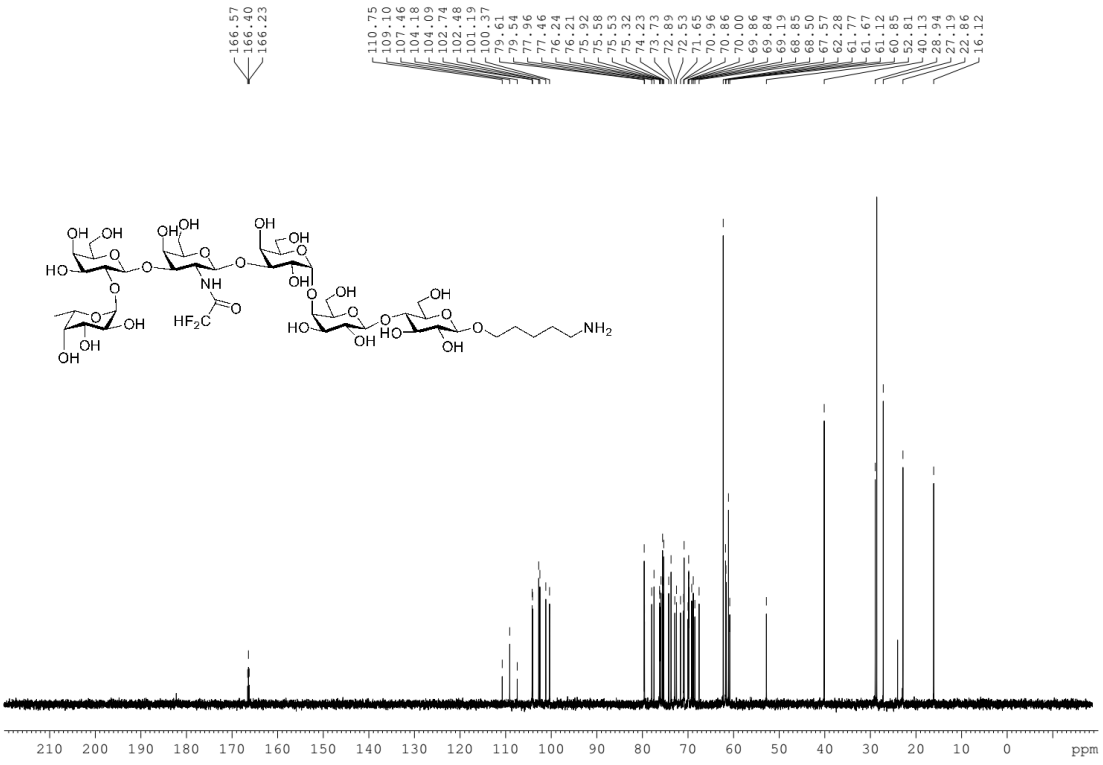
HMBC of **G1**



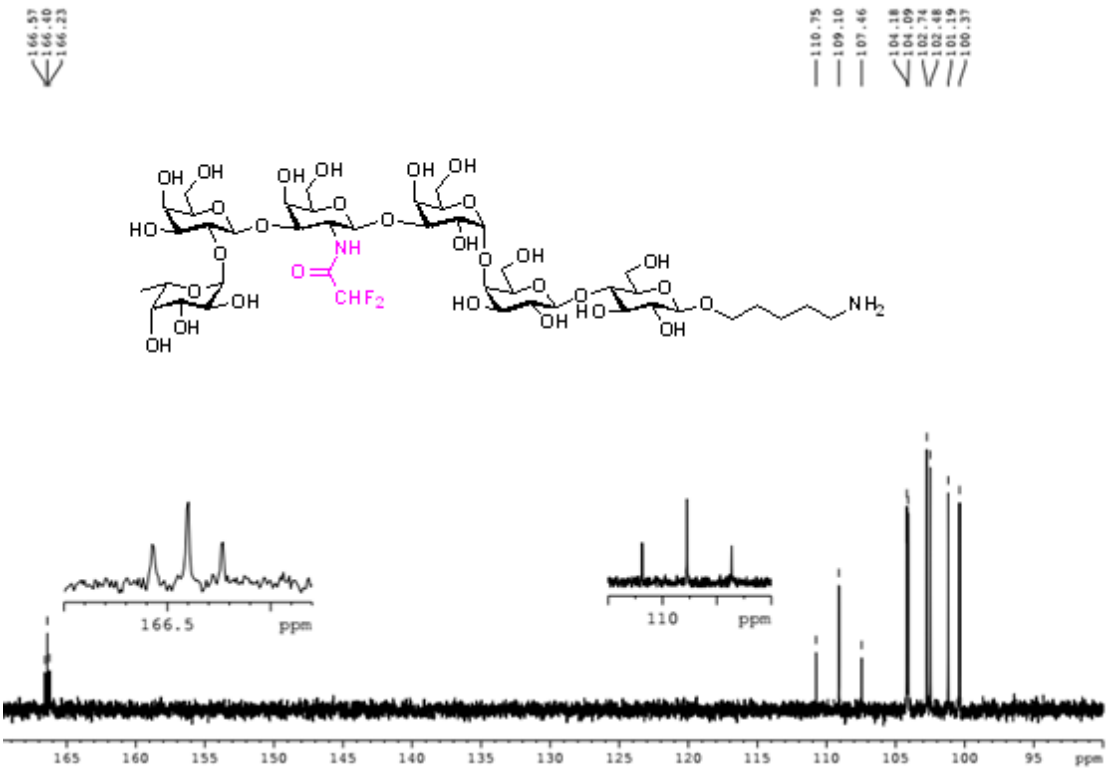
Partial HMBC of **G1**



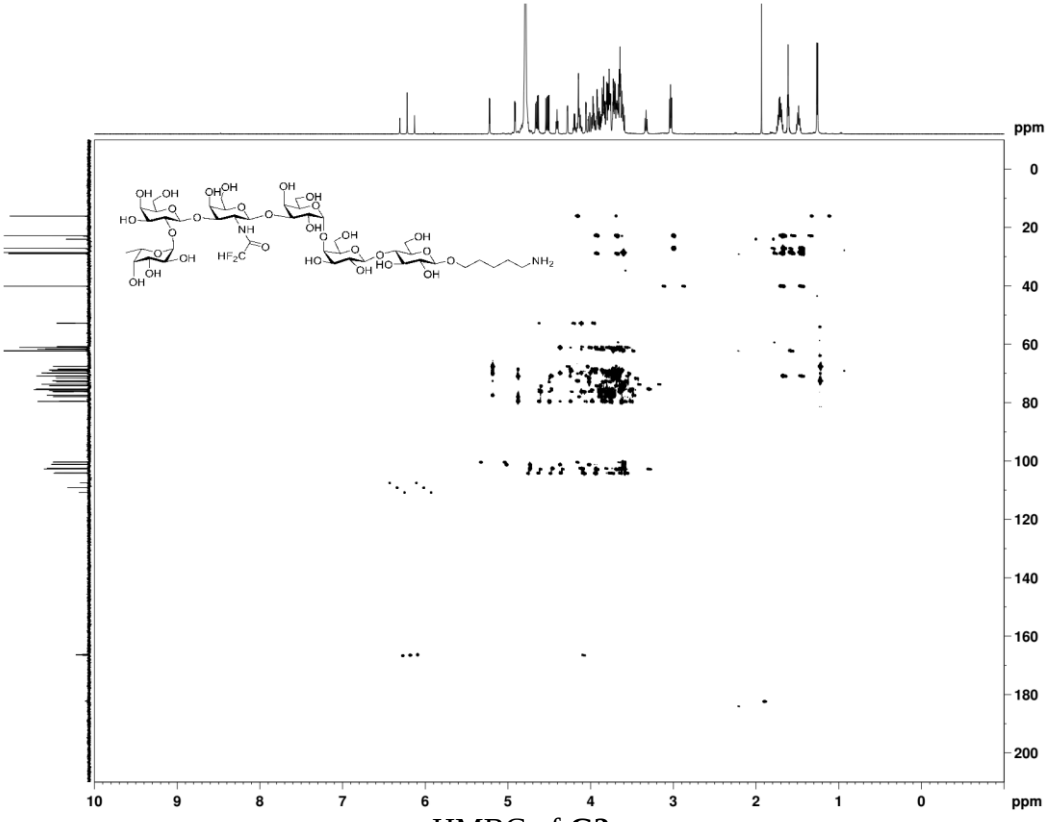
^{1}H NMR (600 MHz, D_2O) of **G2**



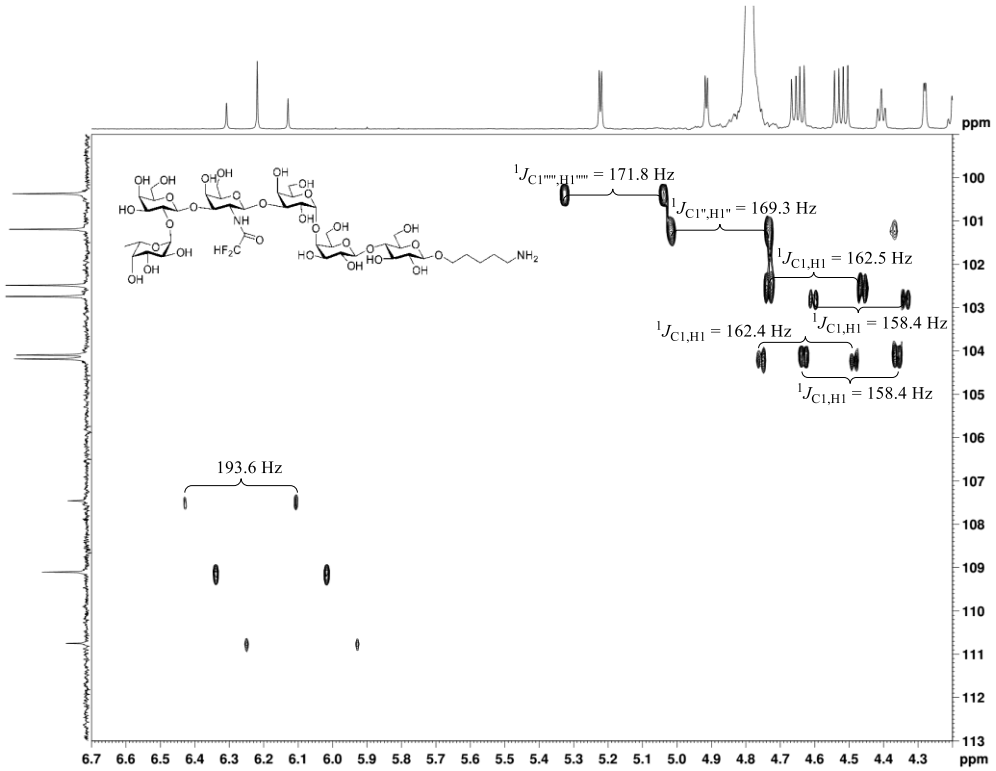
^{13}C NMR (150 MHz, D_2O) of **G2**



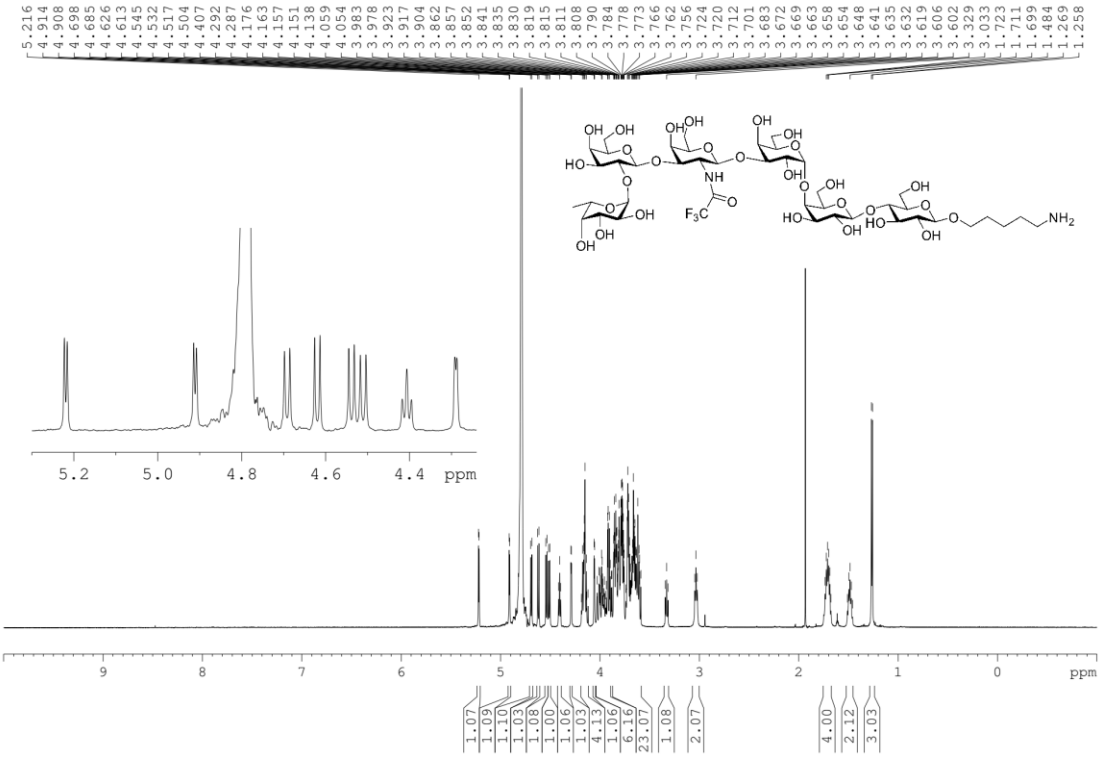
Partial ^{13}C NMR (150 MHz, D_2O) of **G2**



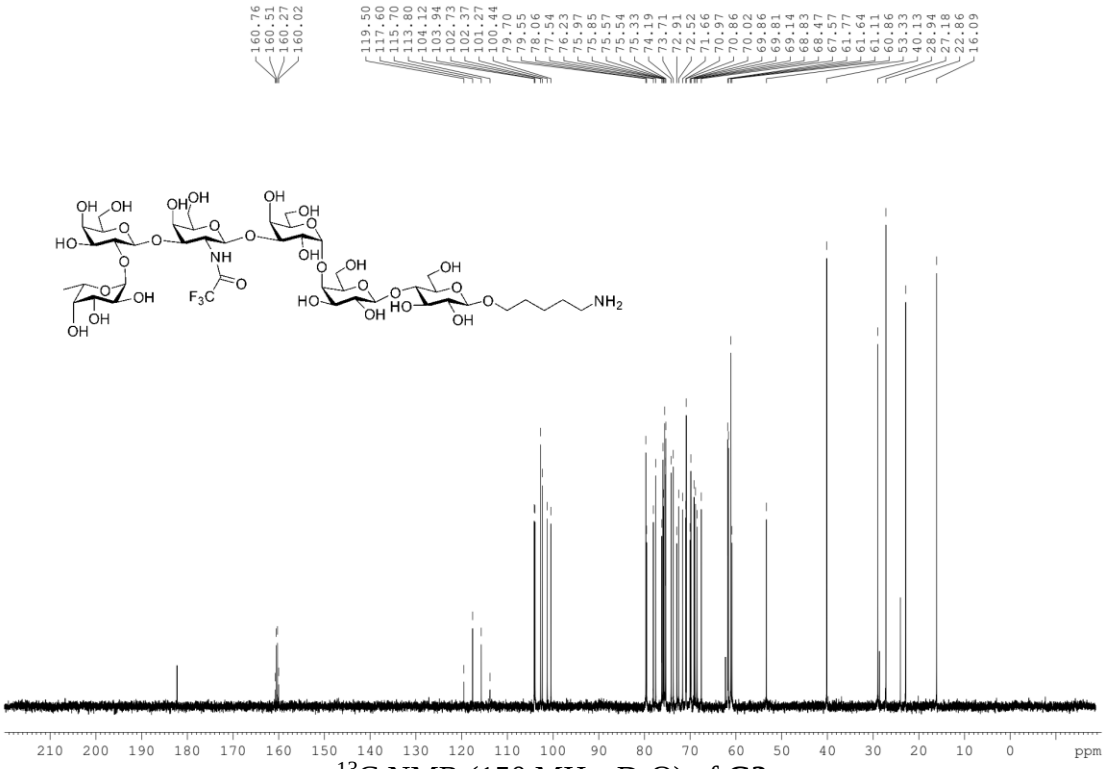
HMBC of **G2**



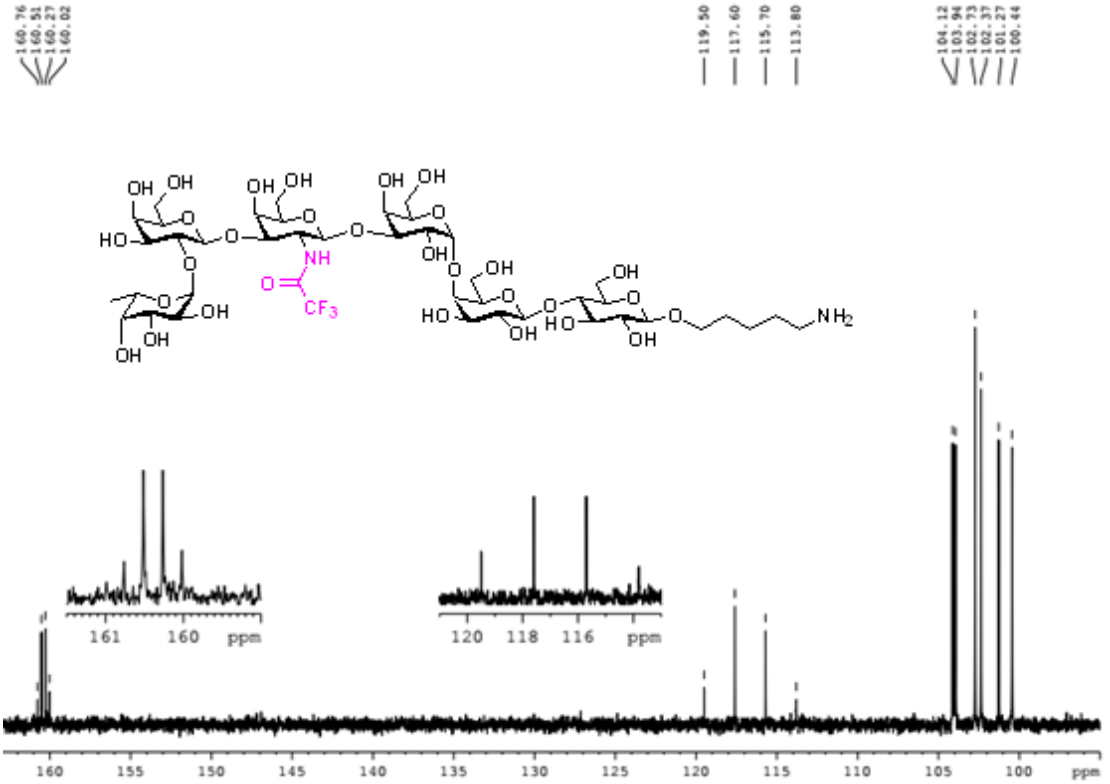
Partial HMBC of **G2**



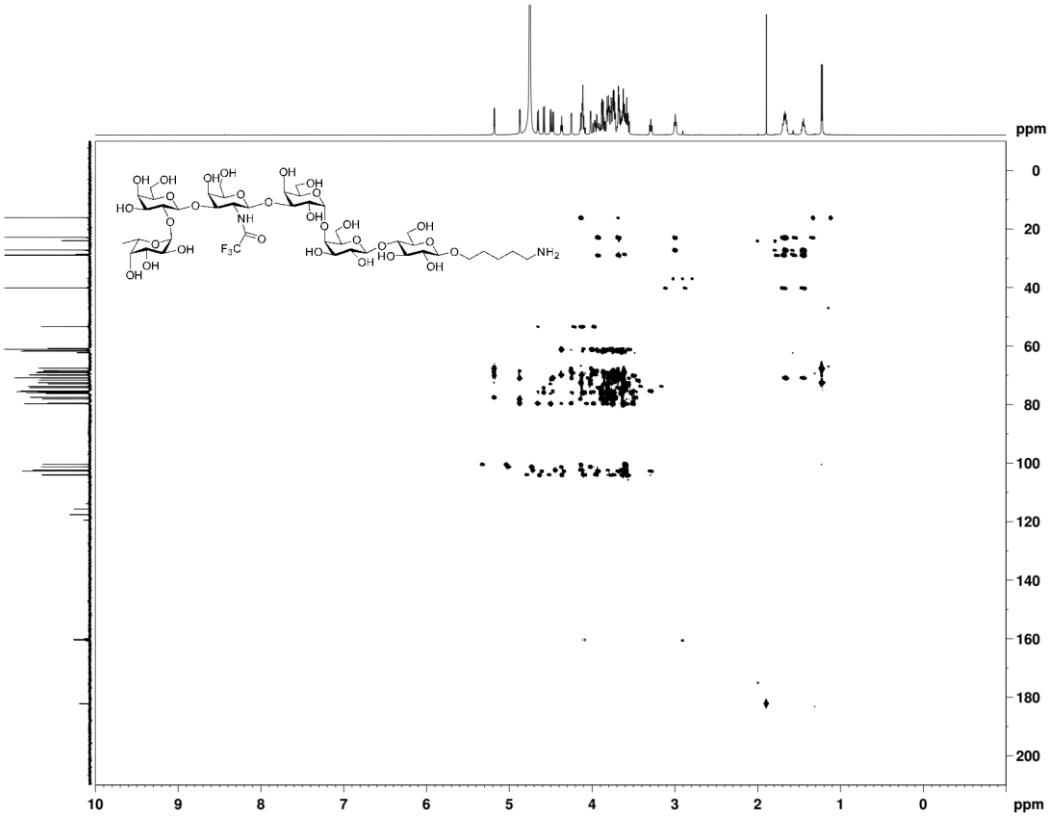
^{1}H NMR (600 MHz, D_2O) of **G3**



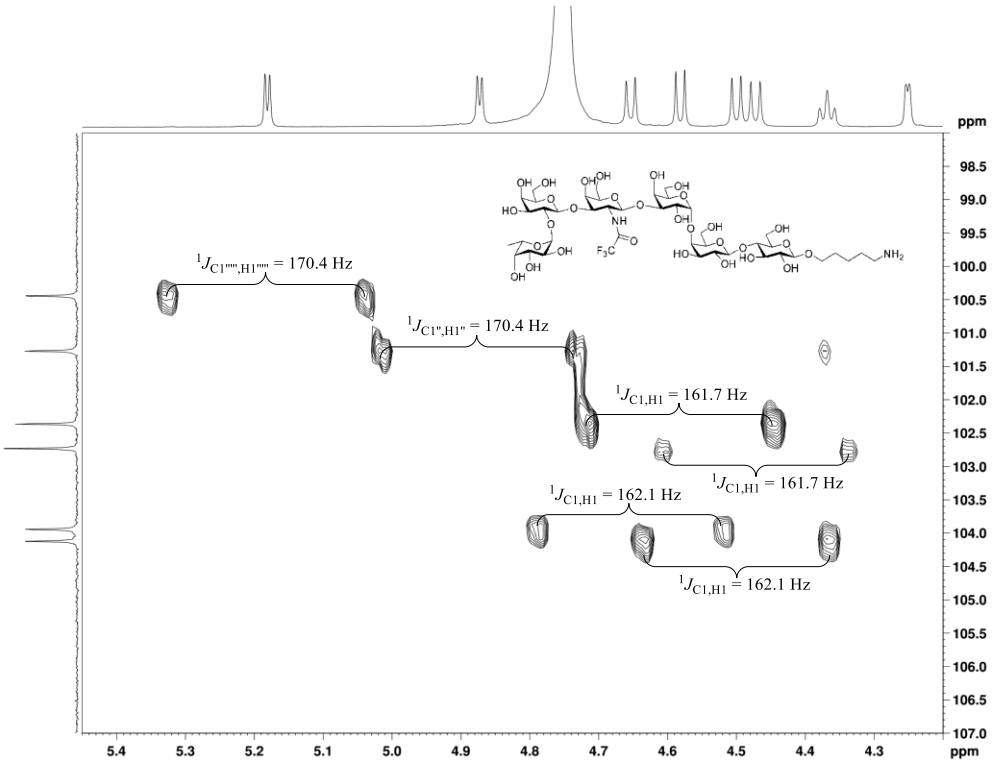
^{13}C NMR (150 MHz, D_2O) of **G3**



Partial ^{13}C NMR (150 MHz, D_2O) of **G3**

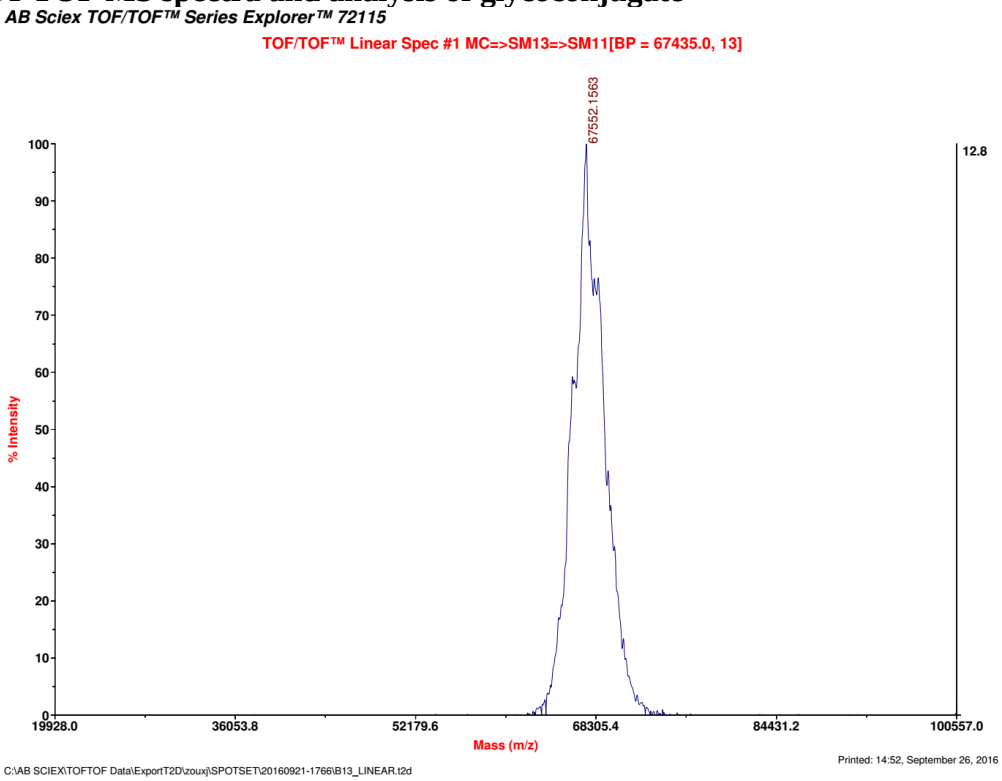


HMBC of **G3**

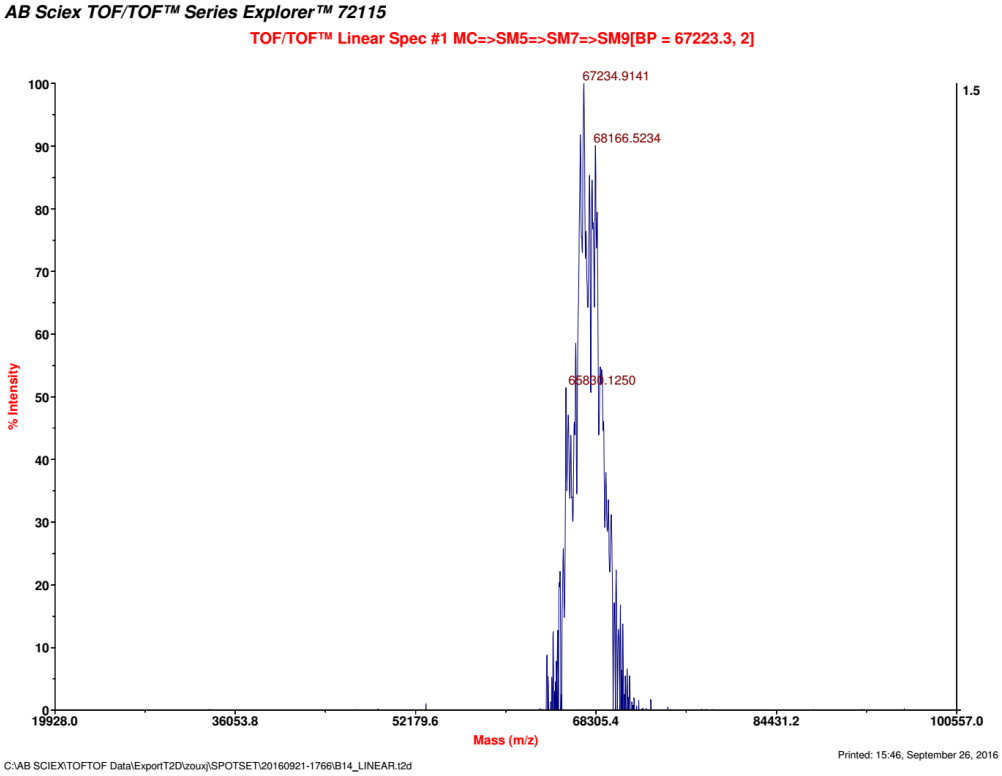


Partial HMBC of **G3**

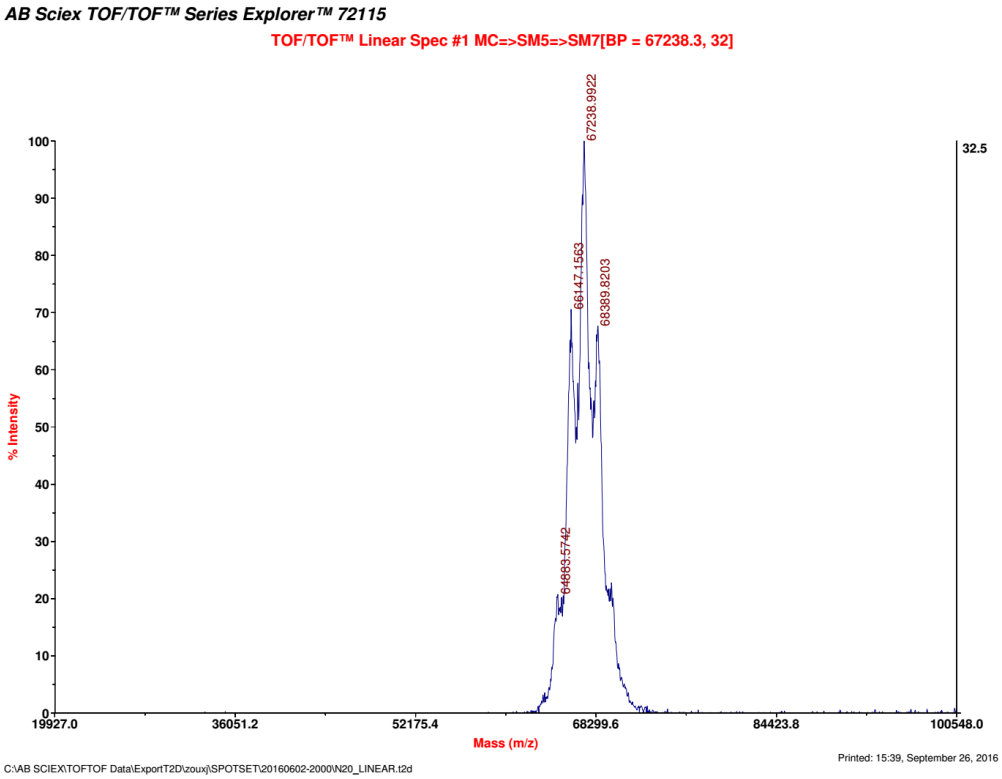
3. MALDI-TOF MS spectra and analysis of glycoconjugate



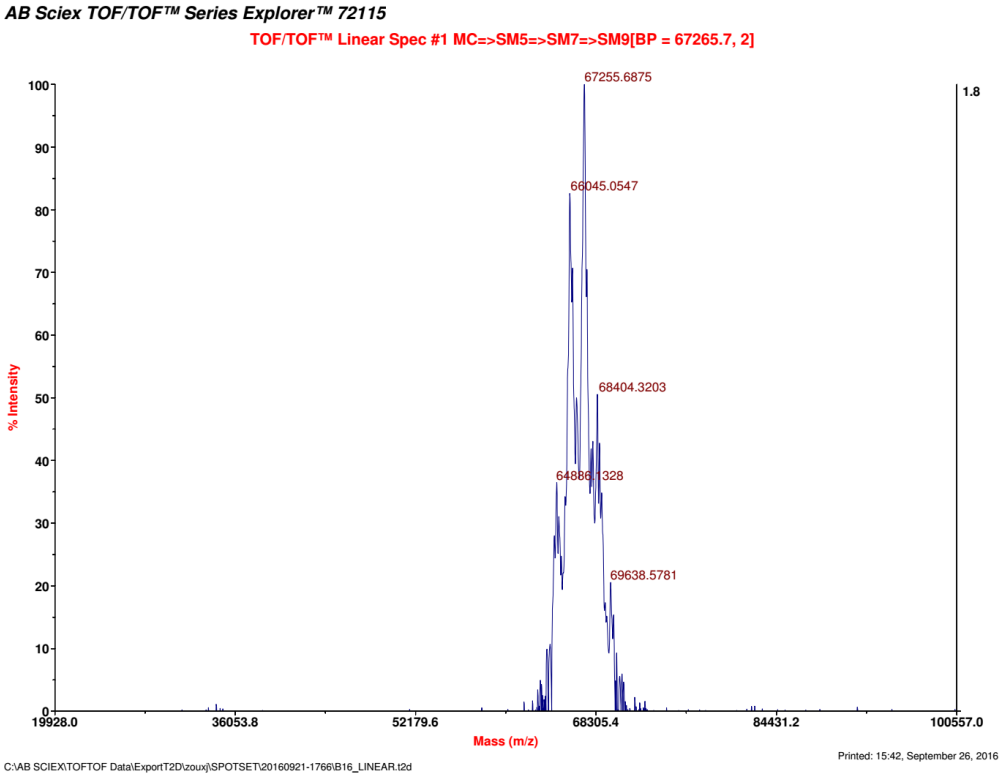
MALDI-TOF MS of **G0**-CRM197



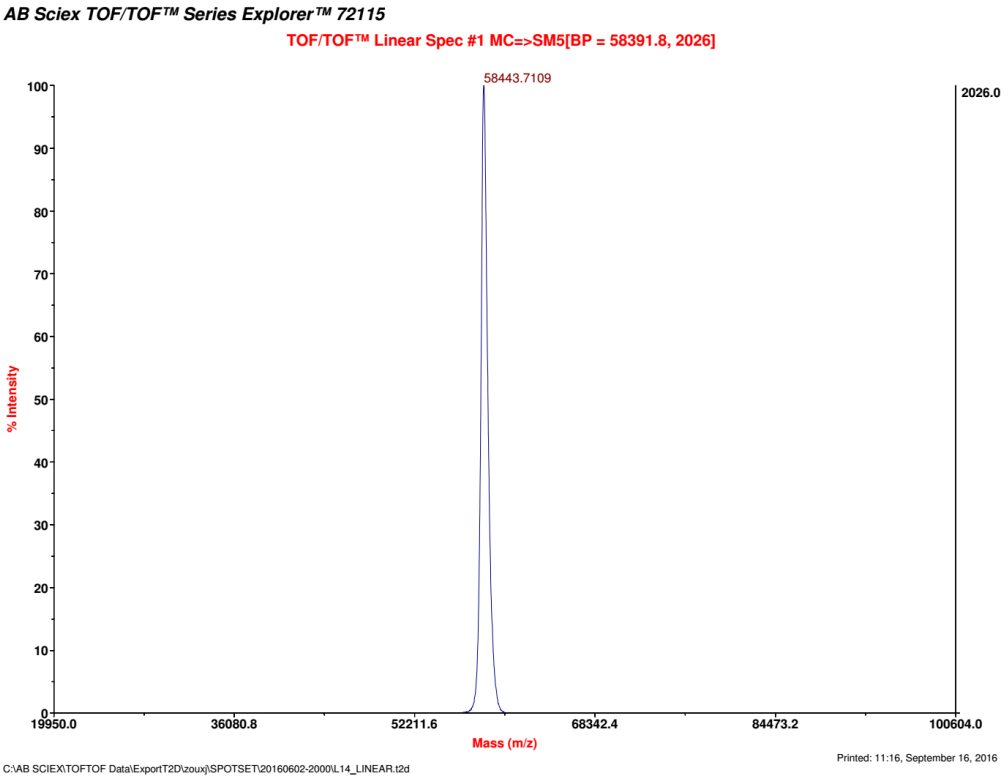
MALDI-TOF MS of **G1**-CRM197



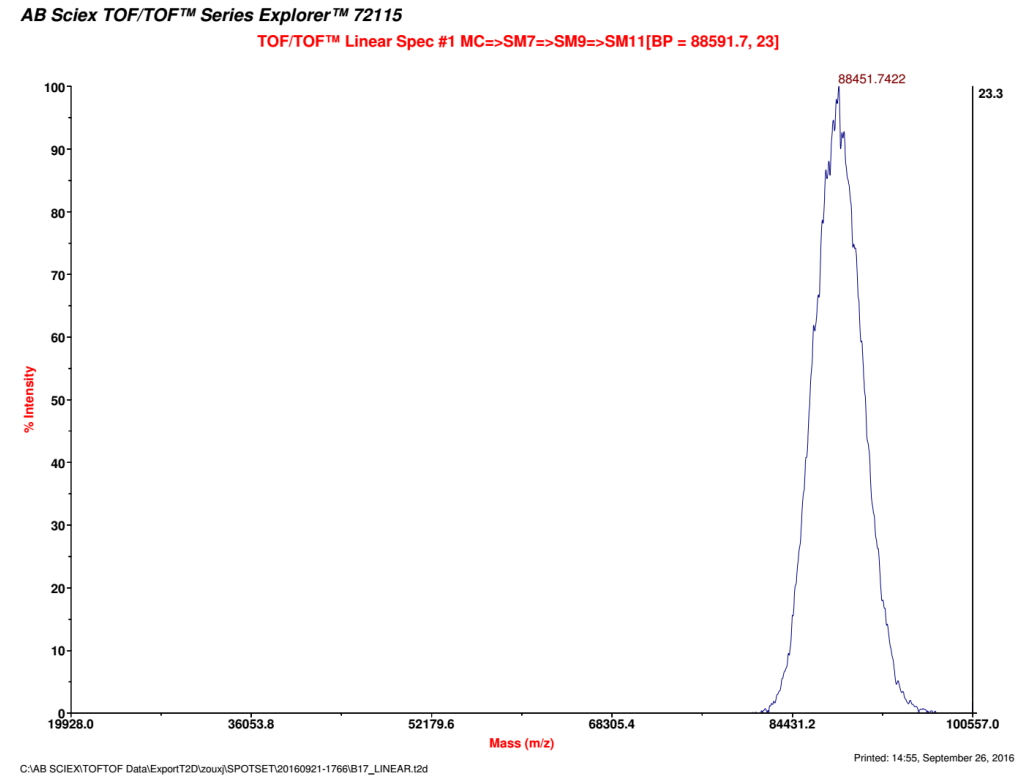
MALDI-TOF MS of **G2**-CRM197



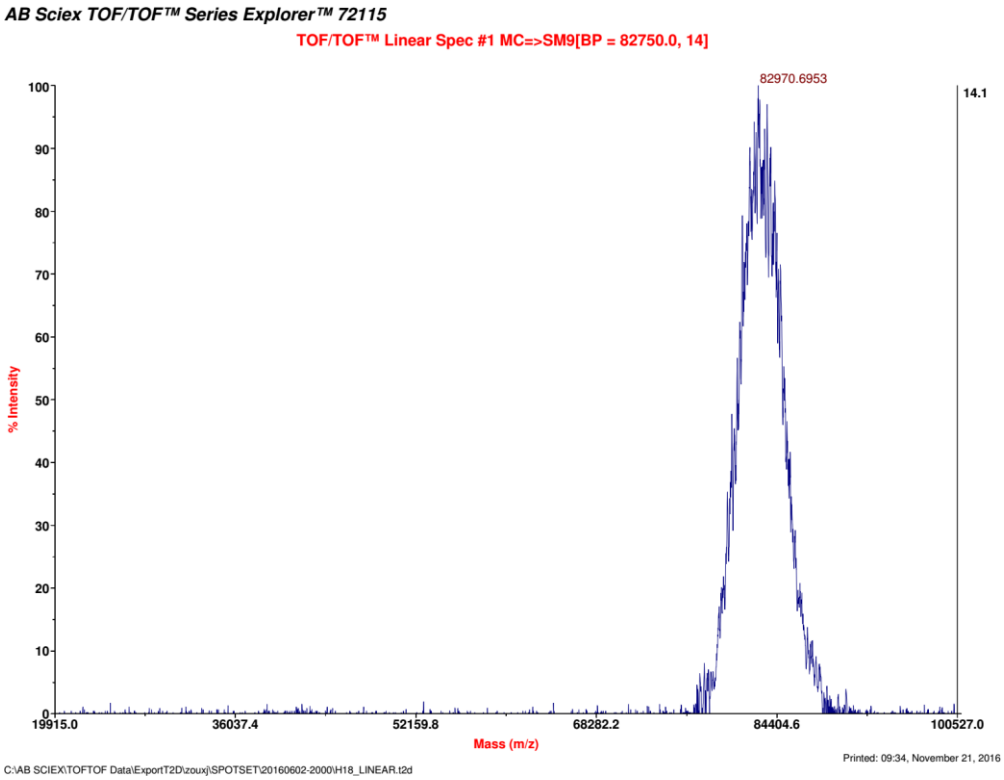
MALDI-TOF MS of **G3**-CRM197



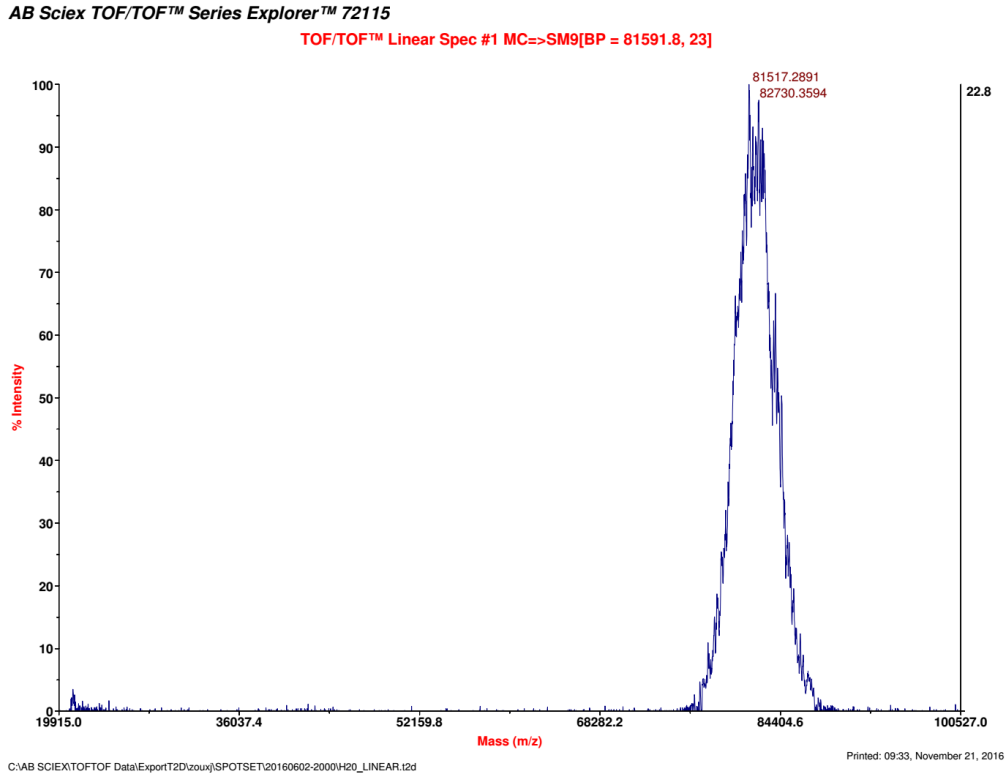
MALDI-TOF MS of CRM197



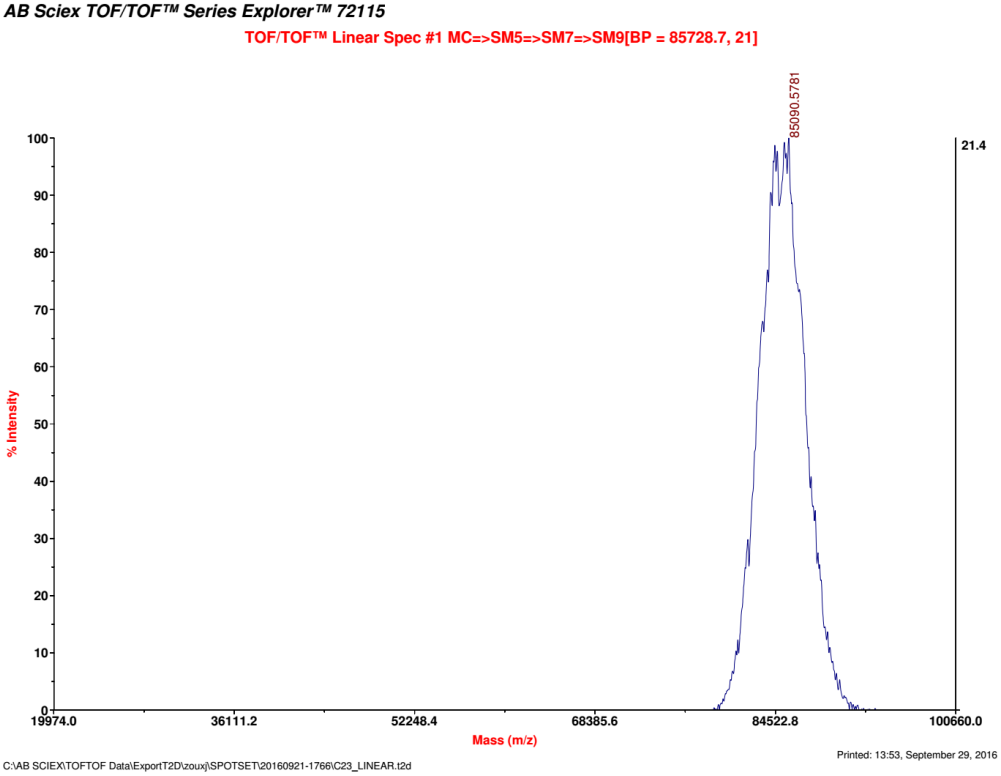
MALDI-TOF MS of **G0**-BSA



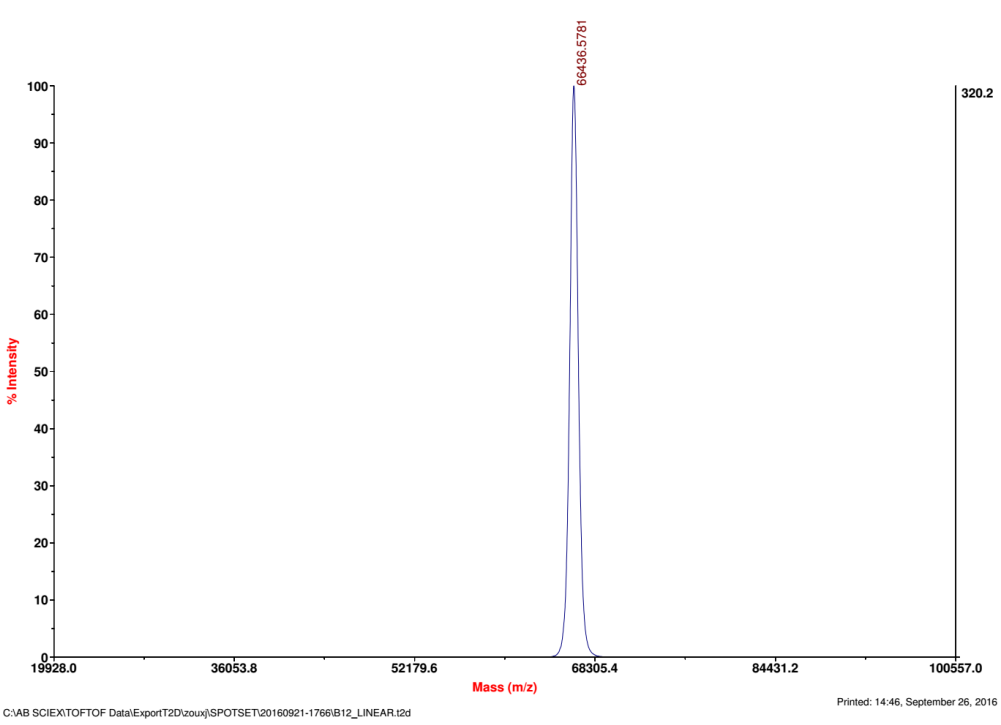
MALDI-TOF MS of **G1**-BSA



MALDI-TOF MS of **G2**-BSA



MALDI-TOF MS of **G3**-BSA



MALDI-TOF MS of BSA

Table S1 MALDI-TOF analysis of glycoconjugates

Sample	Molecular weight	Carbohydrate weight	Carbohydrate molecular weight	Average carbohydrate copies per glycoconjugate (n)	Carbohydrate percentage (%)
G0-CRM197	67552.16	9108.45	1213.19	7.51	13.48%
G1-CRM197	67234.91	8791.20	1231.18	7.14	13.08%
G2-CRM197	67238.99	8795.28	1249.17	7.04	13.08%
G3-CRM197	67255.69	8811.98	1267.16	6.95	13.10%
G0-BSA	88451.74	22015.16	1213.19	18.15	24.89%
G1-BSA	82970.70	16534.12	1231.18	13.43	19.93%
G2-BSA	81517.29	15080.71	1249.17	12.07	18.50%
G3-BSA	85090.58	18654.00	1267.16	14.72	21.92%

*Molecular weight of CRM197: 58443.71; molecular weight of BSA: 66436.58.

4. Immunological results

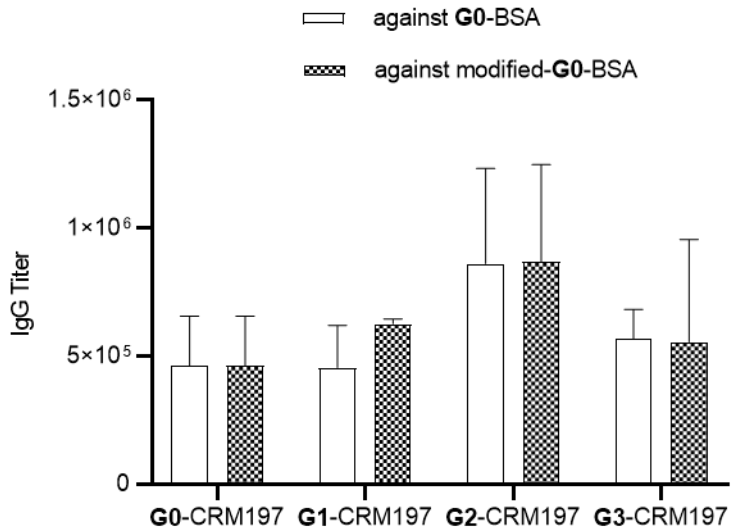


Figure S1. Pooled sera IgG titers against **G0**-BSA (hollow bar) and against modified-**G0**-BSA (dotted bar) after the 4th vaccination with **G0**-CRM197, **G1**-CRM197, **G2**-CRM197 and **G3**-CRM197. The titers were separately detected twice, and three parallel micropores were arranged. Each titer represents the average of two detections. The anti-modified-**G0** antibody titers (sera from 14 days after the 4th vaccination) were determined by ELISA, with the plate coated by the corresponding modified-**G0**-BSA conjugates instead.

Table S2 Reactivity of pooled sera IgG subclasses with **G0**-BSA by ELISA

	IgG1	**IgG2a**	**IgG2b**	**IgG3**
G0-CRM197	1.990 ± 0.138	0.063 ± 0.007	-0.016 ± 0.046	0.027 ± 0.011
G1-CRM197	2.098 ± 0.083	0.188 ± 0.016	0.102 ± 0.022	0.035 ± 0.018
G2-CRM197	2.489 ± 0.207	0.065 ± 0.008	0.315 ± 0.031	0.067 ± 0.021
G3-CRM197	2.347 ± 0.125	0.034 ± 0.021	0.038 ± 0.019	0.044 ± 0.015

Note: samples were measured by ELISA when diluted 1:10 000 for IgG and IgG1, and 1:1000 for other subclasses. Triplicate parallel wells were arranged, and data are shown with mean ± SEM of OD_{490} of them.

5. References

1 C. C. Liu, C. Zhai, X. J. Zheng and X. S. Ye, *ACS Chem. Biol.*, 2016, **11**, 1702–1709.

2 Z. Wang, L. Zhou, K. El-Boubbou, X. S. Ye and X. Huang, *J. Org. Chem.*, 2007, **72**, 6409–6420.

3 J. Feng, R. Hevey and C. C. Ling, *Carbohydr. Res.*, 2011, **346**, 2650–2662.

4 H. Y. Lee, C. Y. Chen, T. I. Tsai, S. T. Li, K. H. Lin, Y. Y. Cheng, C. T. Ren, T. J. R. Cheng, C. Y. Wu and C. H. Wong, *J. Am. Chem. Soc.*, 2014, **136**, 16844–16853.

5 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.