SUPPORTING INFORMATION

A Universal Flow Cytometry Assay for Screening Carbohydrate-Active Enzymes Using Glycan Microspheres

Aarthi Chandrasekaran,[*,a,b] Kai Deng, [a,b] Chung-Yan Koh,[b] Taichi Takasuka,[c] Lai F. Bergeman, [c] Brian G. Fox,[c] Paul D. Adams, [a,d] Anup K. Singh*[,a,b]

[a] Technology Division, Joint BioEnergy Institute, Emeryville, CA; [b] Biotechnology and Bioengineering Department, Sandia National Laboratories, Livermore, CA; [c] Great Lakes Bioenergy Research Center, University of Wisconsin, Madison, WI; [d] Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA.

* Corresponding Author Emails: achandrasekaran@lbl.gov ; aksingh@sandia.gov

Table of Contents

SUPPORTING DATA

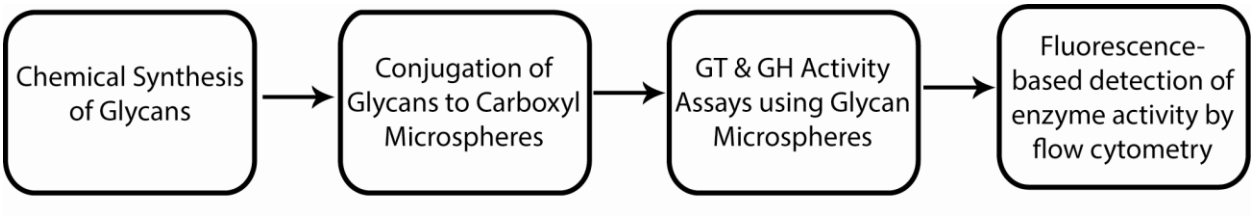


Figure S1: Process flow chart for glycan microsphere-based assay. The assay described in this study uses glycan microspheres to characterize the activities of carbohydrate-active enzymes. The assay consists of four steps: (1) Chemical synthesis of glycan substrates. Four carbohydrates were chemically synthesized for screening GTs (6-ST and GalT) and GHs (cellulases) (Figures 1 and S6). The reducing ends of the substrates were derivatized with a linker with a terminal amine group to enable coupling to microspheres. (2) The glycans were conjugated to microspheres by coupling the amine on the glycan substrates to carboxyl groups on the microspheres using coupling agents such as EDC (Figure S4, methods). These glycan microspheres serve as substrates for the different GT and GH enzymes to be tested. (3) The activity assays for GTs and GHs were performed by incubating the different enzymes with their appropriate glycan-microsphere substrates. The microspheres were then labeled to detect and quantify the products (See Figure 1 and main text for description of assay). (4) Following the activity assay and labeling steps, the fluorescence on the microsphere surface was analyzed using a standard bench-top flow cytometer.

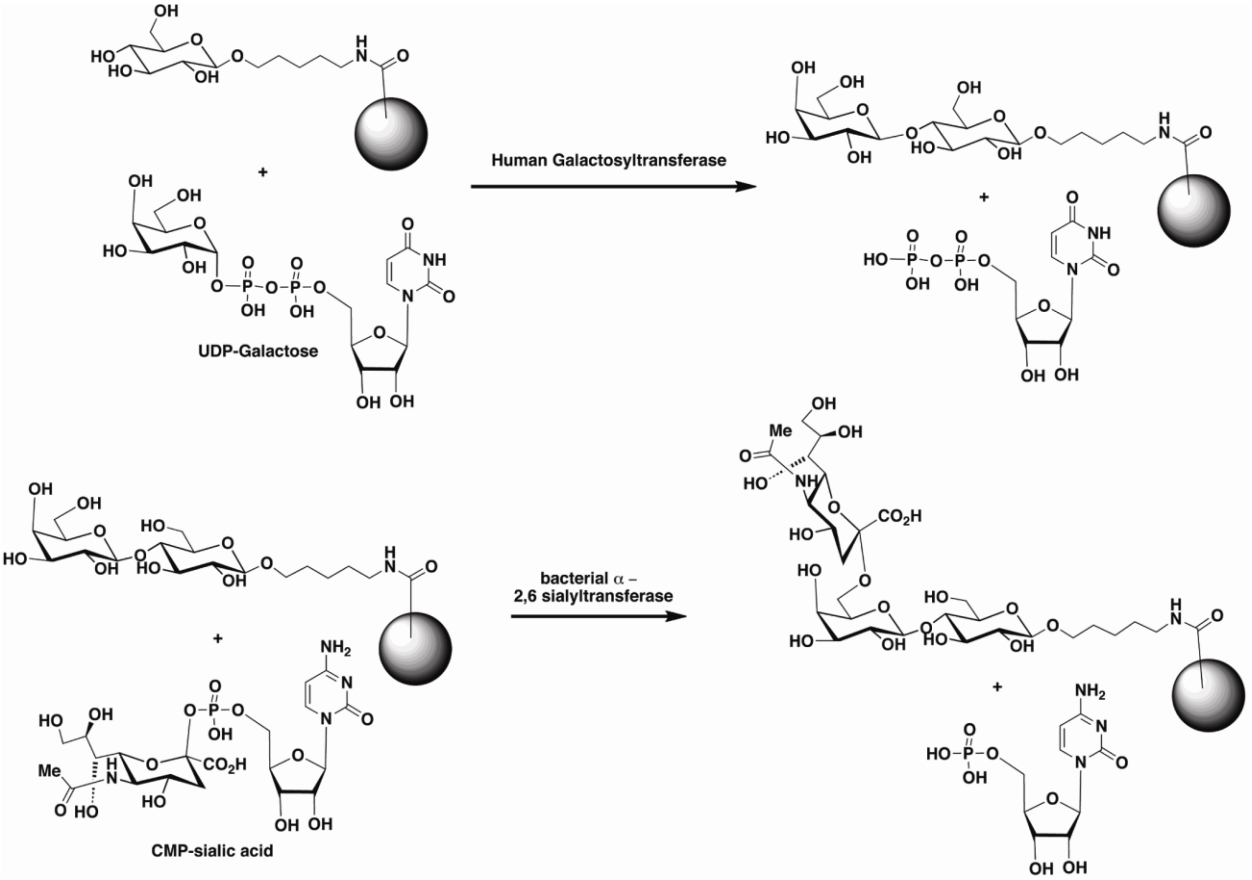


Figure S2: Schematic of enzymatic activities of GTs: GalT and 6-ST. The human galactosyltransferase (top reaction) uses UDP-galactose as the donor sugar and glucose-C5-NH_2 coupled to microspheres as the acceptor substrate. Active GalT enzyme transfers galactose monosaccharide from donor to acceptor sugar resulting in formation of Gal-β1,4-glucose (or lactose) product on the microspheres. The bacterial sialyltransferase (bottom reaction) uses CMP-sialic acid as the donor sugar and lactose- C5-NH_2 coupled to the microspheres as the acceptor substrate. In the presence of active 6-ST, the sialic acid monosaccharide (Neu5Ac) is transferred from the donor to the acceptor sugar to form Neu5Ac-α2,6-Gal-β1,4-glucose (or α2,6-sialyllactose) product on microspheres.

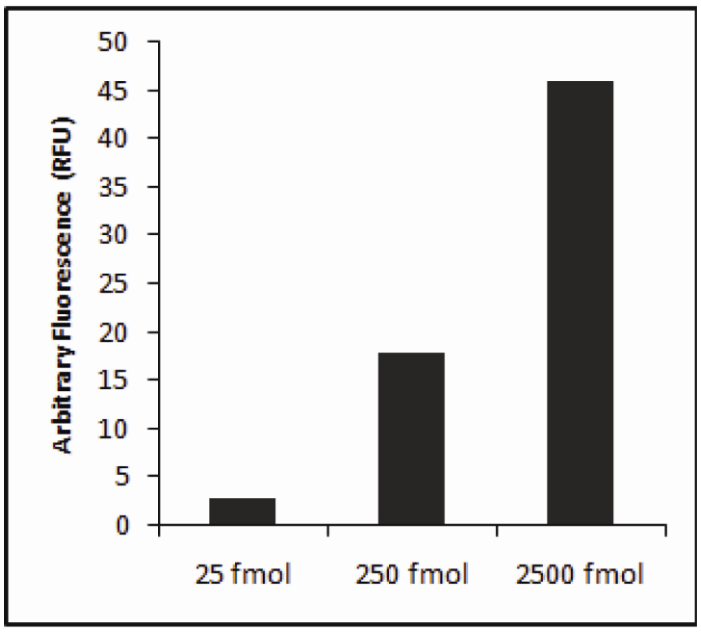


Figure S3: Detection sensitivity of 6-ST assay. In order to quantify the detection limit of the GT assay, the lactose-microspheres were incubated with different concentrations of the 6-ST enzyme for 1 hour at 37°C. After washing, the microspheres were then labeled with 5 µg/mL of fluorescein-SNA lectin to detect the formation of the product. The fluorescence of microspheres was analyzed using a BD FACScan flow cytometer. The analysis was done by collecting a total of 40,000 events per sample and fluorescence was collected on the FL1 channel (which has appropriate filter for FITC/Dylight 488 dyes). The events were then gated based on size on a forward scatter and side scatter plot. The values shown in the graph above were quantified by calculating the median fluorescence values for the size-gated events using FlowJo software and subtracting the median background fluorescence value for the no enzyme control. The GT assay used in this study has a detection limit of 25 fmol of the 6-ST enzyme. In order to demonstrate linearity at the concentrations tested, we fit the data points to a linear regression. The value of R^2 for this linear regression is 0.9279 (with equation: y=0.0335x+17.391).We also repeated this experiment with 3 micron beads to further ensure that the sensitivity of the measurement is independent of bead size. We observed linearity (R^2 = 0.9432) and the same limit of detection of 25fmol with the 3 micron beads.

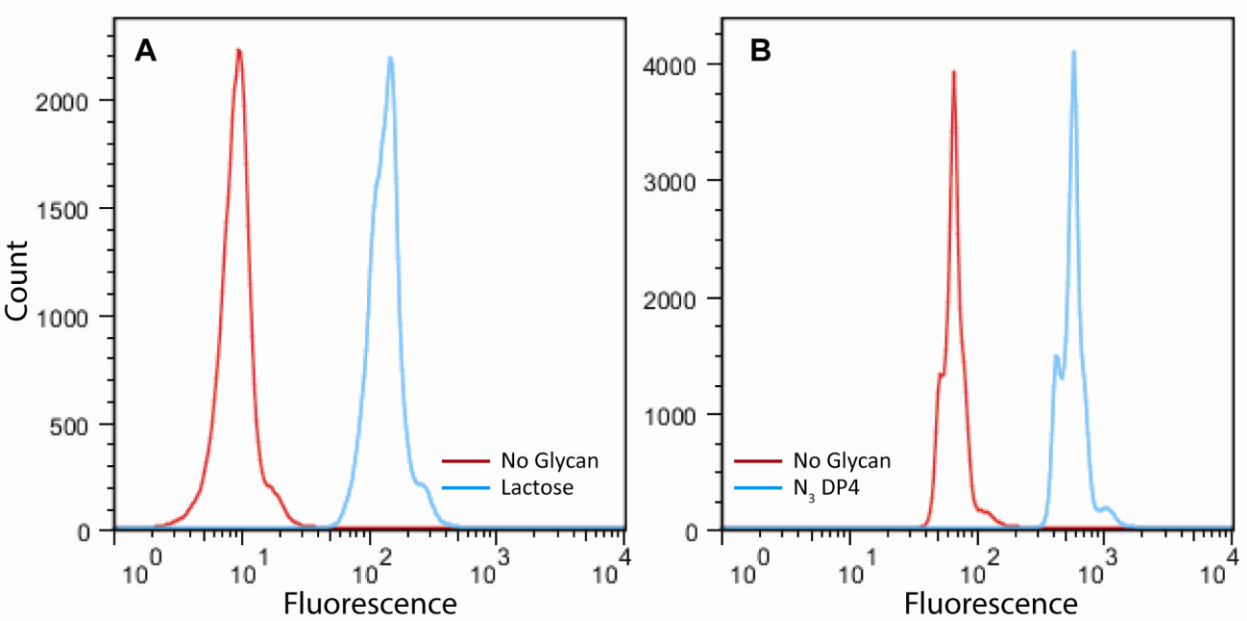


Figure S4: Validation of covalent conjugation of glycan substrates to microspheres. (A) Conjugation of GT acceptor substrate, lactose, to microspheres. The successful conjugation of lactose to the microsphere surface is confirmed using a fluorescent RCA-I lectin that binds specifically to lactose. (B) Depicts the conjugation of the GH substrate to microspheres. The azide moiety is used as a handle to verify the covalent coupling of the cellotetraose (N_3-DP4) substrate (**compound 3**) to microspheres. The azide moieties on the microsphere surface are labeled via Staudinger ligation using fluorescent phosphine probes.

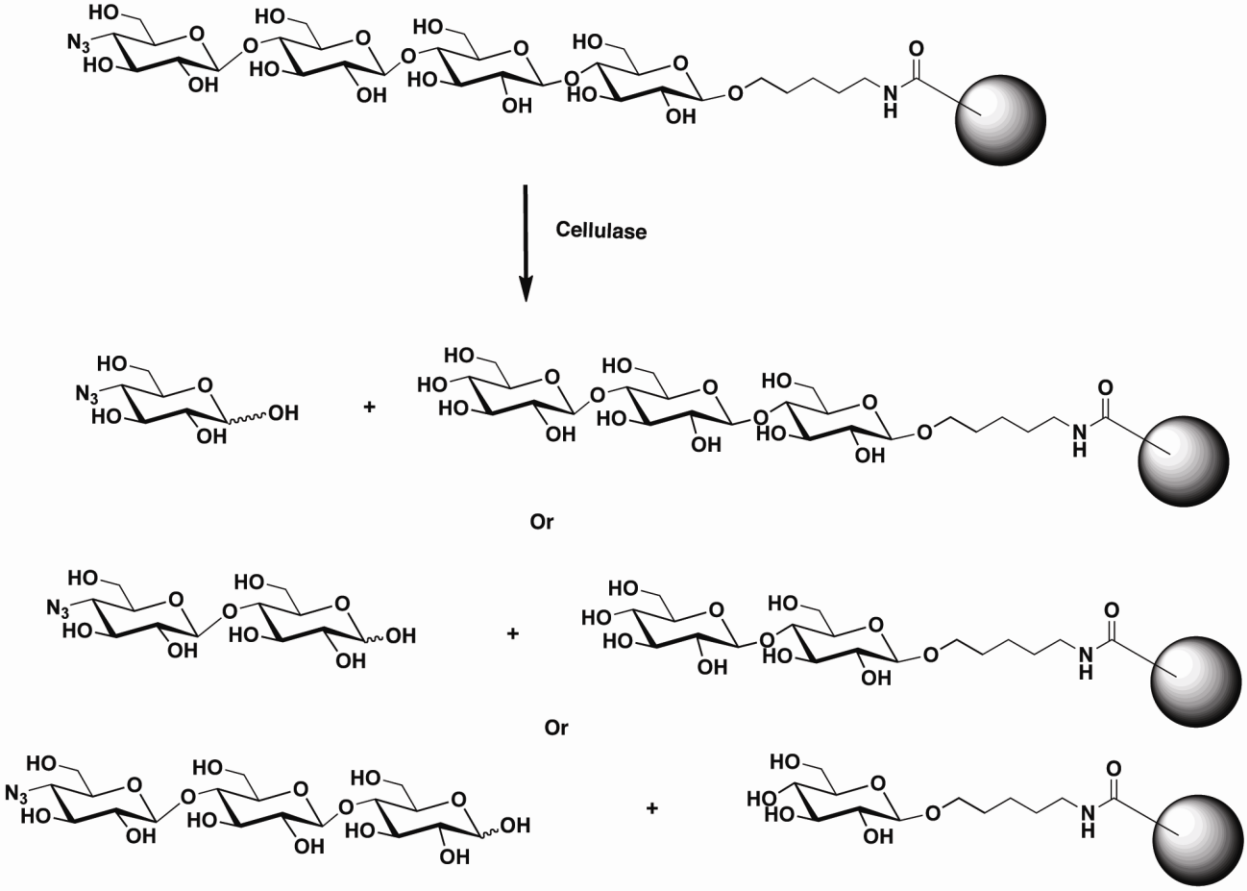


Figure S5: Schematic of cellulase activity. For the cellulase activity assay, **Compound 3** is coupled to the beads (see methods below for coupling protocol) and these glycan derivatized microspheres serve as the substrate for the different purified cellulases. The purified cellulases used in this study cleave the internal glycosidic bonds in the substrate. Therefore, their hydrolytic activity can release a combination of one or more of the following different products: azide-modified glucose, azide-modified cellobiose or azide-modified cellotetraose, from the microsphere surface. In presence of active enzymes, only glycan products without the azide handle primarily remain on the microsphere surface.

MATERIALS AND METHODS

Materials and Reagents:

1μm and 6μm carboxyl polystyrene microspheres were purchased from Polysciences, Inc. (Warrington, PA). The fluorescent lectins (Fluorescein-SNA and Fluorescein-RCA-I) were purchased from Vector Laboratories (Burlingame, CA). The Dylight488-phosphine reagent was obtained from Thermo Scientific (Rockford, IL). The recombinant bacterial *P. damsela* α2,6 sialyltransferase, UDP-Galactose, CMP-sialic acid, 1-(3-Dimethylaminopropyl)-3-ethyl carbodiimide hydochloride (EDC) and 2-(N-Morpholino)ethanesulfonic acid (MES) were purchased from Sigma-Aldrich (St. Louis, MO). The recombinant human β1,4 galactosyltransferase was obtained from R&D systems (Minneapolis, MN). Borate buffer (20X) was purchased from VWR (Radnor, PA).

Recombinant Cellulase Expression: The enzymes CelAcc_CBM3a, CelKcc, CelLcc_CBM3a, and CelRcc_CBM3a were engineered from the following *Clostridium thermocellum* ATCC 27405 genes: CelA (Cthe_0269, Uniprot A3DC29); CelK (Cthe_0412, Uniprot A3DCH1); CelL (Cthe_0405, Uniprot A3DCG4); and CelR (Cthe_0578, Uniprot A3DCY5). PCR primers were designed to remove the signal peptide and dockerin encoding sequences using the boundaries identified in the Uniprot records. The amplified sequences of the genes were ligated by FlexiVector cloning[1] into wheat germ cell-free expression plasmid engineered to produce a C-terminal fusion of the target protein to the CBM3a domain from the *C. thermocellum* scaffoldin protein encoded by Cthe_3077. The CBM3a domain was cloned in a similar manner to the enzyme catalytic domains. All protein coding sequences in the expression plasmids were verified by nucleotide sequencing. Enzyme translation was carried out using WEPRO 2240 wheat germ extract and a Protemist DT-II translation robot (Cell-Free Sciences, Yokohama, Japan) [2] Protein concentrations of translated proteins were determined by using Bio-rad Criterion gel electrophoresis system. Translated proteins were assayed in the translation reaction mixture without further purification. Enzyme assays with filter paper and phosphoric acid swollen cellulose were used to confirm enzyme activity before use in the bead-based assays.[2]

Glycan Substrate Synthesis:

General Information

The substrates used in this study were glucose-C5-amine, lactose-C5-amine, azide modified cellotetraose-C5-amine and azide modified cellobiose-C5-amine. All these compounds were synthesized chemically. Glucose, lactose and cellobiose, were purchased from Sigma-Aldrich (St. Louis, MO). All chemicals were purchased as reagent grade and used without further purification. Flash column chromatography steps were performed on a CombiFlash Rf chromatography system from Teledyne ISCO (Lincoln, NE). Reactions were monitored using analytical thin-layer chromatography (TLC) in EM silica gel 60 F254 plates and/or by staining with acidic ceric ammonium molybdate or ninhydrin. ^{1}H NMR and ^{13}C NMR spectra were recorded on a Bruker AV-600. Chemical shifts (in ppm) were assigned according to the internal standard signal of $CDCl_3$ (δ = 7.26 ppm), CD_3OD (δ = 3.31 ppm), or $CDCl_3$ (δ = 77.16 ppm) and CD_3OD (δ = 49.00 ppm) for ^{13}C NMR. Coupling constants (J) are reported in Hertz, and the splitting patterns are described by using the following abbreviations: s, singlet; d, doublet; dd, doublet of doublets; ddd, doublet of doublet of doublets; t, triplet; dt, doublet of triplets; q, quartet; m, multiple; AB, AB spin system. Infrared (IR) spectra were recorded on a Nicolet MAGNA-IR 850 spectrometer and are reported in frequency of absorption (cm^{-1}). Optical rotation was obtained in a Perkin-Elmer Polarimeter 241 at 20 °C in a 10 cm cell in the stated solvent, values are given in deg cm^2 g^{-1}. High resolution mass spectral data were obtained from the University of California, Berkeley Mass Spectral Facility.

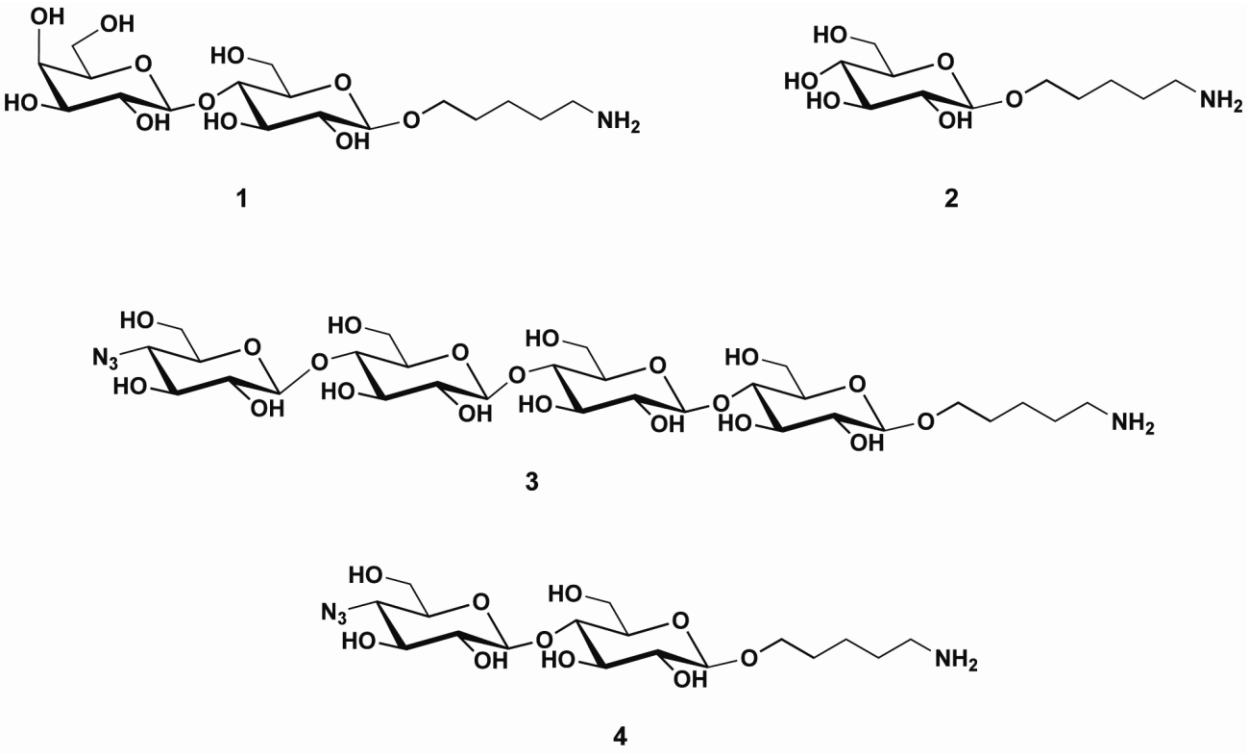


Figure S6: The structures of four glycan-C5-NH_2 substrates used in this study.

Lactose-C5-NH_2 (**Compound 1**) and glucose-C5-NH_2 (**Compound 2**) are known compounds and they are prepared according to literature procedure. [3]

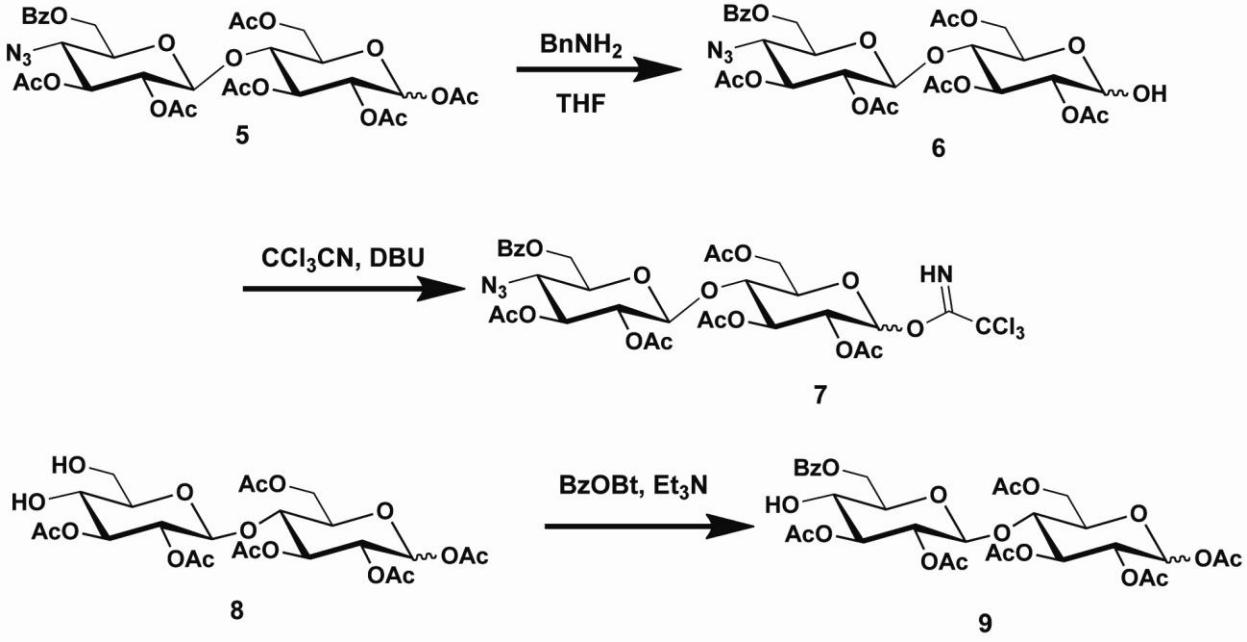


Figure S7: Synthesis of building blocks compounds 7 and 9

Synthesis of disaccharide donor (compound 7, Figure S7)

Compound 7. Benzylamine (0.84 mL, 7.7 mmol) was added to a stirred solution of **Compound 5** [4](3.70 g, 5.1 mmol) in THF (100 mL) at room temperature for 60 h. Subsequently, the solvent was removed and the resulting residue was purified by flash column chromatography to give **Compound 6** (intermediate) (3.02 g, 86% yield). To a stirred solution of **Compound 6** (630 mg, 0.92 mmol) in methylene chloride (15 mL) at 0 ºC was added CCl_3CN (0.90 mL, 9.2 mmol) followed by the addition of DBU (27 μL, 0.2 mmol). The resulting mixture was stirred for 2 h and the solvent was evaporated under reduced pressure. The residue was subjected to flash column chromatography purification to give **Compound 7** (645 mg, 85% yield). IR (thin film) 3334, 3066, 2966, 2116, 1757, 1728, 1368, 1279, 1238, 1062, 1029 cm^{-1}. ^1H NMR (600 MHz, $CDCl_3$) δ (ppm) 8.64 (s, 1H), 8.08-8.05 (m, 2H), 7.62 (m 1H), 7.52-7.49 (m, 2H), 6.46 (d, *J* = 4.2 Hz, 1H), 5.50 (t, *J* = 9.6 Hz, 1H), 5.15 (t, *J* = 9.6 Hz, 1H), 5.00 (dd, *J* = 10.2, 3.6 Hz, 1H), 4.91 (dd, *J* = 9.6, 8.4 Hz, 1H), 4.63 (dd, *J* = 12.0, 1.8 Hz, 1H), 4.57-4.50 (m, 3H), 4.13-4.06 (m, 2H), 3.81 (t, *J* = 9.6 Hz, 1H), 3.75 (t, *J* = 10.2 Hz, 1H), 3.55 (ddd, *J* = 10.2, 4.8, 2.4 Hz, 1H), 2.084 (s, 3H), 2.080 (s, 3H), 2.03 (s, 3H), 1.97(S, 3H), 1.89 (S, 3H).

^{13}C NMR (125 MHz, $CDCl_3$) 171.0, 170.1, 169.9, 169.8, 169.4, 165.9, 160.9, 133.6, 129.6, 129.2, 128.7, 100.9, 92.8, 90.6, 76.2, 74.1, 72.4, 71.8, 71.0, 69.8, 69.2, 63.1, 61.2, 59.9, 20.7, 20.6, 20.5, 20.3.

$[\alpha]_D^{20}$ +70.8 (c 1, $CHCl_3$).

HRMS (ESI) *m/z*: Calcd for $C_{31}H_{35}Cl_3N_4O_{16}$ ($M+Na^+$) 847.1006, found 847.0994.

Synthesis of disaccharide acceptor (compound 9, Figure S7)

Compound 9. To a stirred solution of **compound 8** [5] (1.53g, 2.58 mmol) in CH_2Cl_2 (20 mL) was added 1-benzoyloxybenzotriazole[6] (680 mg, 2.83 mmol), followed by the addition of triethylamine (0.61 mL, 4.40 mmol). The resulting mixture was stirred at room temperature for 12 h and the solvent was evaporated under reduced pressure. The resulting residue was subjected to column purification to give 1.60g **Compound 9** in 88% yield. IR (thin film) 3487, 3023, 2962, 2116, 1752, 1716, 1370, 1238, 1209, 1052, 756, 716 cm^{-1}. ^{1}H NMR (600 MHz, $CDCl_3$) δ (ppm) 8.07-8.03 (m, 2H), 7.64-7.60 (m 1H), 7.52-7.47 (m, 2H), 6.24 (d, *J* = 4.2 Hz, 0.43 H, α isomer), 5.65 (d, *J* = 7.8 Hz, 0.57 H, β isomer), 5.43 (d, *J* = 10.2, 9.6 Hz, 0.43 H, α isomer), 5.22 (dd, *J* = 9.6, 9.6 Hz, 0.57 H, β isomer), 5.06-4.97 (m, 2H), 4.93-4.86 (m, 1H), 4.78-4.74 (m, 1H), 4.57-4.52 (m, 1H), 4.52-4.47 (m, 2H), 4.14-4.09 (m, 1H), 3.99 (m, 0.43 H, α isomer), 3.85-3.78 (m, 1H), 3.73 (m, 0.53 H, β isomer), 3.69-3.62 (m, 1H), 3.62-3.57 (m, 1H), 3.29-3.23 (br, 1H), 2.16(α), 2.11 (α), 2.103 (β), 2.091 (β), 2.048 (β), 2.046(α), 2.037 (α), 2.033 (β), 2.007 (β), 2.005 (α), 1.994 (β),1.991 (α). 13C NMR (125 MHz, CDCl3) 170.29 (2C), 170.27, 169.8, 169.7 (2C), 169.4, 169.3, 169.2, 168.9, 168.8, 133.62, 133.59, 129.70, 129.69, 129.2, 128.6, 128.62, 128.61, 101.0, 100.8, 91.5, 88.9, 76.1, 76.0, 75.5, 75.4, 74.3, 73.5, 72.4, 71.6, 71.5, 70.8, 70.4, 69.31, 69.27, 68.47, 68.43, 63.18, 63.12, 61.6, 61.4, 20.8, 20.72 (2C), 20.71, 20.70 (2C), 20.62, 20.56, 20.53, 20.49, 20.47, 20.38.

$[\alpha]_D^{20}$ +5.6 ° (c 1, $CHCl_3$).

HRMS (ESI) *m/z*: Calcd for $C_{31}H_{38}O_{18}$ ($M+Na^+$) 721.1950, found 721.1938.

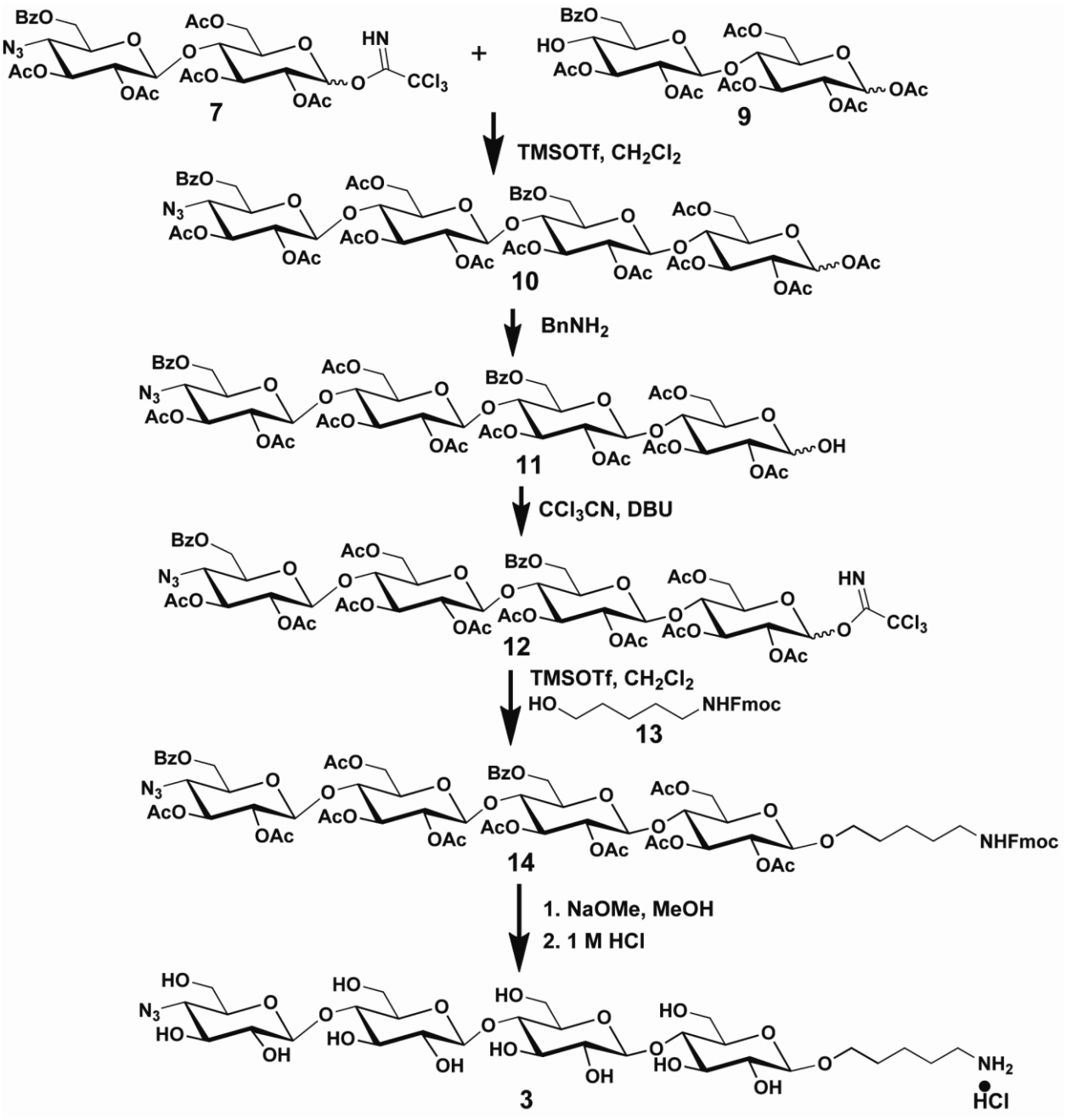


Figure S8: Synthesis of azide modified cellotetraose compound 3

Synthesis of tetrasaccharide 3 (Figure S8)

Compound 10. Disaccahride imidate **Compound 7** (440 mg, 0.53 mmol) and disaccaharide acceptor **Compound 9** (480 mg, 0.69 mmol) were mixed in 20 mL of anhydrous methylene chloride under nitrogen. Then anhydrous 3Å molecular sieve was added. After the resulting mixture was stirred at room temperature for 0.5 h, it was cooled down to -20ºC, followed by

injection of TMSOTf (30 μL, 0.16 mmol). After the reaction mixture was stirred at this temperature for 1h, triethylamine (0.1 mL) was added to quench the reaction. The solvent was removed under reduced pressure and the residue was purified by flash column chromatography to give 530 mg of **Compound 10** in 73% yield as a mixture of α and β anomers with (α/β = 0.43:0.57). With careful column purification, a small amount of pure α and β anomers can be separated for analysis.

Alfa-anomer: IR (thin film) 3485, 3022, 2962, 2921, 2845, 2113, 1755, 1369, 1236, 1053, 801, 758, 716 cm^{-1}. ^{1}H NMR (600 MHz, $CDCl_3$) δ (ppm) 8.08-8.03 (m, 4H), 7.66-7.59 (m 2H), 7.56-7.49 (m, 4H), 6.24 (d, *J* = 3.6 Hz, 1H), 5.43-5.38 (t, *J* = 10.2 Hz, 1H), 5.15 (t, *J* = 9.6 Hz, 1H), 5.11 (t, *J* = 9.6 Hz, 1H), 5.01 (t, *J* = 9.6 Hz, 1H), 4.97-4.94 (dd, *J* = 10.2, 3.6 Hz, 1H), 4.91 (t, *J* = 8.4 Hz, 1H), 4.84 (t, *J* = 8.4 Hz, 1H), 4.81 (t, *J* = 7.8 Hz, 1H), 4.73 (d, *J* = 11.4 Hz, 1H), 4.62-4.58 (dd, *J* = 12.6, 1.8 Hz, 1H), 4.52-4.46 (m, 3H), 4.46-4.42 (d, *J* = 8.4 Hz, 2H), 4.43-4.37 (d, *J* = 10.8 Hz, 1H), 4.29-4.24 (dd, *J* = 12.0, 5.4 Hz, 1H), 4.10-4.03 (dt, *J* = 12.0, 4.2 Hz, 2H), 4.00-3.96 (m, 1H), 3.85 (t, *J* = 8.4 Hz, 1H), 3.76-3.67 (m, 4H), 3.51-3.47 (ddd, *J* = 10.2, 4.2, 1.8 Hz, 1H), 3.46-3.42 (ddd, *J* = 10.2, 4.8, 1.2 Hz, 1H), 2.15 (S, 3H), 2.11 (S, 3H), 2.09 (S, 3H), 2.08 (2 CH_3), 2.04 (S, 3H), 2.02 (S, 3H), 2.01 (S, 3H), 1.97 (S, 3H), 1.96 (S, 3H) 1.87 (s, 3H), 1.86 (s, 3H).
^{13}C NMR (125 MHz, $CDCl_3$) 170.11, 170.06, 169.8, 169.68, 169.65, 169.63, 169.30, 169.28, 168.8, 165.8, 165.6, 133.7, 133.6, 129.62 (2C), 129.55 (2C), 129.2, 129.1, 128.8 (2C), 128.7 (2C), 100.8, 100.6, 100.3, 88.9, 76.11, 76.07, 75.99, 73.9, 73.0, 72.75, 72.62, 72.38, 72.37, 71.9, 71.7, 70.8, 69.28, 69.1, 63.5, 62.9, 62.5, 61.9, 61.1, 59.9, 20.84, 20.73, 20.71, 20.56, 20.53, 20.52, 20.51, 20.47, 20.44, 20.41, 20.37.
$[\alpha]_D^{20}$ +11.2 (c 0.25, $CHCl_3$).

Beta-anomer: IR (thin film) 3481, 3025, 2962, 2868, 2113, 1755, 1368, 1234, 1054, 801, 757, 716 cm^{-1}. ^{1}H NMR (600 MHz, $CDCl_3$) δ (ppm) 8.08-8.02 (m, 4H), 7.66-7.59 (m 2H), 7.56-7.46 (m, 4H), 5.63 (d, *J* = 8.4 Hz, 1H), 5.20 (t, *J* = 9.0 Hz, 1H), 5.15-5.08 (m, 2H), 5.03-4.99 (m, 2H), 4.90-4.87 (dd, *J* = 9.6, 8.4 Hz, 1H), 4.86-4.82 (dd, *J* = 9.6, 6.5 Hz, 1H), 4.82-4.78 (dd, *J* = 9.0, 6.5 Hz, 1H), 4.74-4.70 (m, 1H), 4.61-4.57 (m, 1H), 4.52-4.45 (m, 4H), 4.45-4.41 (dd, *J* = 7.8, 2.4 Hz, 1H), 4.41-4.36 (dd, *J* = 12.0, 1.8 Hz, 1H), 4.27-4.23 (dd, *J* = 12.0, 5.4 Hz, 1H), 4.10-4.03 (m,

2H), 3.86-3.82 (t, *J* = 9.6 Hz, 1H), 3.79-3.74 (m, 1H), 3.73-3.66 (m, 4H), 3.51-3.46 (ddd, *J* = 10.2, 4.2, 1.8 Hz, 1H), 3.46-3.42 (ddd, *J* = 10.2, 4.8, 1.8 Hz, 1H), 2.104 (S, 3H), 2..087 (S, 3H), 2.080 (S, 3H), 2.07 (2 CH_3), 2.021 (S, 3H), 2.016 (S, 3H), 2.009 (S, 3H), 1.98 (S, 3H), 1.95 (S, 3H) 1.86 (s, 3H), 1.85 (s, 3H).

^{13}C NMR (125 MHz, $CDCl_3$) 170.14, 170.07, 169.8, 169.65, 169.63, 169.60, 169.3 (2C), 169.24, 169.20, 168.7, 133.7, 133.6, 129.61, 129.58, 129.57, 129.2, 129.1, 128.8, 128.69, 128.65, 100.7, 100.6, 100.3, 91.6, 76.01, 75.95, 73.88, 73.55, 73.0, 72.7, 72.5, 72.39, 72.35, , 72.21, 71.8, 71.7, 70.4, 64.3, 62.9, 62.4, 61.9, 61.4, 59.9, 20.75, 20.73, 20.70, 20.53, 20.51, 20.49, 20.47, 20.46, 20.44, 20.41, 20.405.

$[\alpha]_D^{20}$ +23.1 (c 0.8, $CHCl_3$).

HRMS (ESI) *m/z*: Calcd for $C_{60}H_{71}N_3O_{33}$ ($M+Na^+$) 1384.3862, found 1384.3848.

Compound 12. Benzylamine (62 μL, 0.55 mmol) was added to a stirred solution of **Compound 10** (500 mg, 0.37 mmol) in THF (20 mL) at room temperature and the resulting mixture was stirred for 60 h. Subsequently, the solvent was removed and the resulting residue was purified by flash column chromatography to give **Compound 11** (intermediate) (410 mg, 85% yield).

To a stirred solution of **Compound 11** (intermediate) (360 mg, 0.27 mmol) in methylene chloride (15 mL) at 0 ºC was added CCl_3CN (0.27 mL, 2.70 mmol) followed by the addition of DBU (8.4 μL, 0.056 mmol). The resulting mixture was stirred for 2 h and the solvent was evaporated under reduced pressure. The residue was subjected to flash column chromatography purification to give **Compound 12** (292 mg, 73% yield).

IR (thin film) 3337, 3031, 2961, 2113, 1755, 1369, 1234, 1056, 798, 756, 716 cm^{-1}. ^1H NMR (600 MHz, $CDCl_3$) δ (ppm) 8.67 (s, 1H), 8.10-8.06 (m, 4H), 7.69-7.64 (m 2H), 7.57-7.53 (m, 4H), 6.49 (d, *J* = 4.2 Hz, 1H), 5.52 (t, *J* = 9.6 Hz, 1H), 5.17 (t, *J* = 9.0 Hz, 1H), 5.14 (t, *J* = 7.2 Hz, 1H), 5.06-5.01 (m, 2H), 4.94 (dd, *J* = 9.6, 8.4 Hz, 1H), 4.88 (dd, *J* = 9.6, 8.4 Hz, 1H), 4.83 (dd, *J* = 9.6, 8.4 Hz, 1H), 4.75 (dd, *J* = 12.0, 1.8 Hz, 1H), 4.62 (dd, *J* = 12.6, 2.4 Hz, 1H), 4.57-4.49 (m, 3H), 4.46 (m, 2H), 4.42 (dd, *J* = 12.0, 1.8 Hz, 1H), 4.31 (dd, *J* = 12.6, 5.4 Hz, 1H), 4.15-4.06 (m, 3H), 3.88 (t, *J* = 9.6 Hz, 1H), 3.82 (t, *J* = 9.6 Hz, 1H), 3.77-3.70 (m, 3H), 3.51 (ddd, *J* = 10.2, 4.2, 3.6 Hz, 1H), 3.47 (ddd, *J* = 7.8, 4.8, 1.8 Hz, 1H), 2.14 (S, 3H), 2.11(S, 3H), 2.10(S, 3H), 2.5 (2 CH_3), 2.04(S, 3H), 2.00(S, 3H), 1.98(S, 3H), 1.90(S, 3H), 1.88(S, 3H).

^{13}C NMR (125 MHz, $CDCl_3$) 170.08, 170.07, 169.84, 169.82, 169.71, 169.64, 169.43, 169.31, 169.26, 169.23, 165.8, 165.6, 160.9, 133.7, 133.6, 129.61, 129.55, 129.22, 129.14, 128.76, 128.69, 100.9, 100.6, 100.3, 92.8, 90.6, 76.2, 76.08, 76.00, 73.9, 73.0, 72.7 (2C), 72.39, 72.35, 70.9, 71.7 (2C), 71.0, 69.8, 69.1, 62.9, 62.5, 61.9, 61.2, 59.9, 20.72, 20.70, 20.56, 20.53 ($2CH_3$), 20.51, 20.47, 20.45, 20.4, 20.3.

$[\alpha]_D^{20}$ +33.8 (c 1, $CHCl_3$).

HRMS (ESI) *m/z*: Calcd for $C_{60}H_{69}Cl_3N_4O_{32}$ ($M+Na^+$) 1485.2853, found 1485.2867.

Compound 14. Tetrasaccharide imidate **compound 12** (101 mg, 0.069 mmol) and **5-(Fmoc-amino)-1-pentanol** (**13**, 30 mg, 0.090 mmol) were mixed in 5 mL of anhydrous methylene chloride under nitrogen. Then 3 Å molecular sieves were added. After the resulting mixture was stirred at room temperature for 0.5 h, it was cooled down to -20 ºC, followed by injection of TMSOTf (10 μL, 0.055 mmol). After the reaction mixture was stirred at this temperature for 1h, triethylamine (50 μL) was added to quench this reaction. The solvent was removed under reduced pressure and the residue was purified by flash column chromatography to give **Compound 14** (40 mg, 36% yield). IR (thin film) 3379, 3068, 2942, 2872, 2108, 1761, 1720, 1451, 1373, 1275, 1234, 1123, 1050, 743, 707 cm^{-1}. ^{1}H NMR (600 MHz, $CDCl_3$) δ (ppm) 8.07-8.03 (m, 4H), 7.77-7.74 (d, *J* = 6.8 Hz, 2H), 7.64-7.60 (m, 2H), 7.60-7.57 (d, *J* = 7.2 Hz, 2H), 7.55-7.49 (m, 4H), 7.41-7.37 (dd, d, *J* = 7.2, 7.2 Hz, 2H), 7.32-7.28 (dt, *J* = 7.2, 0.6 Hz, 2H), 5.18-5.08 (m, 3H), (5.01 t, *J* = 9.0 Hz, 1H), 4.91-4.78 (m, 5H), 4.62-4.57 (dd, *J* = 11.4, 1.8 Hz, 1H), 4.57-4.52 (dd, *J* = 12.0, 2.4 Hz, 1H), 4.50-4.45 (m, 2H), 4.45-4.41 (dd, *J* = 7.8, 5.4 Hz, 1H), 4.50-4.45 (m, 2H), 4.40-4.36 (m, 4H), 4.26-4.18 (m, 2H), 4.08-4.03 (m, 2H), 3.84-3.76 (m, 2H), 3.75-3.64 (m, 4H), 3.57-3.52 (ddd, *J* = 9.6, 4.2, 1.8 Hz, 1H), 3.51-3.47 (ddd, *J* = 10.2, 4.2, 1.8 Hz, 1H), 3.46-3.40 (m, 2H), 3.20-3.13 (m, 2H), 2.10 (s, 3H), 2.08 (s, 3H), 2.07 (s, 3H), 2.02 (s, 9H), 1.98 (s, 3H), 1.95 (s, 3H), 1.86 (s, 3H), 1.85 (s, 3H), 1.60-1.52 (m, 2H), 1.52-1.45 (m, 2H), 1.40-1.30 (m, 2H).

^{13}C NMR (125 MHz, $CDCl_3$) 170.3, 170.1, 169.88, 169.87, 169.71, 169.69, 169.44, 169.36, 169.30, 169.28, 165.89, 165.7, 156.4, 144.0, 141.3, 133.72, 133.68, 129.67, 129.61, 129.26, 129.18, 128.84, 128.75, 127.66, 127.01, 125.0, 120.0, 100.66, 100.61, 100.3, 76.6, 76.2, 76.1, 73.9, 73.0, 77.78, 77.77, 72.66, 72.43, 72.39, 72.20, 71.9, 71.76, 71.73, 71.58, 69.8, 66.51, 66.50,

64.36, 62.96, 62.5, 61.9, 61.6, 59.9, 47.3, 40.9, 30.7, 29.5, 28.96, 23.0, 21.0, 20.86, 20.76, 20.64, 20.60, 20.57, 20.53, 20.50, 20.47, 19.1, 14.2, 13.7.

$[\alpha]_D^{20}$ +7.5 (c 1, $CHCl_3$).

HRMS (ESI) *m/z*: Calcd for $C_{78}H_{90}N_4O_{34}$ ($M+Na^+$) 1649.5329, found 1649.5284.

Compound 3. NaOMe in MeOH (9 μL; 25% w/w) was added to **Compound 14** (21 mg, 0.013 mmol) in methanol (1.5 mL). After the resulting mixture was stirred at room temperature for 20 h, 1N HCl (0.2 mL) was added. The solvent was evaporated under reduced pressure to give a white solid substance, which was purifed by Combiflash to give (7.8 mg; **Compound 3**) in 78% yield. IR (thin film) 3357, 2926, 2873, 2113, 1153, 1065, 1024 cm^{-1}. ^1H NMR (600 MHz, CD_3OD) δ (ppm) 4.46 (d, *J* = 8.4 Hz, 1H), 4.44 (d, *J* = 7.8 Hz, 1H), 4.39 (d, *J* = 7.8 Hz, 1H), 4.30 (d, *J* = 8.4 Hz, 1H), 3.94-3.83 (m, 5H), 3.82-3.79 (dd, *J* = 10.2, 1.8 Hz, 1H), 3.71-3.66 (m, 5H), 3.82-3.79 (dd, *J* = 10.2, 1.8 Hz, 1H), 3.71-3.66 (dd, *J* = 12.6, 4.8 Hz, 1H), 3.61-3.45 (m, 9H), 3.43-3.36 (m, 2H), 3.33-3.20 (m, 6H), 2.96-2.91 (m, 2H), 1.73-1.65 (m, 4H), 1.55-1.48 (m, 2H). ^{13}C NMR (125 MHz, CD_3OD) 102.1, 102.01, 101.98, 101.8, 78.3, 77.71, 77.69, 74.8, 74.3, 74.10, 74.09, 73.9, 73.78, 73.76, 72.7, 72.6, 72.31, 72.27, 67.9, 60.8, 60.1, 59.9, 59.5, 59.1, 38.4, 30.5, 27.6, 25.8, 21.6.

$[\alpha]_D^{20}$ +21.6 (c 0.8, MeOH).

HRMS (ESI) *m/z*: Calcd for $C_{29}H_{52}N_4O_{20}$ ($M+H^+$) 777.3248, found 777.3238.

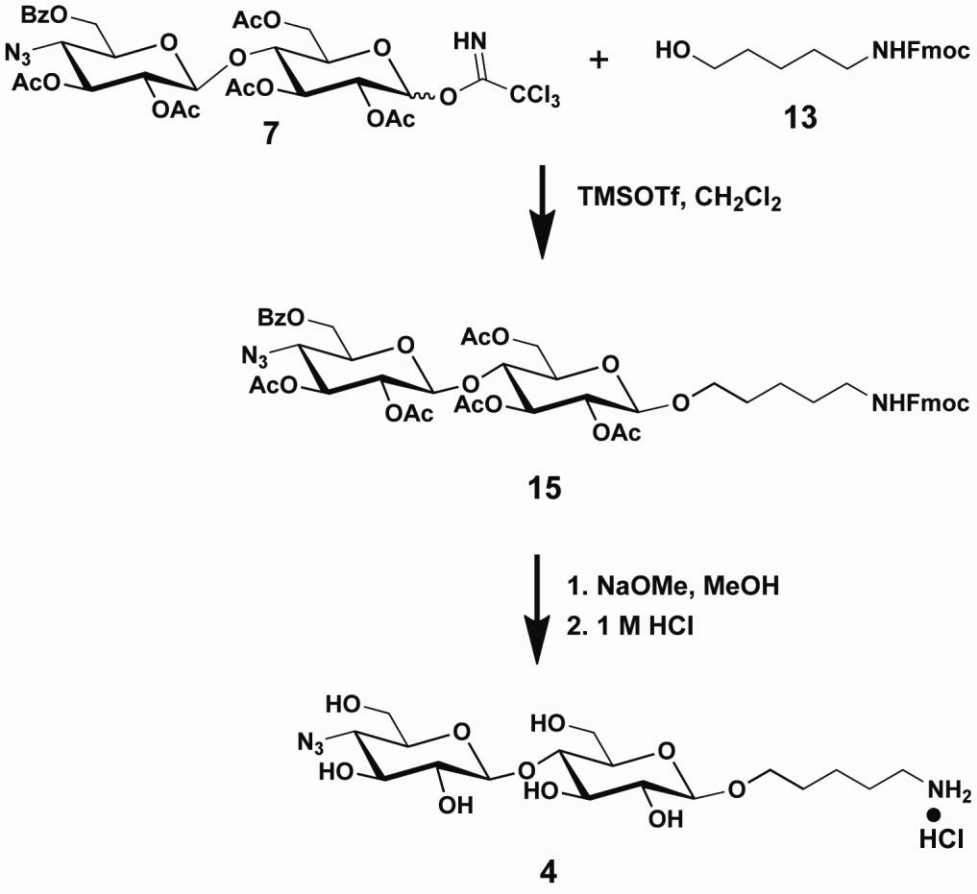


Figure S9: Synthesis of azide modified cellobiose compound 4

Synthesis of disaccharide 4 (Figure S9)

Compound 15. Disaccharide imidate **Compound 7** (154 mg, 0.186 mmol) and **5-(Fmoc-amino)-1-pentanol** (**13**, 79 mg, 0.24 mmol) were mixed in 7.0 mL of anhydrous methylene chloride under nitrogen. Then 3 Å molecular sieves were added. After the resulting mixture was stirred at room temperature for 0.5 h, it was cooled down to -20ºC, followed by injection of TMSOTf (27 μL, 0.15 mmol). After the reaction mixture was stirred at this temperature for 1h, triethylamine (50 μL) was added to quench this reaction. The solvent was removed under reduced pressure and the residue was purified by flash column chromatography to give **Compound 15** (72 mg, 39% yield). IR (thin film) 3393, 3015, 2941, 2864, 2112, 1755, 1707, 1522, 1451, 1368, 1235, 1054, 759, 715 cm^{-1}. ^1H NMR (600 MHz, $CDCl_3$) δ (ppm) 8.07-8.04 (m, 2H), 7.77-7.74 (d, *J* = 7.8 Hz, 2H), 7.64-7.60 (m, 1H), 7.60-7.56 (d, *J* = 7.2 Hz, 2H), 7.55-7.50 (m, 2H), 7.42-7.37 (dd, d, *J* = 7.2, 7.2 Hz, 2H), 7.32-7.27 (dt, *J* = 7.2, 0.6 Hz, 2H), 5.19-

5.13 (AB, *J* = 9.6 Hz, 2H), 4.65-4.60 (d, *J* = 10.8, 1H), 4.60-4.55 (dd, *J* = 12.0, 1.8 Hz, 1H), 4.54-4.46 (m, 2H), 4.44-4.37 (m, 2H), 4.25-4.19 (t, *J* = 7.2 Hz, 1H), 4.10-4.05 (m, 1H), 3.83-3.78 (m, 1H), 3.77-3.70 (m, 2H), 3.59-3.55 (ddd, *J* = 9.6, 4.8, 1.8 Hz, 1H), 3.54-3.50 (m, 1H), 3.47-3.40 (m, 1H), 3.20-3.13 (m, 2H), 2.10 (s, 3H), 2.09 (s, 3H), 2.04 (s, 3H), 1.99 (s, 3H), 1.88 (s, 3H), 1.60-1.53 (m, 2H), 1.52-1.45 (m, 2H), 1.40-1.30 (m, 2H).

^{13}C NMR (125 MHz, $CDCl_3$) 170.3, 169.9, 169.8, 169.4, 169.3, 165.9, 156.4, 144.0 (2C), 141.3 (2C), 133.6, 129.6 (2C), 129.2, 128.7 (2C), 127.6 (2C), 127.0 (2C), 125.0, 124.98, 100.67, 100.58, 76.7, 73.9, 72.7, 72.31, 72.27, 71.8, 71.6, 69.7, 66.5, 63.0, 61.7, 59.9, 47.3, 40.84, 40.83, 29.4, 28.9, 23.0, 20.8, 20.60, 20.56, 20.53, 20.50.

$[\alpha]_D^{20}$ +7.6 (c 1, $CHCl_3$).

HRMS (ESI) *m/z*: Calcd for $C_{49}H_{56}N_4O_{18}$ ($M+Na^+$) 1011.3482, found 1011.3471.

Compound 4. NaOMe in MeOH (7.8 μL; 25% w/w) was added to **Compound 15** (26 mg, 0.026 mmol) in methanol (1.5 mL). After the resulting mixture was stirred at room temperature for 18 h, 1N HCl (0.2 mL) was added. The solvent was evaporated under reduced pressure to give a white solid substance, which was purifed by Combiflash to give (9.2 mg; **Compound 4**) in 79% yield.

IR (thin film) 3362, 2926, 2874, 2112, 1153, 1059, 1024 cm^{-1}. ^{1}H NMR (600 MHz, CD_3OD) δ (ppm) 4.40-4.38 (d, *J* = 7.8 Hz, 1H), 4.30-4.27 (d, *J* = 7.8 Hz, 1H), 3.93-3.89 (m, 1H), 3.88-3.86 (m, 2H), 3.82-3.79 (dd, *J* = 10.2, 1.8 Hz, 1H), 3.71-3.79 (dd, *J* = 10.2, 1.8 Hz, 1H), 3.60-3.55 (m, 9H), 3.43-3.36 (m, 2H), 3.33-3.20 (m, 6H), 2.96-2.91 (m, 2H), 1.73-1.65 (m, 4H), 1.55-1.48 (m, 2H). ^{13}C NMR (125 MHz, CD_3OD) 103.0, 102.7, 79.1, 75.7, 75.04, 75.00, 74.8, 73.6, 73.4, 68.8, 61.8, 60.8, 39.2, 31.4, 28.5, 26.7, 22.5.

$[\alpha]_D^{20}$ +17.0 (c 0.5, MeOH).

HRMS (ESI) *m/z*: Calcd for $C_{17}H_{32}N_4O_{10}$ ($M+H^+$) 453.2191, found 453.2181.

Glycan Substrate Conjugation to Microspheres

The carboxyl polystyrene microspheres (10^7-10^8 beads) were washed twice with 50 mM MES buffer, pH6.0 and were activated with EDC dissolved in MES buffer for 20 minutes at room temperature (RT). After activation, the microspheres were washed twice with 100mM borate

buffer (pH ~8.3) and were gently sonicated for one minute to disperse any aggregates. Microspheres were then incubated on an end-to-end shaker with 500-600 nmoles of the glycan substrates (**Compounds 1**, **2**, **3** or **4** in main text) in 1 mL of borate buffer for four hours at RT. Subsequently, microspheres were washed with borate buffer followed by incubation with 50mM glycine for 30 minutes at RT to block any unreacted carboxyl sites on the microsphere surface. The glycan derivatized microspheres were then washed and stored in water at 4°C for subsequent use.

Enzyme Assays and Detection:

GT assays

For the 6-ST assay, lactose- (**Compound 1**) derivatized microspheres were incubated with recombinant 6-ST (~1.4 μg/reaction) and 1mM CMP-sialic acid for 1 hour at 37°C with shaking in 100mM Tris-HCl buffer (pH 8.0). For the GalT assay, glucose- (**Compound 2**) microspheres were incubated with recombinant GalT (~ 0.2 μg/reaction) and 1mM UDP-galactose for 15 minutes at RT in assay buffer (10 mM Tris with 10 mM $MnCl_2$, pH 7.5). After completion of enzyme reactions, beads were thoroughly washed with PBS followed by incubation with fluorescent lectins, fluorescein labeled SNA (for 6-ST assay) and fluorescein labeled RCA-I (for GalT assay) at 5-10 μg/mL in PBS for 1hr at RT. Beads were then washed in PBS and analyzed using a BD FACScan flow cytometer. A total of 40,000 events were collected per sample and fluorescence was collected on the FL1 channel (which has appropriate filters for FITC/Dylight 488 dyes). Using FlowJo software, the events were then gated based on size on a forward scatter and side scatter plot. Overlays were then generated with FL1 fluorescence values for the size-gated events. For the time course experiments, lactose- microspheres were incubated with the 6-ST enzyme for the indicated times and were then labeled with the fluorescein-SNA lectin. Microspheres were analyzed on a standard flow cytometer and events were gated based on size (as described above). Using FlowJo software, median FL1 fluorescence values were calculated for the size-gated events (40,000 ± 2000 events per sample) for each time point. The median FL1 fluorescence value for the no enzyme control sample was then subtracted from the values for the enzyme incubated samples. The fluorescence values obtained were then normalized to the maximum value. The % maximum fluorescence vs. time plot was generated using Origin software.

GH assays

Microspheres derivatized with **Compound 3** were washed with 50mM phosphate buffer (pH 7.0) and then incubated with the recombinant cellulases (10 µg/reaction) for 1hr at 37°C with shaking. After completion of enzyme reactions, beads were thoroughly washed with PBS followed by incubation with the phosphine reagent, Dylight 488-phosphine (at 100 µM in PBS) for 2hr at 37°C. Beads were then washed with PBS-Tween20 and analyzed on a BD FACScan. A total of 40,000 events were collected per sample and fluorescence was collected on the FL1 channel (which has appropriate filters for FITC/Dylight 488 dyes). Using FlowJo software, the events were then gated based on size (forward scatter and side scatter). Overlays with FL1 fluorescence were then generated for the size-gated events. In order to quantify the FL1 fluorescence for each sample, the median of fluorescence values for size-gated events for each sample were calculated using FlowJo and normalized to the median values of the no enzyme control (which has maximum median fluorescence). For the time course experiments, the microspheres were incubated with CelAcc_CBM3a enzyme (10µg/reaction) for the indicated time points. Microspheres were washed, labeled and analyzed using a BD FACScan and FlowJo software as described above.

For the multiplexed GH assays, 1µm microspheres derivatized with **Compound 4** and 6µm microspheres derivatized with **Compound 3** were mixed (in a 1:6 volume ratio) and washed with 50mM sodium acetate buffer, pH 4.6. This volume ratio was chosen to keep the surface area of the two different bead sizes (and thereby the number of azide-cellobiose (compound 4) substrate molecules and azide-cellotetraose (compound 3) substrate molecules) constant for the multiplexed assay. Subsequently the microsphere mixture was incubated with a commercial CTEC2 cocktail (~ 0.2 µl cocktail/reaction) in acetate buffer for 2hr at 50°C with shaking. Microspheres were washed, labeled with phosphine reagent and analyzed on a BD FACScan using FlowJo software (as mentioned above).

REFERENCES

(1) P. G. Blommel, P. A. Martin, K. D. Seder, R. L. Wrobel, B. G. Fox, *Methods Mol. Biol.* **2009**, *498*, 55-73.

(2) T. E. Takasuka, J. A. Walker, L. F. Bergeman, K. A. Vander Meulen, S.-I. Makino, N. L. Elsen, B. G. Fox, *Methods Mol. Biol.* **2012**, *In Press*.

(3) (a) K. Deng, K. W. George, W. Reindl, J. D. Keasling, P. D. Adams, T. S. Lee, A. K. Singh, T. R. Northen, *Rapid Commun. Mass Spectrom.* **2012**, *26*, 611-615 (For synthesis of lactose **Compound 1**). (b) C. Wendeln, A. Heile, H. F. Arlinghaus, B. J. Ravoo, *Langmuir* **2010**, *26*, 4933-4940 (For synthesis of glucose **Compound 2**).

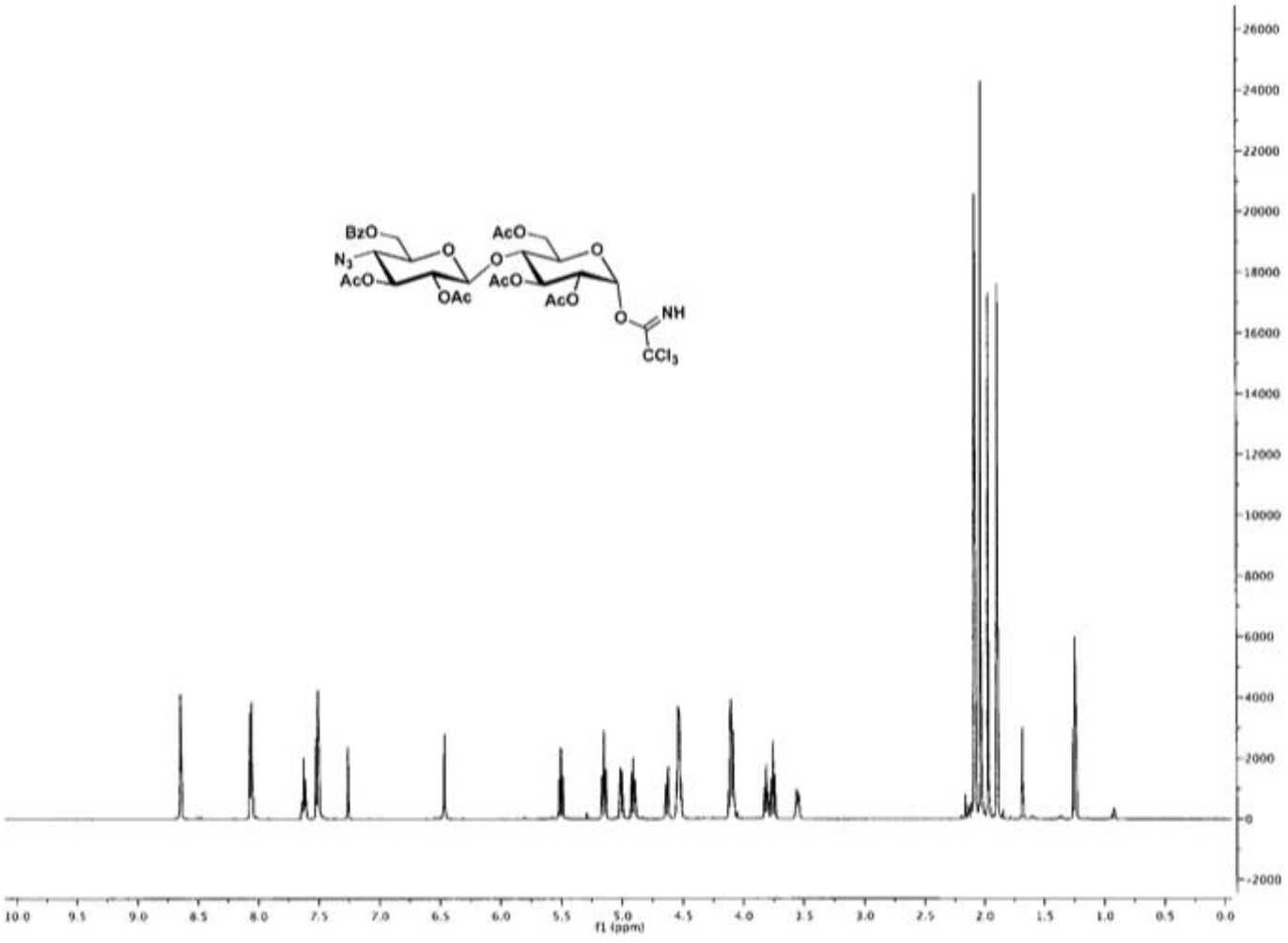
(4) V. Boyer, S. Fort, T. P. Frandsen, M. Schulein, S. Cottaz, H. Driguez, *Chemistry* **2002**, *8*, 1389-1394.

(5) D. Day, M. Du, R. Mendez, D. Maunder, **2008**. **WO 2008000082 A1**.

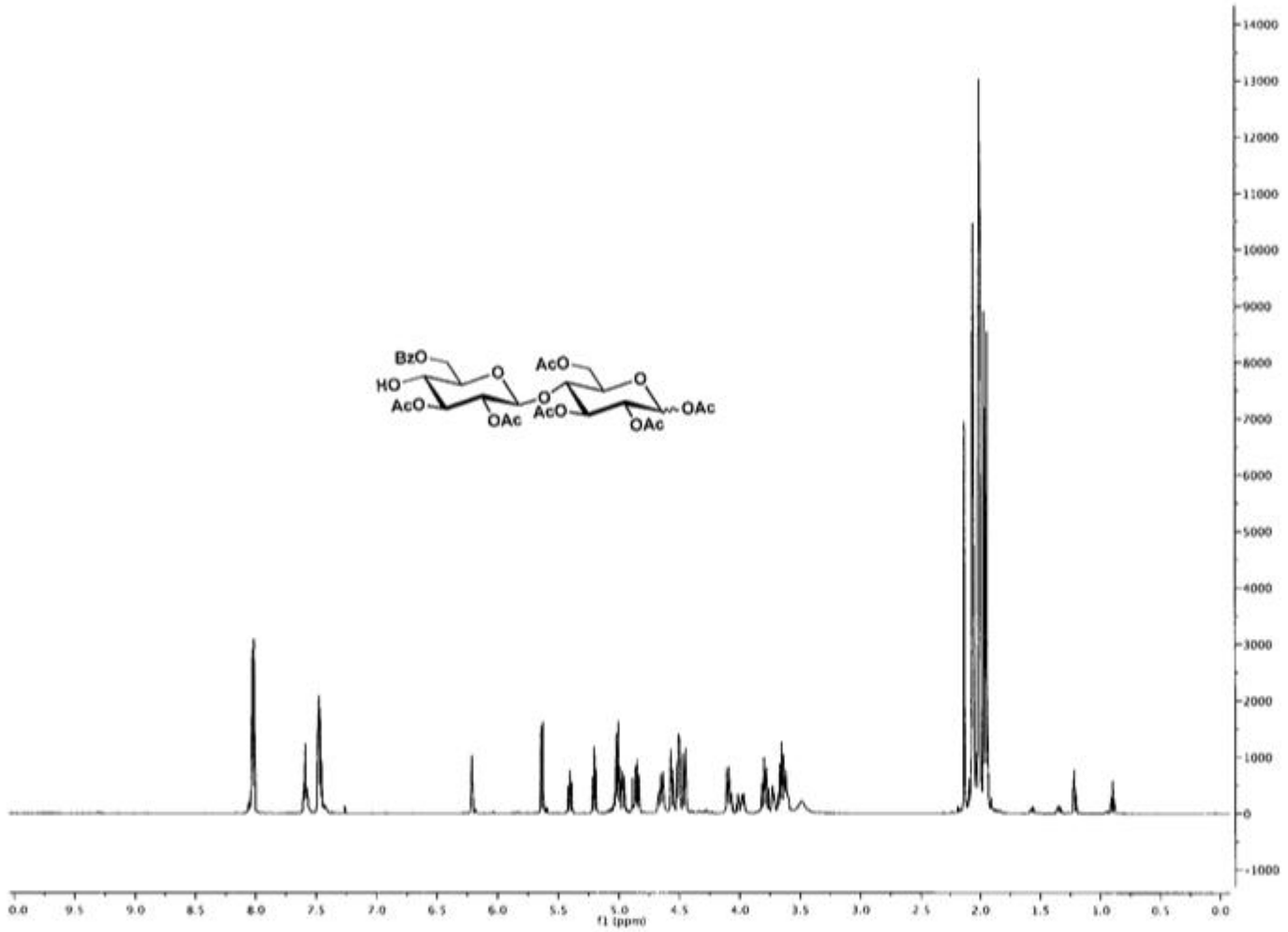
(6) K. H. Verschueren, K. Pumpor, S. Anemuller, S. Chen, J. R. Mesters, R. Hilgenfeld, *Chem. Biol.* **2008**, *15*, 597-606.

Proton NMR Spectra of Glycan Compounds

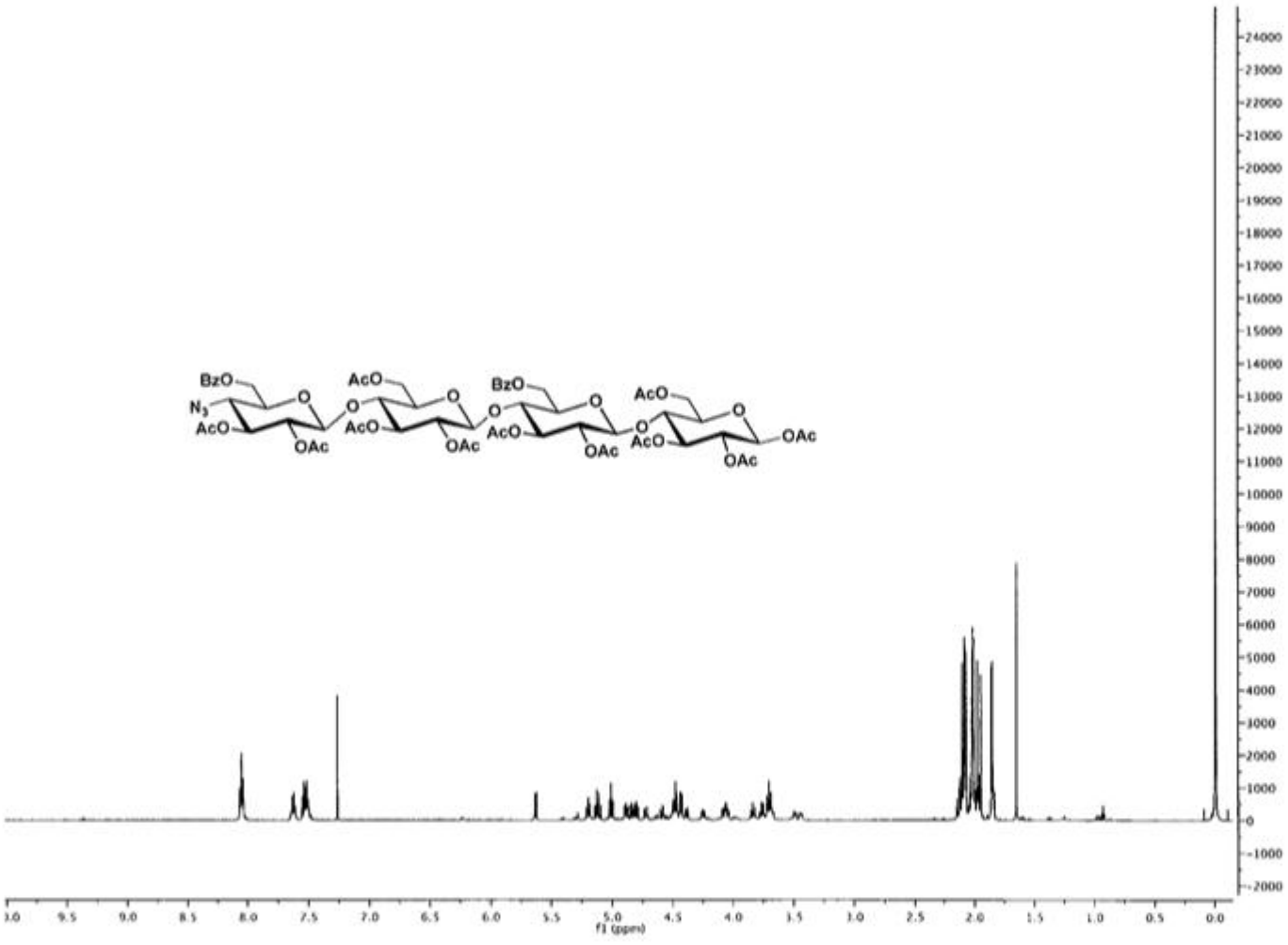
Compound 7 (^{1}H NMR, 600 MHz, $CDCl_3$)



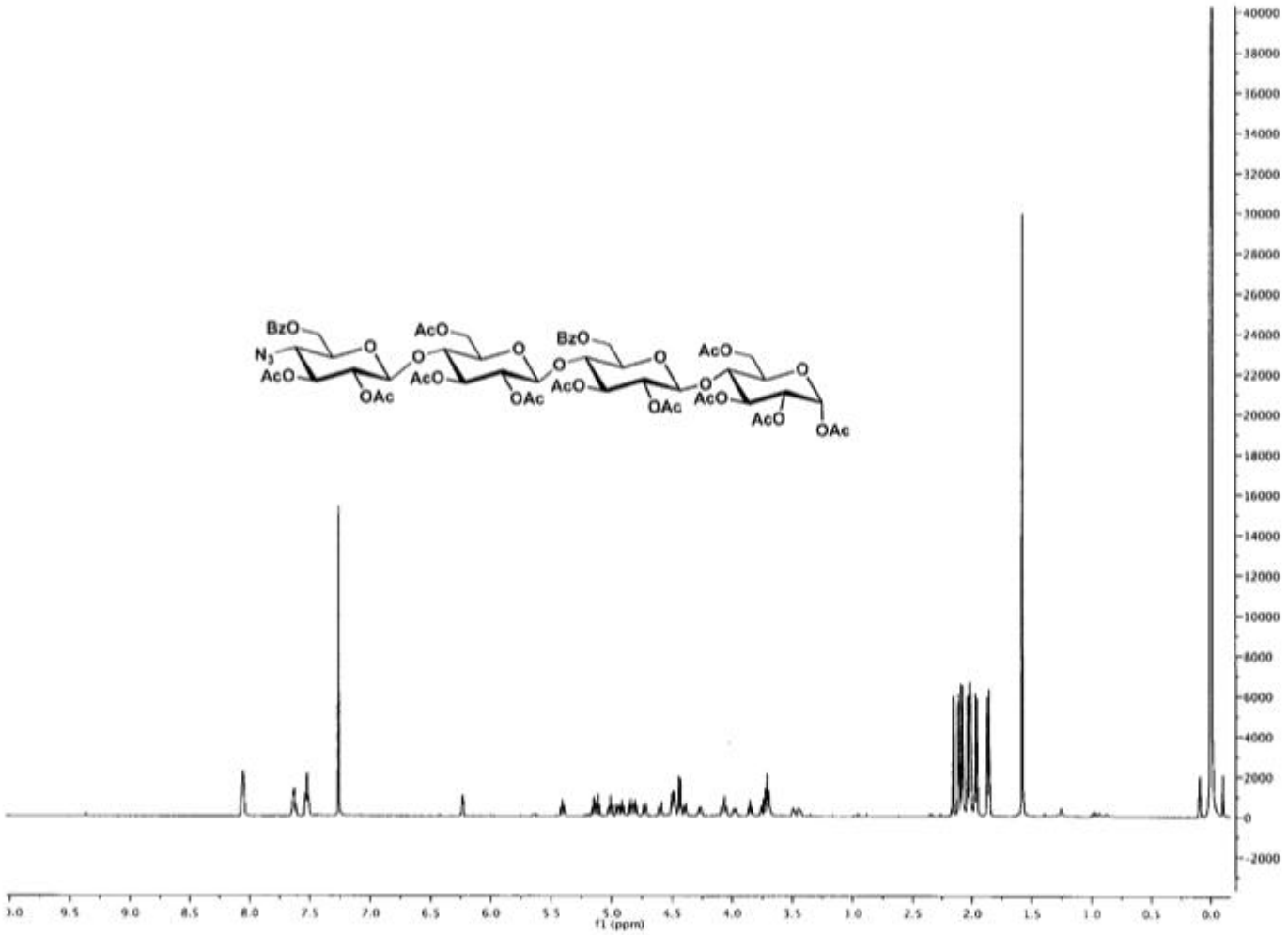
Compound 9 (^{1}H NMR, 600 MHz, $CDCl_3$)



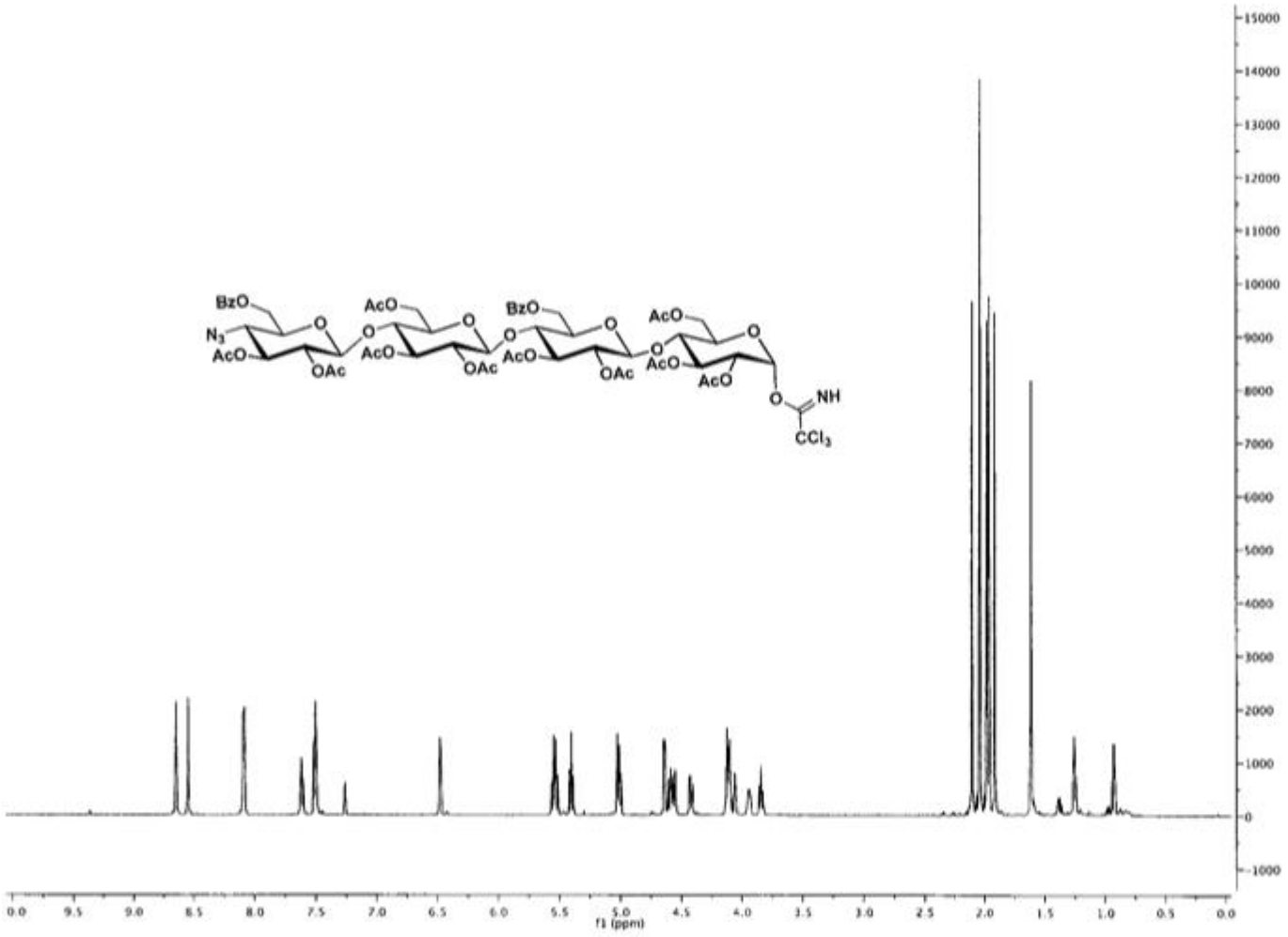
Compound 10 α-anomer (^{1}H NMR, 600 MHz, $CDCl_3$)



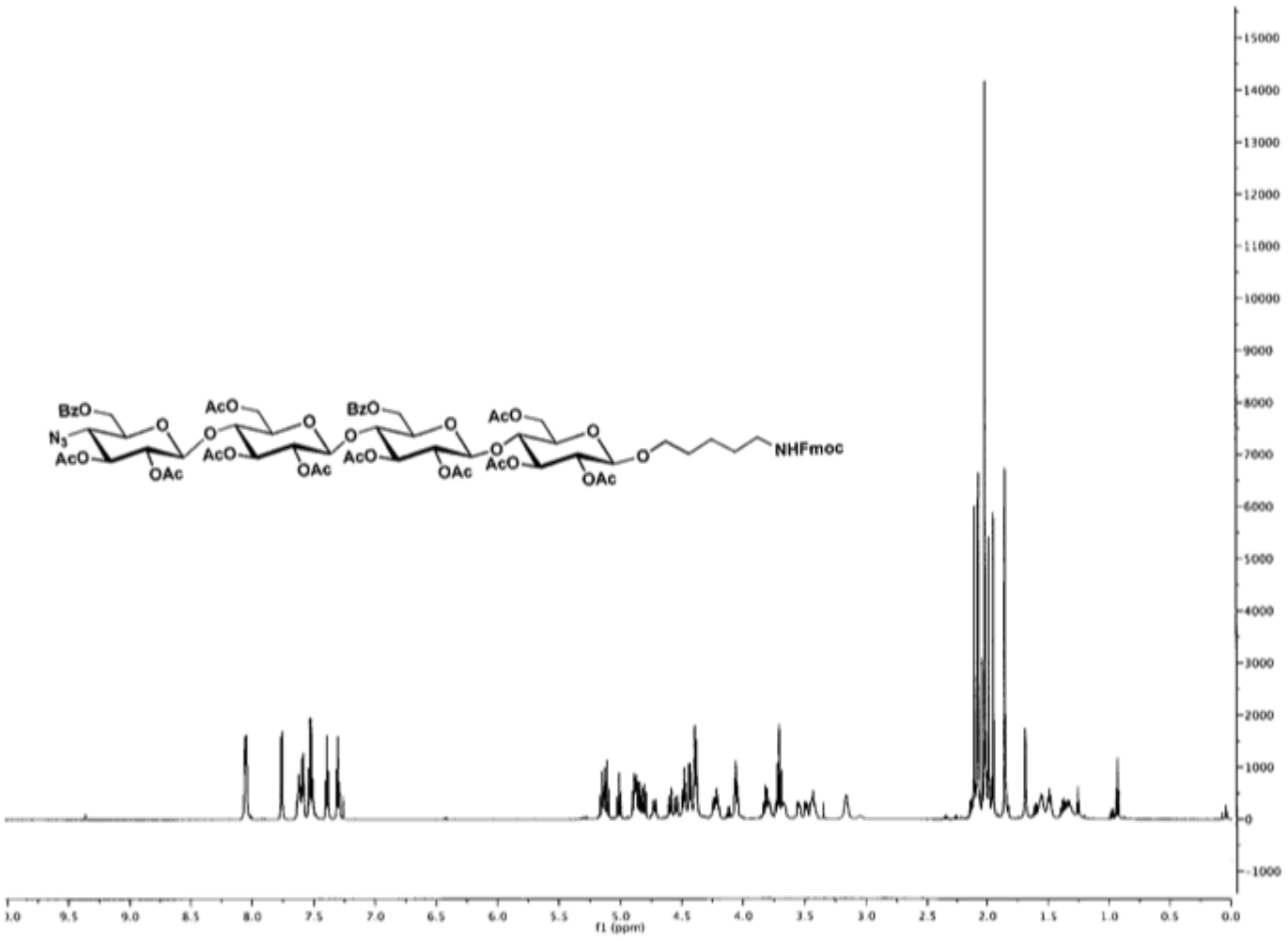
Compound 10 β-anomer (^{1}H NMR, 600 MHz, $CDCl_3$)



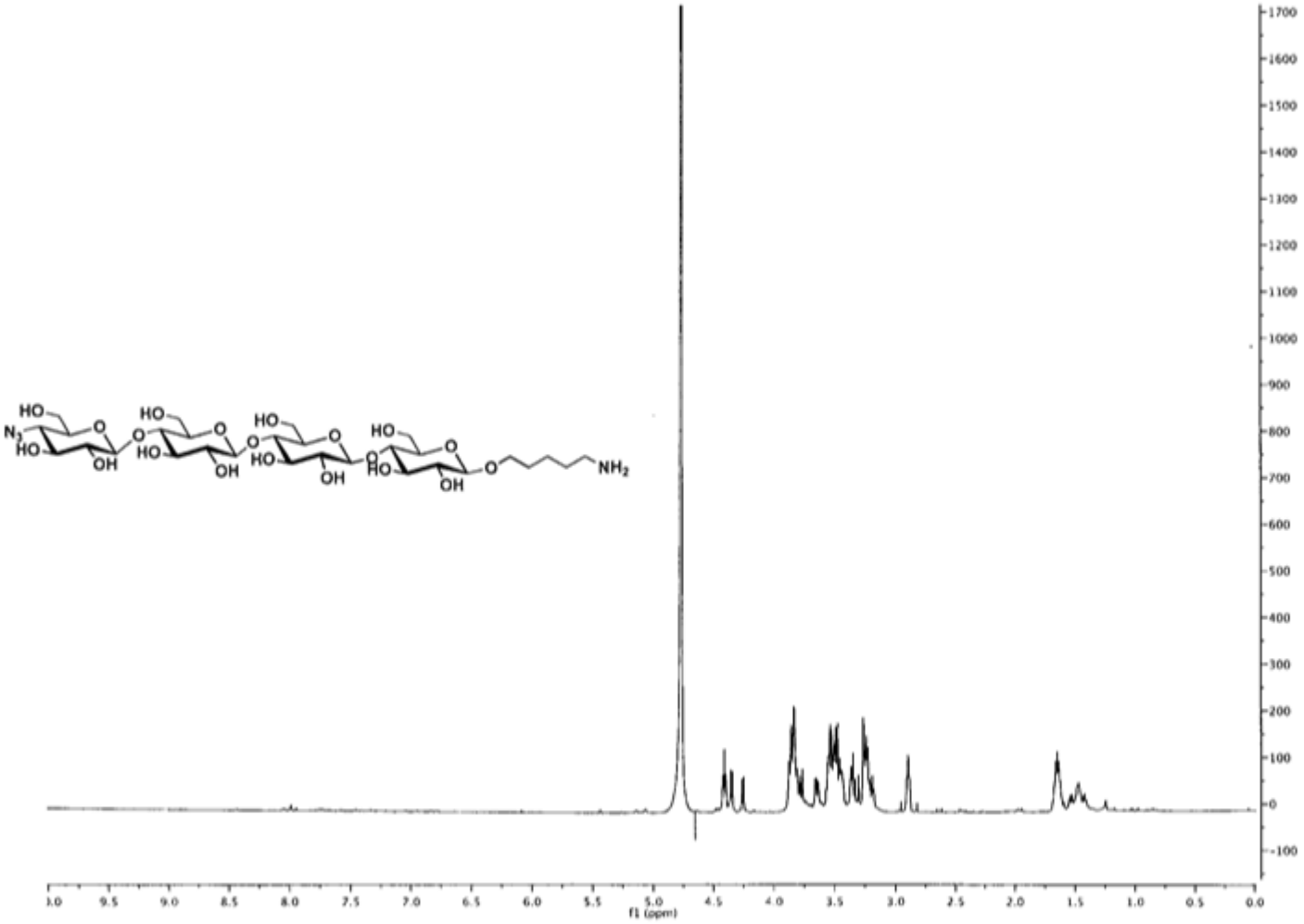
Compound 12 (^{1}H NMR, 600 MHz, $CDCl_3$)



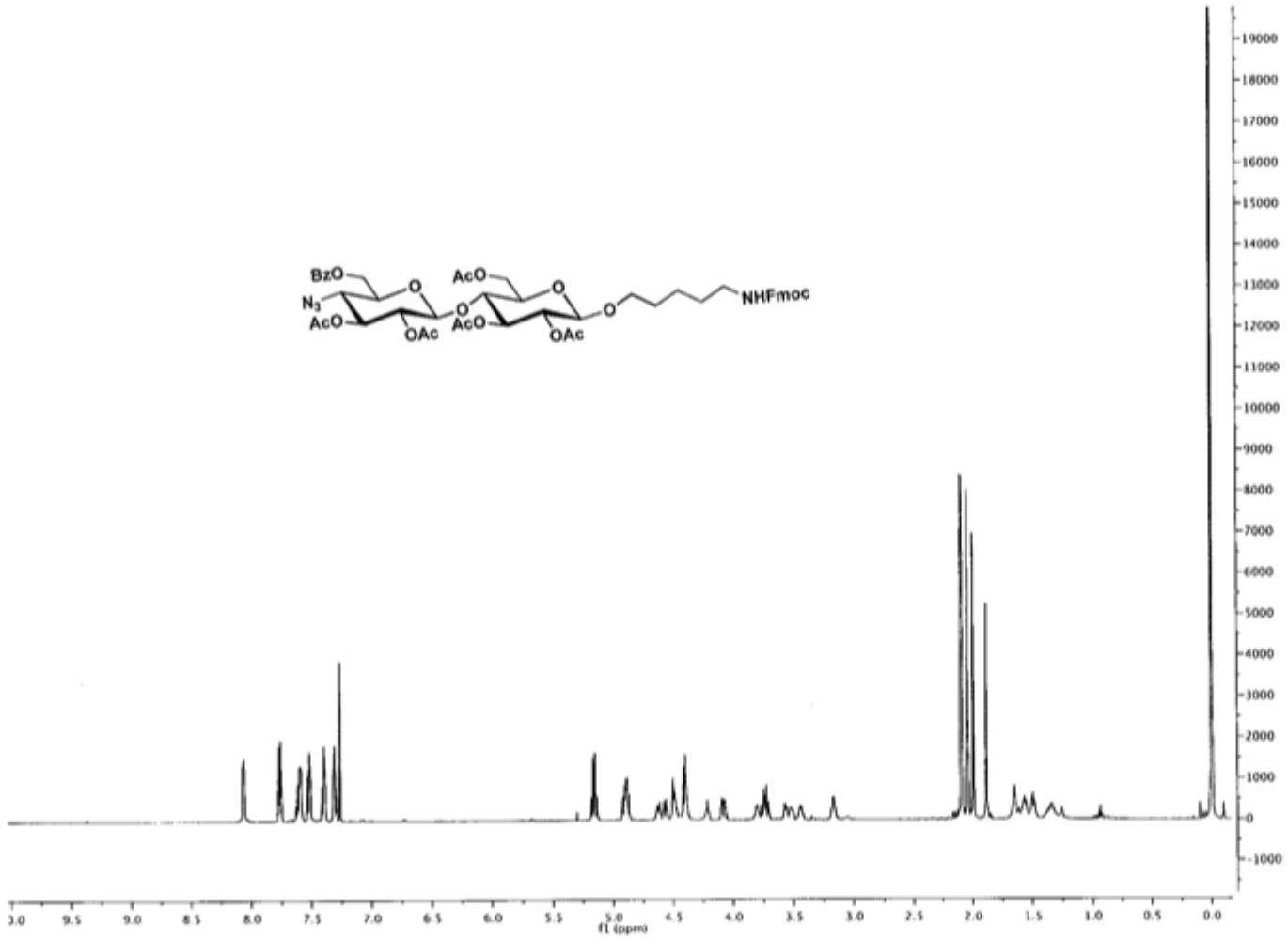
Compound 14 (^{1}H NMR, 600 MHz, $CDCl_3$)



Compound 3 (^{1}H NMR, 600 MHz, CD_3OD)



Compound 15 (^{1}H NMR, 600 MHz, $CDCl_3$)



Compound 4 (^{1}H NMR, 600 MHz, CD_3OD)