Supplementary Material (ESI) for Chemical Communications

Supporting Information

for

Highly Efficient Chemoenzymatic Synthesis of β1–4-Linked Galactosides with Promiscuous Bacterial β1–4-Galactosyltransferases

Kam Lau, Vireak Thon, Hai Yu, Li Ding, Yi Chen, Musleh M. Muthana, Denton Wong, Ronald Huang, and Xi Chen*

Department of Chemistry, University of California, One Shields Avenue, Davis, CA 95616

*To whom correspondence should be addressed: Tel: +1-530-754-6037; Fax: +1-530-752-8995; e-mail: chen@chem.ucdavis.edu

Table of Contents

Cloning and expression of *Neisseria meningitidis* β1–4-galactosyltransferase encoded by *lgtB* gene (NmLgtB), *Helicobacter pylori* β1–4-galactosyltransferase (Hp1–4GalT) and *Pasteurella multocida* inorganic pyrophosphatase (PmPpA)

Bacterial strains, plasmids, and materials

Electrocompetent *E. coli* DH5α cells and chemically competent *Escherichia coli* BL21 (DE3) cells were purchased from Invitrogen (Carlsbad, CA). Genomic DNA of *Neisseria meningitidis* serogroup B strain MC58 (ATCC BAA-335D), *Helicobacter pylori* strain J99 genomic DNA (ATCC 700824D), and *Pasteurella multocida* strain P-1059 (ATCC 15742) were from American Type Culture Collection (ATCC, Manassas, VA). Vector plasmids pET15b and pET22b(+) were purchased from Novagen (EMD Biosciences, Inc. Madison, WI). Ni^{2+}–NTA agarose (nickel–nitrilotriacetic acid–agarose), QIAprep spin miniprep kit, and QIAEX II gel extraction kit were from Qiagen (Valencia, CA). Herculase enhanced DNA polymerase was from Stratagene (La Jolla, CA). T4 DNA ligase, 1 kb DNA ladder, and *BamH*I restriction enzyme were obtained from Promega (Madison, WI). *Nde*I restriction enzyme and vector plasmid pMAL-c4X were from New England Biolabs, Inc. (Beverly, MA). Bicinchoninic acid (BCA) protein assay kit was from Pierce Biotechnology, Inc. (Rockford, IL). Uridine 5'-diphosphogalactose (UDP-Gal), 4-methylumbelliferyl-2-acetamido-2-deoxy-β-D-glucopyranoside (GlcNAcβMU), 4-methylumbelliferyl-β-D-glucopyranoside (GlcβMU), *N*-acetyl-D-glucosamine (GlcNAc), and D-glucose (Glc) were all purchased from Sigma (St. Louis, MO).

Cloning

Full length NmLgtB (GenBank accession number AE002098) was cloned from genomic DNA of *Neisseria meningitidis* (ATCC BAA-335D) in pET15b vector and expressed in *Escherichia coli* as an N-terminal His_6-tagged fusion protein. Primers used were: forward primer 5'GATCCATATGCAAAACCACGTTATCAGC3' (*Nde*I restriction site is underlined) and reverse primer 5'CGCGGATCCTTATTGGAAAGGCACAATGAAC3' (*BamH*I restriction site is underlined). **Full length Hp1–4GalT** (GenBank accession number AB035971) was cloned from *Helicobacter pylori* str. J99 genomic DNA (ATCC 700824D) in pET15b (NdeI and BamHI sites) or pET22b(+) (NdeI and XhoI sites) as N-His_6- or C-His_6-tagged recombinant fusion protein (the expression level was low, <0.4 mg/L culture, and these enzymes were not used for the synthesis). It was also cloned in pMAL-c4X vector and expressed in *E. coli* as an N-terminal maltose binding protein (MBP)-tagged and C-His_6-tagged fusion protein. Primers were: forward primer 5'-CGCGGATCCATGCGTGTTTTTATCATTTCTTTAAATC-3' (*BamH*I restriction site is underlined); reverse primer 5'-CGGCCTGCAGTTA**GTGGTGGTGGTGGTGGTG**TACAAA CTGCCACCATTTC-3' (*PstI* restriction site is underlined; codons for C-His_6-tag are in bold). **Full length PmPpA** encoded by *PM1191* gene was cloned from *Pasteurella multocida* strain P-1059 genomic DNA (AE004439) into pET15b or pET22b(+) vector and N-His_6-tagged or C-His_6-tagged fusion proteins respectively. The primers used were in vector were: forward primer for both vectors 5'-GATCCATATGGGTTTAGAAACCGTACCGGC-3' (NdeI restriction site is underlined); reverse primer for pET15b vector 5'-CGCGGATCCTTATGCCCAAAACAGGACGTTTT-3' (*BamH*I restriction site is underlined) and for pET22b(+) vector 5'-CCGCTCGAGTTTTTTTGCACGTTCAAATG-3' (*Xho*I restriction site is underlined).

PCR amplification of the target gene was performed in a 50 μL reaction containing genomic DNA (1 μg), forward and reverse primers (1 μM each), 10 × Herculase buffer (5 μL), dNTP mixture (1 mM), and 5 units (1 μL) of Herculase enhanced DNA polymerase. The reaction mixture was subjected to 30 cycles of amplification with an annealing temperature of 50–54°C. The resulting PCR product was digested with restriction enzymes, purified, and ligated with predigested pET15b, pET22b(+), or pMAL-c4X vector. Ligation product was transformed into electrocompetent *E. coli* DH5α cells. Selected clones were grown for minipreps and positive clones were verified by restriction mapping and DNA sequencing performed by Davis Sequencing Facility in the University of California-Davis.

Expression

Positive plasmid was selected and subsequently transformed into *E. coli* BL21 (DE3) (for NmLgtB and PmPpA) or Origami B (DE3) (for Hp1–4GalT) chemically competent cells. Plasmid-bearing *E. coli* strain was cultured in LB medium (10 g/L tryptone, 5 g/L yeast extract, and 10 g/L NaCl) with ampicillin (100 μg/mL) (for PmPpA); LB medium with ampicillin (100 μg/mL) and kanamycin (15 μg/mL) (for Hp1–4GalT); or LB medium containing 0.2% glucose, 0.5 M NaCl (total medium salt concentration), and 1 mM betaine (LBNB)[1] supplemented with ampicillin (100 μg/mL) (for NmLgtB). For expression of NmLgtB in LBNB medium, *E. coli* cells were first subjected to overnight adaptation of high salt concentration (0.5 M NaCl) prior to induction by isopropyl-1-thio-β-D-galactopyranoside (IPTG). Overexpression of the target protein was achieved by inducing the *E. coli* culture with 0.1 mM (for NmLgtB and PmPpA) or 0.3 mM (for Hp1–4GalT) of IPTG when the $OD_{600\ nm}$ of the culture reached 0.8–1.0 and incubating at 20°C for 16–20 h (for NmLgtB and PmPpA) or 25°C for 18 h (for Hp1–4GalT) with vigorous shaking at 250 rpm in a C25KC incubator shaker (New Brunswick Scientific, Edison, NJ).

Protein purification

His_6-tagged target proteins were purified from cell lysate using Ni^{2+}–NTA affinity column. To obtain cell lysate, cells were harvested by centrifugation at 4,000 rpm (Sorvall) at 4°C for 3 h. The cell pellet was resuspended in lysis buffer (pH 8.0, 100 mM Tris–HCl containing 0.1% Triton X-100). Lysozyme (100 μg/mL) and DNaseI (3 μg/mL) were then added to the cell suspension. The mixture was incubated at 37ºC for 50 min with vigorous shaking (200 rpm). Cell lysate was obtained by centrifugation at 12,000 rpm (Sorvall) at 4°C for 30 min as the supernatant. Purification of His-tagged proteins from cell lysate supernatant was achieved using an AKTA FPLC system (GE Healthcare) equipped with a HisTrap™ FF 5 mL Ni^{2+}–NTA affinity column. The column was pre-equilibrated with 10 column volumes of binding buffer (5 mM imidazole, 0.5 M NaCl, 50 mM Tris–HCl, pH 7.5) before the lysate was loaded. Followed by washing with 10 column volumes of binding buffer, 10 column volumes of washing buffer (50 mM imidazole, 0.5 M NaCl, 50 mM Tris–HCl, pH 7.5), target His-tagged protein was eluted with 10 column volumes of elution buffer (200 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl, pH 7.5). Fractions containing purified enzyme were collected and dialyzed against a dialysis buffer (20 mM Tris-HCl, pH 7.5, 10% glycerol). The purified recombinant enzyme was stored at -20°C. Typically, 12.5 mg of NmLgtB, 5 mg of Hp1–4GalT, and 432–488 mg of PmPpA can be purified from one liter cell culture. Sodium dodecylsulfate–polyacrylamide gel electrophoresis (SDS–PAGE) indicated the successful expression and purification of targeted proteins (**Figure S1**).

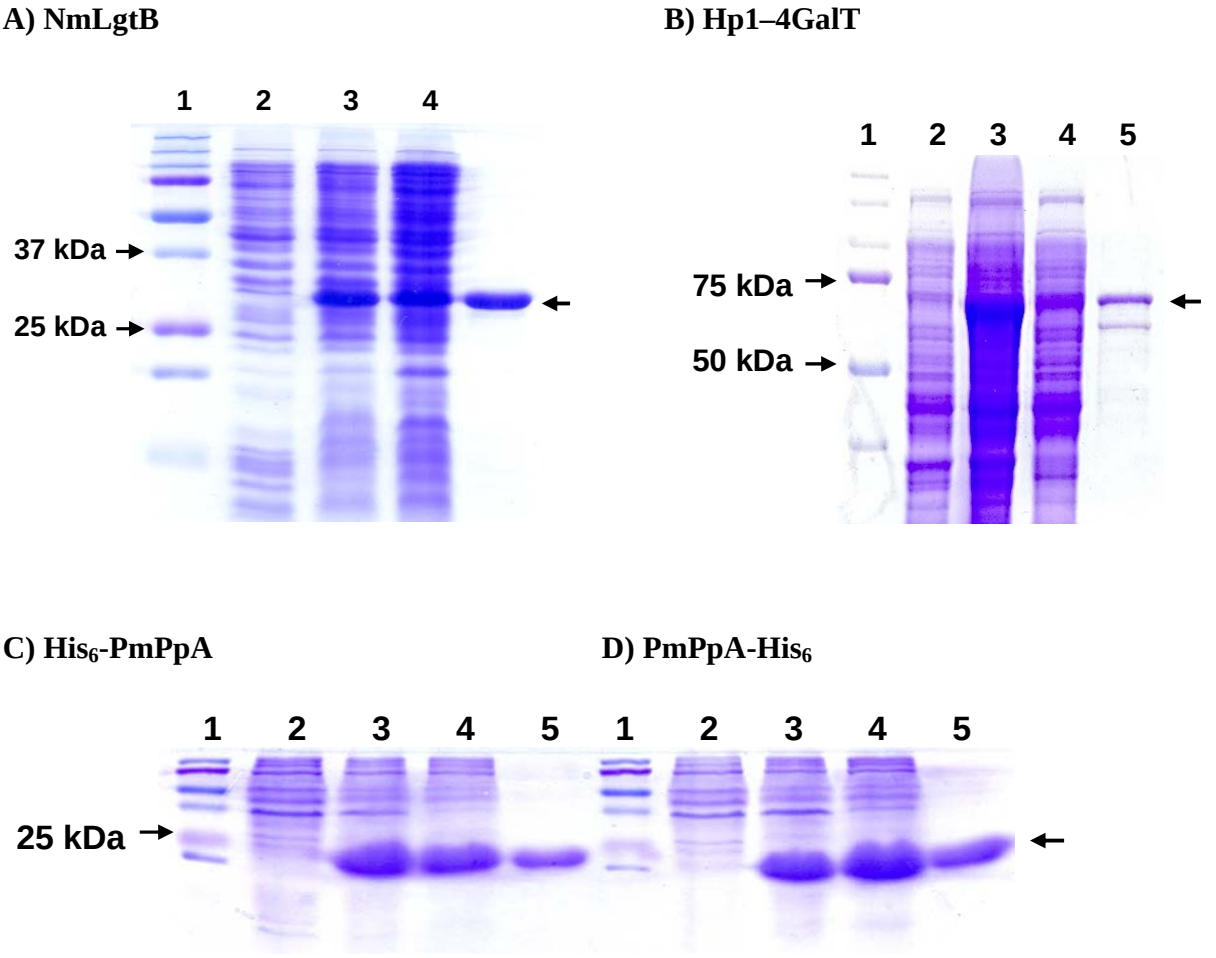


Figure S1. SDS–PAGE analysis of NmLgtB (A), Hp1–4GalT (B), His_6-PmPpA (C), and PmPpA-His_6 (D). Lanes: 1, Bio-Rad Precision Plus Protein Standards (10–250 kDa); 2, whole cell extraction before induction; 3, whole cell extraction after induction; 4, lysate (supernatant) after induction; 5, Ni^{2+}-column purified protein.

Quantification of purified protein

The concentration of purified enzyme was obtained in a 96-well plate using a Bicinchoninic acid (BCA) Protein Assay Kit (Pierce Biotechnology, Rockford, IL) with bovine serum albumin as a protein standard. The absorbance of samples was measured at 562 nm by a BioTek Synergy™ HT Multi-Mode Microplate Reader.

Detection of NmLgtB activity using TLC (Glc and GlcNAc) and HPLC (Glcβ-MU and GlcNAcβ-MU) analyses

The enzymatic activity of purified NmLgtB was detected by thin-layer chromatography (TLC) analysis using GlcNAc and Glc as acceptor substrates and also by high performance liquid chromatography (HPLC) analysis using fluorescent acceptors GlcNAcβMU and GlcβMU. Reactions were carried out in a total volume of 20 μL at 37°C for 60 min. For TLC analysis, the reaction mixtures contained Tris-HCl buffer (pH 7.0, 150 mM), $MnCl_2$ (10 mM), Glc or GlcNAc

(10 mM), and UDP-Gal (10 mM) and the reactions were analyzed using isopropanol:NH_4OH:water (7:3:2) as mobile phase for TLC and the plates were stained with *p*-anisaldehyde sugar stain. For HPLC analysis, the reaction mixtures contained MES buffer (pH 7.0, 150 mM), GlcβMU (1 mM) and/or GlcNAcβMU (1 mM), $MnCl_2$ (10 mM), and UDP-Gal (2 mM). As shown in **Figure S2**, when the reaction was carried out in the presence of both GlcβMU and GlcNAcβMU, both LacβMU and LacNAcβMU formed with similar percentage conversions, indicating both GlcβMU and GlcNAcβMU are similarly good acceptors for NmLgtB.

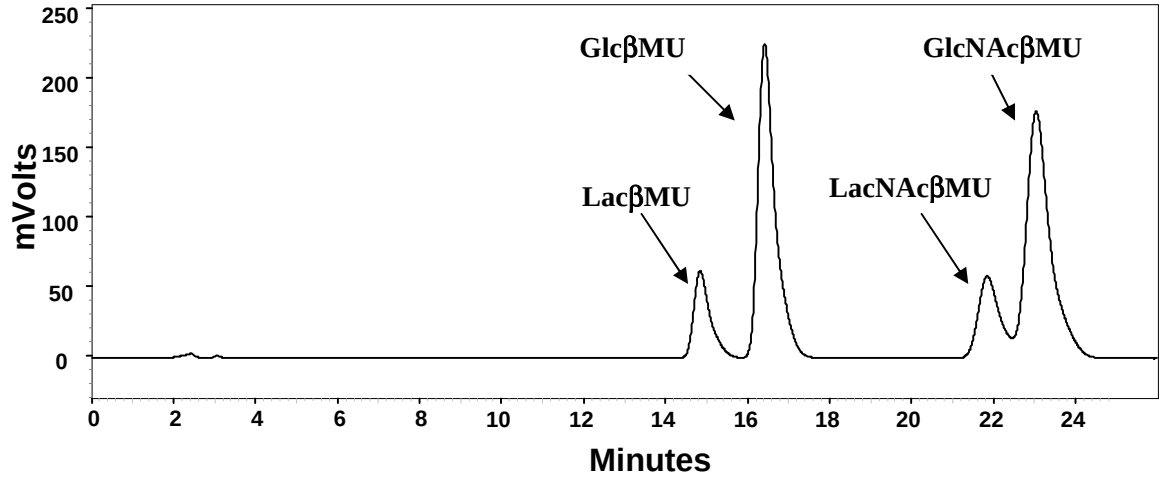
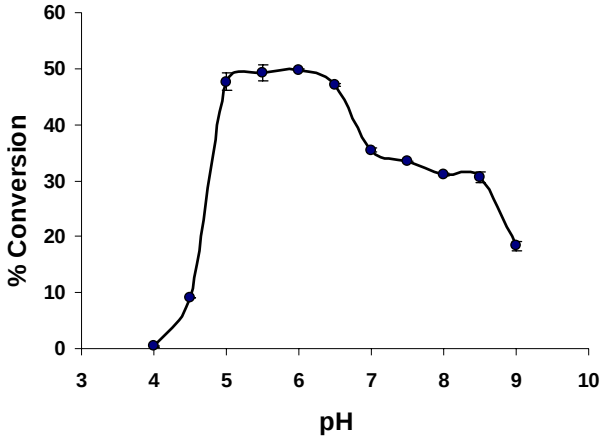


Figure S2. HPLC analysis of NmLgtB reactions using both GlcβMU (1 mM) and GlcNAcβMU (1 mM) as acceptor substrates. Percentage conversions to LacβMU and LacNAcβMU are 26% and 29%, respectively.

NmLgtB pH profile by HPLC assays

Typical enzymatic assays were performed in a total volume of 20 μL in a buffer (250 mM) with pH varying from 4.0–9.0 containing 10 mM $MnCl_2$, 1 mM GlcNAcβMU, 1 mM UDP-Gal, and 8.9 μg enzyme in an Eppendorf's tube. Reactions were allowed to proceed for 15 min at 37°C and quenched by the addition of ice-cold 12% acetonitrile (580 μL) to make 30-fold dilution.

Figure S3. The pH profile of NmLgtB by HPLC analysis using GlcNAcβMU as the acceptor substrate. Buffers (250 mM) used were: acetate-NaOH, pH 4.0–5.0; MES, pH 5.5–6.5; Tris-HCl, pH 7.0–9.0.



The samples were then kept on ice until analyzed by a Shimadzu LC-2010A HPLC system equipped with a membrane on-line degasser, a temperature control unit and a fluorescence detector. A reverse phase Premier C18 column (250 × 4.6 mm I.D., 5 μm particle size, Shimadzu) protected with a C18 guard column cartridge was used. The mobile phase was 12% acetonitrile. The fluorescent compounds GlcNAcβMU and LacNAcβMU were detected by excitation at 325 nm and emission at 372 nm.[2] All assays were carried out in duplicate.

As shown in **Figure S3**, the pH profile of NmLgtB using GlcNAcβMU as an acceptor indicates an optimal pH range of 5.0 to 6.5. The highest activity was observed in MES buffer at pH 5.5. Medium activity was observed in the pH range of 7.0–8.5. Low activity was seen at pH 9.0 or 4.5.

Effects of Co^{2+}, Mg^{2+}, and Mn^{2+} on NmLgtB activity

Different concentrations of $MgCl_2$, $MnCl_2$, and $CoCl_2$ were used at pH 6.0 to analyze the effect of metal ions on the LgtB activity. The reactions were carried out in the presence of 1 mM GlcβMU 1 mM UPD-Gal, and 8.9 μg enzyme for 15 min at 37°C. The reactions were quenched by the addition of ice-cold 12% acetonitrile (580 μL) to make 30-fold dilution, and subsequently analyzed by HPLC. Reactions without the metal ions were used as a control. As shown in **Figure S4**, divalent metals such as Mn^{2+}, Mg^{2+}, or Co^{2+} can effectively activate the activity of NmLgtB although the effect of Mg^{2+} is lower compared to Mn^{2+} or Co^{2+}. At concentrations of less than 3 mM, Co^{2+} showed slightly higher activation than Mn^{2+}. This result is similar to that for the truncated bovine β1–4GalT1, showing faster activation at lower concentration of Co^{2+}.[3] At metal ion concentration higher than 3 mM, the activation of NmLgtB by Mn^{2+} is slightly higher than that by Co^{2+}. NmLgtB activity reaches a plateau with Mn^{2+}>Co^{2+}>Mg^{2+} when the concentration of these metal ions are in the range of 5–50 mM.

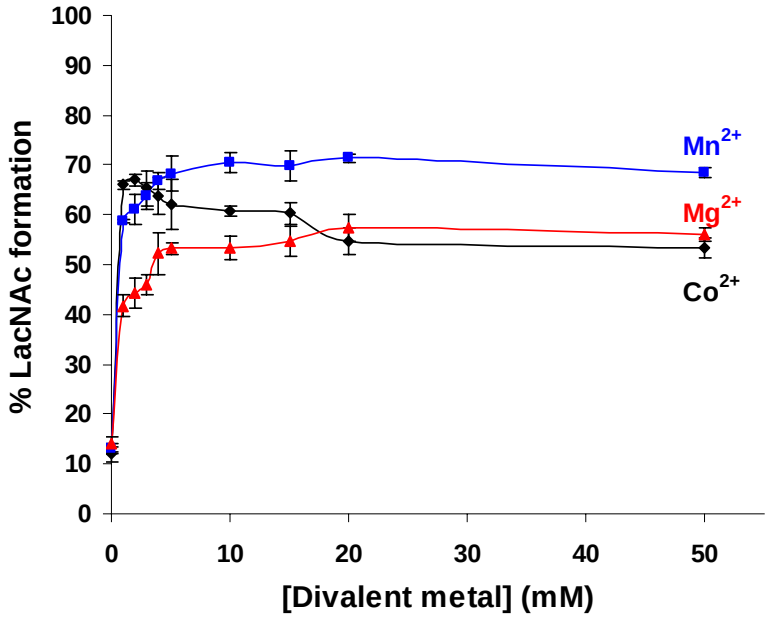


Figure S4. Effects of metal ions Mn^{2+}, Mg^{2+}, and Co^{2+} and their concentrations on the activity of NmLgtB by HPLC analysis.

NmLgtB kinetics by HPLC assays

The enzymatic assays were carried out *in vitro* in a total volume of 20 μL containing 150 mM MES-KOH (pH 6.0), 10 mM $MnCl_2$, 4-methylumbelliferyl-2-acetamido-2-deoxy-β-D-glucopyranoside (GlcNAcβMU) or 4-methylumbelliferyl-β-D-glucopyranoside (GlcβMU), uridine 5'-diphospho-α-D-galactopyranoside (UDP-Gal), and 8.3 μM enzyme in an Eppendorf tube. Reactions were allowed to proceed for 10 min at 37°C and were quenched by the addition of ice-cold 12% acetonitrile (20:1 dilution). Michaelis-Menten kinetic parameters were obtained by varying the concentrations of the acceptor substrate GlcNAcβMU (0.5, 1.0, 2.0, 4.0, and 5.0 mM) or GlcβMU (0.5, 1.25, 2.5, 4.0, and 8.0 mM) in the presence of a fixed concentration (1 mM) of UDP-Gal, or varying the concentrations of the sugar nucleotide donor UDP-Gal (0.05, 0.1, 0.25, 0.5, 1.0, and 2.0 mM) with a fixed concentration (0.35 mM) of GlcNAcβMU or GlcβMU . The double reciprocal Lineweaver-Burk plots were obtained from the average values of duplicate assay results.

The apparent K_m values of the acceptors GlcNAcβMU and GlcβMU are 1.3 ± 0.3 mM and 1.6 ± 0.2 mM, respectively for NmLgtB (**Table 1**). The apparent K_m value of UDP-Gal is 26 ± 4 μM. This K_m value of UPD-Gal is comparable to those of the non-histidine tagged *LgtB* from *Neisseria meningitidis* MC58 (25 μM),[4] human placenta β1–4GalT (25 μM),[5] and truncated bovine milk GT-d129 (60 μM).[6] The acceptor GlcNAcβMU K_m value (1.3 ± 0.3 mM) of NmLgtB is comparable to those of GlcNAcβ1–4GlcNAc (1.59 mM) for human β1–4GalT[7] and GlcNAcβ-FCHASE (0.6 mM) for the previsously reported recombinant *Neisseria meningitidis* LgtB.[4] In the case of GlcβOR acceptor, β1–4-galactosyltransferase from *H. pylori* showed activity, but to a lesser degree than LgtB.[8,9] Mammalian β1–4-galactosyltransferases are able to utilize glucose or GlcβOR effectively but only in the presence of α-lactalbumin. For example, the K_m value for a mammalian β1–4-galactosyltransferase was reported to be 5.9 mM for glucose as an acceptor.[10]

Table 1. Apparent kinetic parameters of NmLgtB

Parameters/substrates	UDP-Gal	GlcNAcβMU	GlcβMU
V_{max} (mM/min)	$(1.5 \pm 0.1) \times 10^{-2}$	$(8.4 \pm 0.7) \times 10^{-2}$	$(8.7 \pm 0.3) \times 10^{-2}$
K_m (mM)	$(2.6 \pm 0.4) \times 10^{-2}$	1.3 ± 0.3	1.6 ± 0.2
k_{cat} (min^{-1})	1.8 ± 0.1	10 ± 1	11 ± 1
k_{cat}/K_m ($min^{-1}mM^{-1}$)	74 ± 3	7.8 ± 0.1	7.1 ± 0.6

General methods for compound purification and characterization:

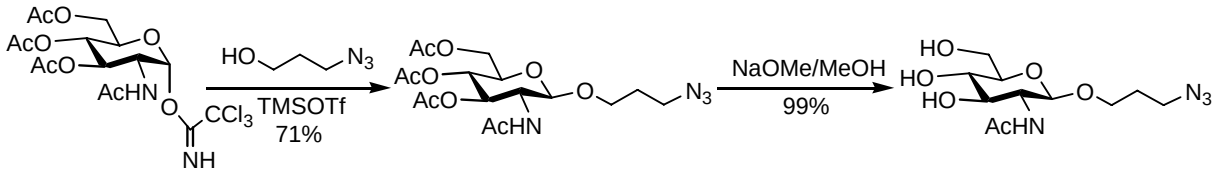
Chemicals were purchased and used without further purification. ^{1}H NMR and ^{13}C NMR spectra were recorded on Mercury (300 MHz), Inova (400 MHz), and Varian VNMRS 600 MHz spectrometers. High resolution electrospray ionization (ESI) mass spectra were obtained at the Mass Spectrometry Facility in the University of California, Davis. Silica gel 60 Å (Sorbent Technologies) was used for flash column chromatography. Thin-layer chromatography (Sorbent Technologies) was performed on silica gel plates using anisaldehyde sugar stain or 5% sulfuric acid in ethanol stain for detection. Gel filtration

chromatography was performed with a column (100 cm × 2.5 cm) packed with BioGel P-2 Fine resins.

Synthesis of GlcNAc and Glc derivatives

GlcNAc **1**, GlcNAcβMU **4**, and Glucose **7** were obtained from commercial suppliers.

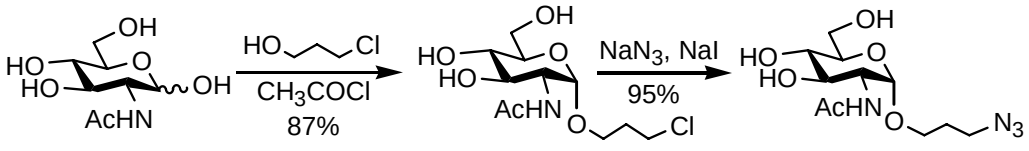
Synthesis of *GlcNAcβProN$_3$*



Peracetylated GlcNAc trichloroacetimidate (2.98 g, 6.06 mmol) was dissolved in CH_2Cl_2 (100 ml) with 4 Å molecular sieves (2 g). 3-Azidopropanol (1.83 g, 18.2 mmol) was added and the reaction mixture was cooled down to -40°C. After the mixture was stirred for 30 min, TMSOTf (0.2 eq) was added dropwise. The mixture was allowed to warm up to 0°C and stirred for 5 h. The mixture was filtered over Celite and concentrated. The crude material was purified by flash silica gel column (Hexane:EtOAc = 1:2, v/v) to yield preacetylated GlcNAcβProN$_3$ (1.89 g, 71%). ^{1}H NMR (400 MHz, D_2O) δ 6.52 (d, *J* = 8.8 Hz, 1H), 5.19 (t, *J* = 10 Hz, 1H), 4.94 (t, *J* = 10 Hz, 1H), 4.60 (d, *J* = 8 Hz, 1H), 4.18–4.01 (m, 2H), 3.85–3.48 (m, 4H), 3.29–3.26 (m, 2H), 1.97, 1.92, 1.91, 1.85 (s, 12H), 1.81–1.69 (m, 2H). ^{13}C NMR (100 MHz, D_2O) δ 170.94, 170.82, 169.61, 101.02, 72.61, 71.79, 69.01, 66.45, 62.40, 54.50, 48.15, 29.04, 23.32, 20.89, 20.83, 20.76.

Peracetylated GlcNAcβProN$_3$ (1.89 g, 4.39 mmol) was dissolved in dry methanol (50 mL) containing a catalytic amount of sodium methoxide. The resulted mixture was stirred at r.t. for overnight. The reaction mixture was then neutralized with DOWEX HCR-W2 (H^+) resin, filtered and concentration to give product **2** GlcNAcβProN$_3$ (1.32 g, 99%). ^{1}H NMR (400 MHz, D_2O) δ 4.49 (d, *J* = 8.4 Hz, 1H), 3.97–3.89 (m, 2H), 3.75–3.62 (m, 3H), 3.53 (t, *J* = 8.4 Hz, 1H), 3.43–3.36 (m, 4H), 2.04 (s, 3H), 1.86–1.80 (m, 2H). ^{13}C NMR (100 MHz, D_2O) δ 177.23, 103.86, 78.56, 76.45, 72.62, 69.78, 63.44, 58.28, 50.49, 30.82, 24.86.

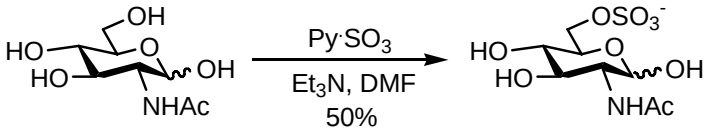
Synthesis of GlcNAcαProN$_3$



To a solution of *N*-acetyl-D-glucosamine (1.0 g, 4.52 mmol) in 3-chloropropanol (15 mL) was added dropwisely acetyl chloride (0.43 g, 5.48 mmol) at 0°C. The reaction mixture was heated at 80°C for 6 h. The solution was concentrated and the residue was purified by silica gel chromatography to yield 3-chloropropyl GlcNAc (1.17 g, 87%) as a syrup residue. 3-Chloropropyl GlcNAc (1.0 g, 3.36 mmol) was then dissolved in CH_3CN (3 mL) by heating the

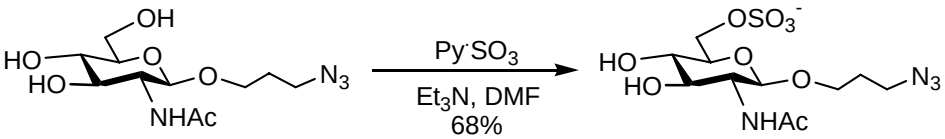
solution. NaN_3 (0.23 g, 3.53 mmol) and NaI (60 mg, 0.4 mmol) were then added. The reaction mixture was heated at 60°C for 7 h. The reaction mixture was concentrated and the residue was purified using a short silica gel column with EtOAc:MeOH = 5:1 as mobile phase to yield GlcNAcαProN_3 (0.97 g, 95%) as an solid. ^{1}H NMR (600 MHz, D_2O) δ 4.75 (d, 1H, *J* = 4.2 Hz, H-1), 3.74–3.57 (m, 5H), 3.42–3.31 (m, 3H), 1.93 (s, 3H), 1.78 (m, 2H); ^{13}C NMR (150 MHz, D_2O) δ 174.55, 97.11, 72.07, 71.16, 70.21, 65.12, 60.76, 53.93, 48.38, 28.18, 22.16.

Synthesis of *GlcNAc6S*



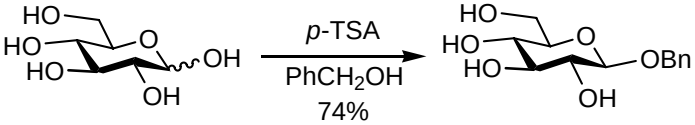
GlcNAc (300 mg, 1.36 mmol) was dissolved in 15 mL of dry DMF. Dry Et_3N (5 mL) and sulfur trioxide pyridine (1.2 eq.) were then added at 0°C. After stirred for overnight at rt, the reaction was quenched by adding MeOH and concentrated. The residue was purified by flash column chromatography (EtOAc:MeOH:H_2O = 6:2:1, by volume) to afford 6-*O*-sulfo-GlcNAc (204 mg, 50%).[11] ^{1}H NMR (400 MHz, D_2O) δ 5.06 (d, *J* = 2.8 Hz, 0.6H), 4.60 (d, *J* = 8.4 Hz, 0.4H), 4.21–3.35 (m, 6H), 1.91 (m, 3H). ^{13}C NMR (100 MHz, D_2O) δ 174.70, 91.13, 70.84, 69.96, 69.81, 67.36, 54.13, 22.14.

Synthesis of *GlcNAc6SβProN_3*



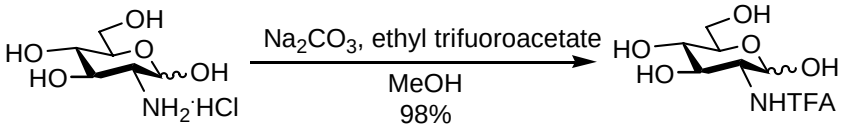
GlcNAcβProN_3 (200 mg, 0.66 mmol) was dissolved in 15 mL dry DMF. Dry Et_3N (5 mL) and sulfur trioxide pyridine (259 mg, 1.63 mmol) were added at 0°C. After stirred for overnight at rt, the reaction was quenched by adding MeOH and concentrated. The residue was purified by flash column chromatography (EtOAc:MeOH:H_2O = 7:2:1, by volume) to afford 6-*O*-sulfo-GlcNAcβProN_3 (171 mg, 68%).[11] ^{1}H NMR (600 MHz, D_2O) δ 4.37 (d, *J* = 8.4 Hz, 1H), 4.19 (dd, *J* = 1.8 Hz and 11.4 Hz, 1H), 4.07 (dd, *J* = 5.4 Hz and 11.4 Hz, 1H), 3.81–3.79 (m, 1H), 3.56–3.33 (m, 5H), 3.23–3.20 (m, 2H), 1.89 (s, 3H), 1.69–1.67 (m, 2H). ^{13}C HMR (150 MHz, D_2O) δ 174.70, 101.40, 73.79, 73.75, 69.66, 67.40, 67.14, 55.64, 47.94, 47.94, 28.27, 22.32.

Synthesis of *GlcβOBn*



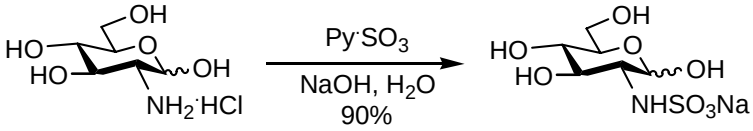
Glucose (12 g, 66.6 mmol) was dissolved in 20 mL of benzyl Alcohol followed by an addition of *p*-toluenesulfonic acid (*p*-TSA, 3.8 g, 20.0 mmol). The reaction mixture was refluxed at 80°C for five hours. The residue was purified by flash column chromatography (EtOAc:MeOH = 5:1) to afford GlcβOBn (13.4 g, 74%). The GlcβOBn was further isolated by (EtOAc:MeOH = 9:1). ^{1}H NMR (400 MHz, D_2O) δ 7.43–7.36 (m, 5H), 4.89 (d, *J* = 11.6 Hz, 1H), 4.70 (d, *J* = 11.6 Hz, 1H), 4.47 (d, *J* = 7.6 Hz, 1H), 3.89 (d, *J* = 12.4 Hz, 1H), 3.72–3.27 (m, 5H). ^{13}C NMR (100 MHz, D_2O) δ 136.73, 128.86, 128.61, 101.34, 76.05, 75.92, 73.25, 71.56, 69.79, 60.91.

Synthesis of *GlcNHTFA*



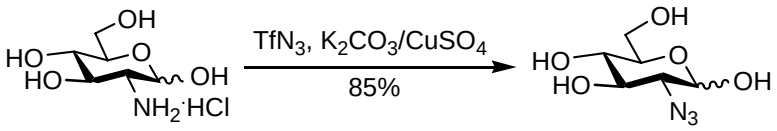
Glucosamine HCl (1.02 g, 4.73 mmol) was dissolved in dry MeOH (10 mL). Na_2CO_3 (1.00 g, 9.46 mmol) was added into the solution, followed by ethyl trifluoroacetate (1.34 mL, 9.46 mmol). The mixture was stirred at room temperature for overnight. The residue was purified by flash column chromatography (EtOAc:MeOH:H_2O = 10:2:0.2, by volume) to afford GlcNHTFA (1.28 g, 98%). ^{1}H NMR (600 MHz, D_2O) δ 5.30 (d, *J* = 3.6 Hz, 0.6H), 4.85 (d, *J* = 8.4 Hz, 0.4H), 4.04–3.73 (m, 4H), 3.69–3.51(m, 2H). ^{13}C NMR (150 MHz, D_2O) δ 160.17 (*J* = 37.4 Hz), 159.92 (*J* = 37.4 Hz), 118.92 (*J* = 284.5 Hz), 118.86 (*J* = 284.5 Hz), 94.43, 90.58, 76.22, 73.45, 71.78, 70.43, 70.30, 70.06, 60.91, 60.81, 57.47, 54.95.

Synthesis of *GlcNS*



Glucosamine hydrochloride (1.1 g, 5.10 mmol) was dissolved in 30 mL water, and the solution was adjusted to pH = 9.5 by the addition of 2N NaOH (aq). Sulfur trioxide-pyridine complex (0.97 g, 6.12 mmol) was added in three equal portions during 35 minutes intervals at room temperature, and the pH was maintained at 9.5 throughout the whole process using 2N NaOH (aq). After 2 hrs, the reaction mixture was concentrated and purified using silica gel column (EtOAc:MeOH:H_2O = 5:2:1, by volume) to obtain the $GlcNHSO_3Na$ in 90% yield (1.36 g).[12] ^{1}H NMR (600 MHz, D_2O) δ 5.29 (d, *J* = 3.6 Hz, 0.6H), 4.54 (d, *J* = 8.4 Hz, 0.4H), 3.75–3.58 (m, 3H), 3.48–3.28 (m, 2H), 3.06 (dd, *J* = 3 Hz and 7.2 Hz, 0.6 H), 2.84 (dd, *J* = 2.4 Hz and 8.4 Hz). ^{13}C NMR (150 MHz, D_2O) δ 95.72, 91.43, 75.93, 74.85, 71.47, 71.27, 70.27, 70.04, 61.13, 60.90, 60.79, 58.13.

Synthesis of *$GlcN_3$*



Sodium azide (3.0 g, 46.12 mmol) was dissolved in 8 mL of water and 13 mL of dichloromethane. Reaction flask was cooled to 0°C and trifluoromethane-sulfonic anhydride (2.61 mL, 9.25 mmol) was added drop wisely. After stirring the reaction mixture for two hours, the sample was extracted twice with CH_2Cl_2 (15 mL × 2). The combined organic phase was washed with saturated sodium carbonate. Glucosamine hydrochloride (1.3 g, 6.03 mmol) was dissolved in 20 mL of water. Potassium carbonate (1 g, 7.24 mmol), $CuSO_4{\cdot}5H_2O$ (14 mg, 0.06 mmol), methanol (30 mL), and 30 mL of the extract from above were added and the mixture was stirred for overnight. The product was purified by flash column chromatography (EtOAc:MeOH = 2:1, by volume) to afford $GlcN_3$ (1.04 g, 85%), ^{1}H NMR (600 MHz, D_2O) δ 5.31 (d, *J* = 3 Hz, 0.4H), 4.67 (d, *J* = 8.4 Hz, 0.6H), 3.88–3.69 (m, 3H), 3.49–3.42 (m, 2.4H), 3.25 (t, *J* = 8.4, 0.6 H). ^{13}C NMR (150 MHz, D_2O) δ 92.58, 88.67, 73.49, 71.85, 69.08, 68.98, 67.28, 66.97, 64.39, 61.05, 58.11, 57.95.

Enzymatic synthesis of β1–4 linked galactosides

General one-pot four-enzyme approach

General procedures for one-pot four-enzyme preparative synthesis of β1–4-linked galactosides, **12–22**: Glc-1-P (1.3 eq.), monosaccharides **1–11** (1 eq.) and UTP (1.3 eq.) were dissolved in water in a 50 mL centrifuge tube containing Tris-HCl buffer (100 mM, pH 7.5) and $MgCl_2$ (20 mM). After addition of appropriate amounts of GalU (2.4–4.6 mg), GalE (2.8–3.5 mg), inorganic pyrophosphatase (1 mg), and a β1–4-galactosyltransferase (2–4 mg for NmLgtB or 1–2 mg for Hp1–4GalT), water was added to bring the final volume of the reaction mixture to 10 mL. The reaction was carried out by incubating the solution in an incubator for 24 to 48 h at 37°C with gentle shaking. Product formation was monitored by TLC (EtOAc:MeOH:H_2O = 5:2:1 by volume, *p*-anisaldehyde sugar stain). The reaction was stopped by adding the same volume of ice-cold ethanol and incubating at 4°C for 30 min. The mixture was then concentrated and passed through a BioGel P-2 gel filtration column to obtain the desired product. Silica gel purification (EtOAc:MeOH:H_2O = 5:2:0.5) was applied when necessary to achieve further purification.

Compound **12**, *Galβ1–4GlcNAc*. Yield 64% for NmLgtB, 67% for Hp1–4GalT, white foam; ^{1}H NMR (600 MHz, D_2O) δ 5.19 (m, 0.6H), 4.71 (dd, *J* = 1.8 Hz and 7.8 Hz, 0.4H), 4.46 (d, *J* = 7.2 Hz, 1H), 3.97–3.65 (m, 10 H), 3.59–3.51 (m, 2H), 2.03 (m, 3H) ^{13}C HMR (150 MHz, D_2O) δ 174.83, 174.57, 103.05, 102.99, 94.99, 90.66, 78.91, 78.46, 75.48, 74.97, 72.64, 71.10, 70.39, 69.40, 68.68, 61.16, 60.20, 60.08, 56.33, 53.84, 22.32, 22.03.22.03. HRMS (ESI) m/z calculated for $C_{14}H_{26}NO_{11}$ (M+H) 384.1506, measured 384.1505.

Compound **13,** *Galβ1–4GlcNAcβProN*$_3$. Yield 85% for NmLgtB, 70% for Hp1–4GalT, white foam; ^{1}H NMR (400 MHz, D_2O) δ 4.33 (d, *J* = 6.8 Hz, 1H), 4.30 (d, *J* = 6.8 Hz, 1H), 3.83–3.75 (m, 3H), 3.68–3.36 (m, 10H), 3.21 (t, *J* = 6.8 Hz, 2H), 1.87 (s, 3H), 1.68–1.65 (m, 2H). ^{13}C HMR (100 MHz, D_2O) δ 174.61, 103.07, 101.27, 78.69, 75.52, 74.91, 72.70, 72.54, 71.14, 68.72,

67.29, 61.19, 60.26, 55.26, 47.96, 28.29, 22.37. HRMS (ESI) m/z calculated for $C_{17}H_{31}N_4O_{11}$ (M+H) 467.1989, measured 467.1982.

Compound **14**, *Galβ1–4GlcNAcαProN*$_3$. Yield 51% for NmLgtB, white foam; ^{1}H NMR (600 MHz, D_2O) δ 4.87 (d, *J* = 3.6 Hz, 1H), 4.46 (d, *J* = 7.8 Hz, 1H), 3.94–3.69 (m, 11H), 3.66 (dd, *J* = 3.6 and 10.2 Hz, 1H), 3.54–3.41 (m, 4H), 2.03 (s, 3H), 1.90–1.86 (m, 2H). ^{13}C HMR (150MHz, D_2O) δ 174.51, 103.07, 96.77, 78.91, 75.51, 72.69, 71.13, 70.71, 69.70, 68.70, 65.14, 61.18, 60.05, 53.52, 48.25, 28.10, 22.03. HRMS (ESI) m/z calculated for $C_{17}H_{31}N_4O_{11}$ (M+H) 467.1989, measured 467.1988.

Compound **15**, *Galβ1–4GlcNAcβMU*. Yield 68% for NmLgtB, 66% for Hp1–4GalT, white foam; ^{1}H NMR (600 MHz, D_2O) δ 7.39 (d, *J* = 8.4 Hz, 1H), 6.86 (dd, *J* = 9.0 and 2.4 Hz, 1H), 6.70 (d, *J* = 1.8 Hz, 1H), 5.96 (s, 1H), 5.22 (d, *J* = 8.4 Hz, 1H), 4.55 (d, *J* = 8.4 Hz, 1H), 4.11–4.04 (m, 2H), 3.96–3.70 (m, 12H), 3.61–3.58 (m, 1H), 2.21 (s, 3H), 2.10 (s, 3H). ^{13}C HMR (150 MHz, D_2O) δ 174.98, 164.10, 159.45, 155.93, 126.68, 113.94, 111.35, 103.11, 98.75, 75.60, 75.23, 72.70, 72.26, 71.22, 69.75, 68.81, 61.67, 61.31, 60.04, 59.58, 55.09, 22.54, 18.14. HRMS (ESI) m/z calculated for $C_{24}H_{32}NO_{13}$ (M+H) 542.1874, measured 542.1866.

Comopound **16**, *Galβ1–4GlcNAc6S*. Yield 46% for Hp1–4GalT, light yellow foam; ^{1}H NMR (600 MHz, D_2O) δ 5.04 (d, *J* = 3.2 Hz, 0.4H), 4.37 (m, 0.6H), 4.24–4.13 (m, 2H), 4.02 (m, 0.4), 3.90–3.88 (m, 0.6H), 3.84–3.71(m, 3H), 3.68–3.56 (m, 3H), 3.54–3.46 (m, 2H), 3.40–3.34 (m, 2H), 1.84 (s, 3H). ^{13}C HMR (150 MHz, D_2O) δ 174.86, 174.57, 102.70, 97.40, 97.18, 95.10, 77.91, 77.74, 74.04, 73.76, 73.25, 72.80, 72.20, 71.15, 70.70, 69.42, 68.77, 68.45, 68.04, 67.61, 66.63, 66.57, 61.20, 61.08, 56.26, 53.73, 22.32, 22.04. HRMS (ESI) m/z calculated for $C_{14}H_{25}NO_{14}S^-Na^+$ (M+Na) 486.0893, measured 486.0887.

Compound **17**, *Galβ1–4GlcNAc6SβProN*$_3$. Yield 70% for Hp1–4GalT, white foam; ^{1}H NMR (600 MHz, D_2O) δ 4.43–4.40 (m, 2H), 4.28 (d, *J* = 10.8 Hz, 1H), 4.20 (dd, *J* = 2.4 Hz and 10.8 Hz, 1H), 3.85–3.80 (m, 2H), 3.67–3.54 (m, 9H), 3.40 (t, *J* = 10.8 Hz, 1H), 3.25–3.24 (s, 2H), 1.91 (s, 3H), 1.72 (m, 2H). ^{13}C HMR (150 MHz, D_2O) δ 174.65, 102.67, 101.32, 77.54, 75.50, 72.69, 72.63, 72.42, 71.14, 68.76, 67.41, 66.43, 61.20, 55.24, 47.93, 28.26, 22.32. HRMS (ESI) m/z calculated for $C_{17}H_{30}N_4O_{14}S^-Na^+$ (M+Na) 569.1377, measured 569.1358.

Compound **18**, *Galβ1–4Glc*. Yield 78% for NmLgtB, white foam; ^{1}H NMR (300 MHz, D_2O) δ 5.18 (d, *J* = 3.6 Hz, 0.4H), 4.41 (d, *J* = 7.8 Hz, 0.6H), 3.79–3.75 (m, 3H), 3.69–3.33 (m, 10H). ^{13}C HMR (75 MHz, D_2O) δ 103.00, 91.93, 78.49, 75.47, 72.62, 71.52, 71.25, 71.08, 70.21, 68.67, 61.18, 60.04. HRMS (ESI) m/z calculated for $C_{12}H_{23}O_{11}$ (M+H) 343.1240, measured 343.1242.

Compound **19**, *Galβ1–4GlcβOBn*. Yield 83% for for NmLgtB, white foam; ^{1}H NMR (600 MHz, D_2O) δ 7.51–7.46 (m, 5H), 4.98 (d, *J* = 12.0 Hz, 1H), 4.80 (d, *J* = 12.0 Hz, 1H), 4.60 (d, *J* =7.8 Hz, 1H), 4.49 (d, *J* = 7.8 Hz, 1H), 4.04–3.57 (m, 11H), 3.40 (t, *J* = 8.4 Hz, 1H). ^{13}C HMR (150 MHz, D_2O) δ 136.75 (1C), 128.92 (4C), 128.68 (1C), 103.13, 101.25, 78.54, 76.48, 74.97, 74.64, 73.05, 72.77, 71.73, 71.16, 68.76, 61.21, 60.33. HRMS (ESI) m/z calculated for $C_{19}H_{29}O_{11}$ (M+H) 433.1710, measured 433.1709.

Compound **20,** *Galβ1–4GlcNHTFA*. Yield 68% for NmLgtB, white foam; ^{1}H NMR (600 MHz, D_2O) δ 5.11 (d, *J* = 2.4 Hz, 0.3H), 4.31 (dd, *J* = 1.8 Hz, 0.7H), 3.86–3.71 (m, 3H), 3.68–3.55 (m, 5H), 3.51–3.46 (m, 2H), 3.39–3.35 (m, 1H), 3.05–3.01 (m, 2H). ^{13}C HMR (150 MHz, D_2O) δ 165.08, 103.08, 103.03, 94.28, 90.16, 78.81, 78.24, 75.52, 75.08, 72.64, 72.19, 72.05, 71.12, 70.40, 68.95, 68.70, 62.61, 61.18, 60.15, 60.04, 56.88, 54.49. HRMS (ESI) m/z calculated for $C_{14}H_{22}NO_{11}F_3Na$ (M+Na) 460.1043, measured 460.1035.

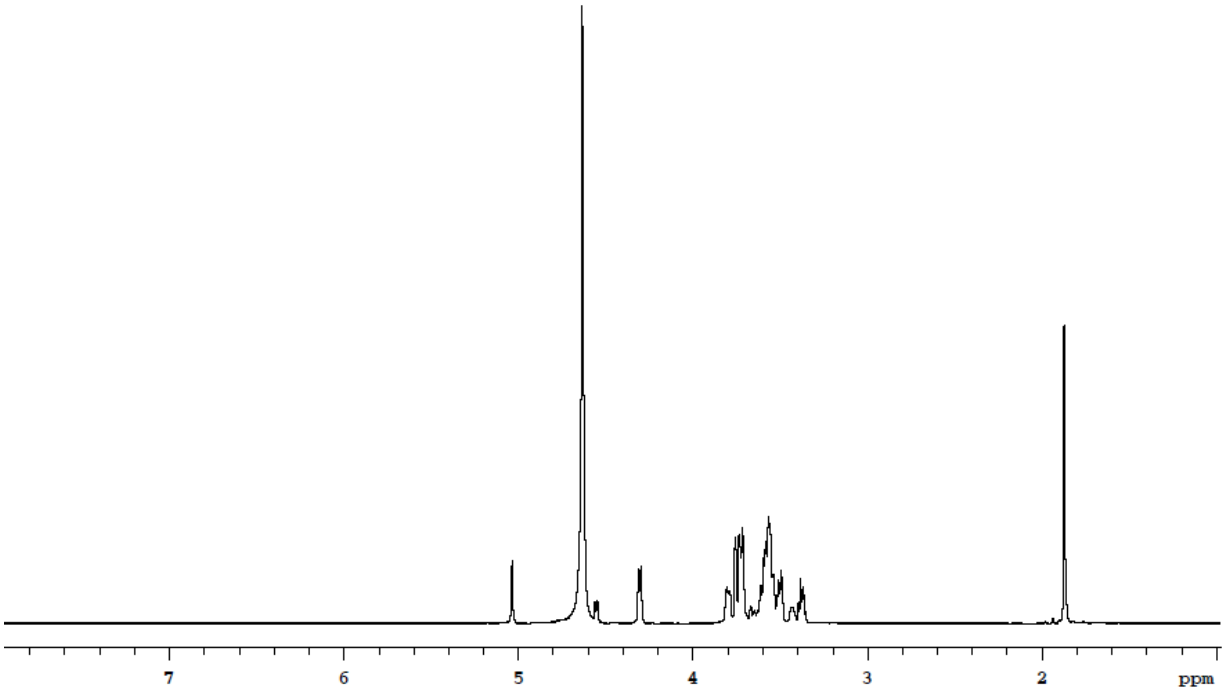
Compound **21**, *Galβ1–4GlcNS*. Yield 74% for NmLgtB, white foam; ^{1}H NMR (600 MHz, D_2O) δ 5.44 (d, *J* = 3.6 Hz, 0.8H), 4.71 (d, *J* = 8.4 Hz, 0.2H), 4.45 (d, *J* = 8.4 Hz, 1H), 3.94–3.83 (m, 4H), 3.80–3.68 (m, 5H), 3.66–3.61 (m, 2H), 3.54–3.51 (m, 1H). ^{13}C HMR (150 MHz, D_2O) δ 102.94, 102.84, 95.85, 91.13, 78.58, 78.10, 75.51, 74.96, 73.00, 72.67, 72.20, 71.16, 70.27, 69.72, 68.77, 62.62, 61.22, 60.87, 60.16, 59.46, 57.80. HRMS (ESI) m/z calculated for $C_{12}H_{23}NO_{13}S^{-}Na^{+}$ (M+Na) 444.0788, measured 444.0778.

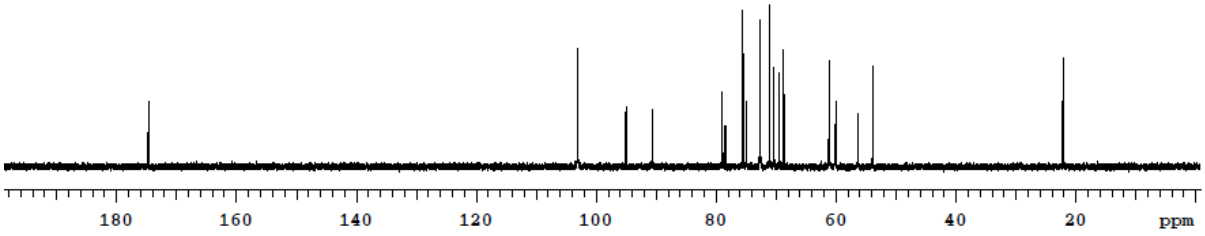
Compound **22**, *Galβ1–4GlcN$_3$*. Yield 81 % for for NmLgtB, white foam; ^{1}H NMR (600 MHz, D_2O) δ 5.29 (d, *J* = 3.6 Hz, 0.6H), 4.54 (d, *J* = 8.4 Hz, 0.4H), 3.75–3.66 (m, 3H), 3.62–3.56 (m, 2H), 3.48–3.41 (m, 1H), 3.33–3.27 (m, 2H), 3.06 (dd, *J* = 3.0 and 10.2 Hz, 0.6H), 2.85–2.82 (m, 0.4H). ^{13}C HMR (150 MHz, D_2O) δ 103.08, 95.11, 91.09, 78.57, 78.19, 75.51, 75.50, 75.02, 72.65, 72.21, 71.11, 71.08, 70.46, 70.33, 72.65, 72.21, 71.11, 71.08, 68.71, 68.68, 63.26, 62.62, 61.19, 61.16. HRMS (ESI) m/z calculated for $C_{12}H_{22}N_3O_{10}$ (M+H) 368.1305, measured 368.1305.

References:

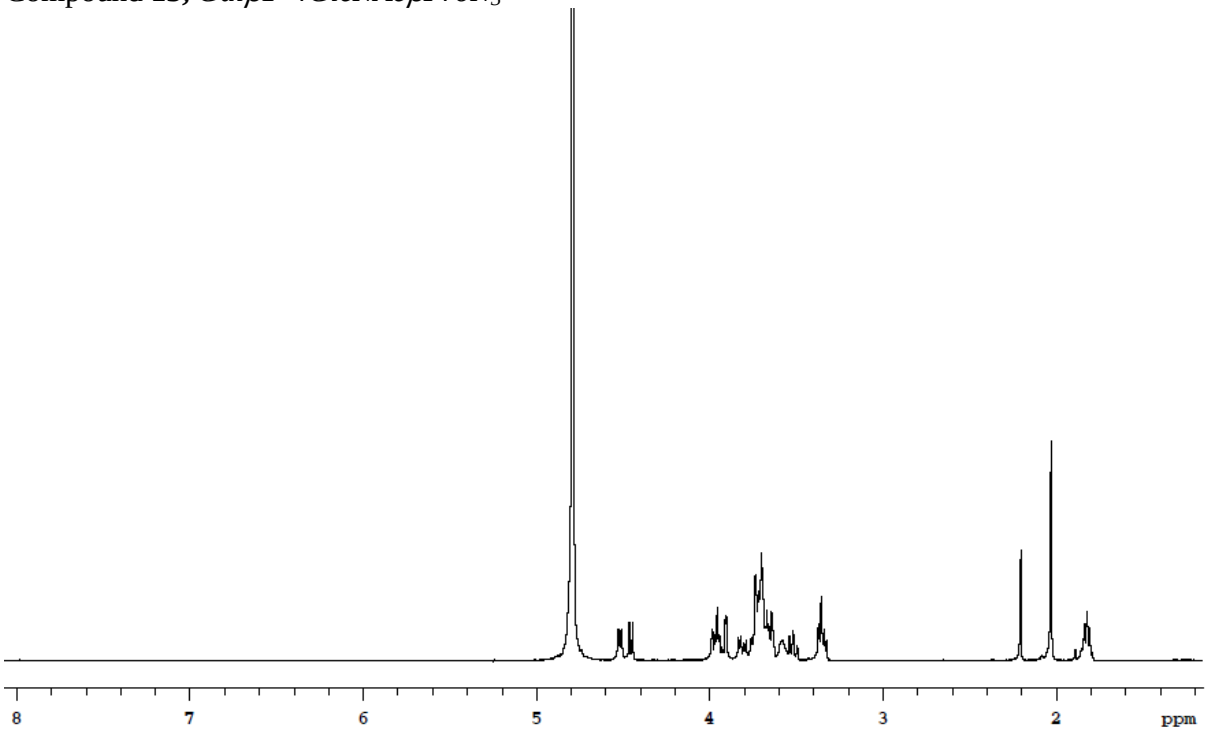
1. Oganesyan, N.; Ankoudinova, I.; Kim, S. H.; Kim, R., Effect of osmotic stress and heat shock in recombinant protein overexpression and crystallization. *Protein Expr. Purif.* **2007,** *52*, 280–285.
2. Yu, H.; Chokhawala, H.; Karpel, R.; Yu, H.; Wu, B.; Zhang, J.; Zhang, Y.; Jia, Q.; Chen, X., A multifunctional *Pasteurella multocida* sialyltransferase: a powerful tool for the synthesis of sialoside libraries. *J. Am. Chem. Soc.* **2005,** *127*, 17618–17619.
3. Boeggeman, E.; Qasba, P. K. Studies on the metal binding sites in the catalytic domain of beta1,4-galactosyltransferase. *Glycobiology* **2002**, *12*, 395–407.
4. Wakarchuk, W. W.; Cunningham, A.; Watson, D. C.; Young, N. M., Role of paired basic residues in the expression of active recombinant galactosyltransferases from the bacterial pathogen *Neisseria meningitidis*. *Protein Eng.* **1998,** *11*, 295–302.
5. Nakazawa, K.; Furukawa, K.; Narimatsu, H.; Kobata, A., Kinetic study of human beta-1,4-galactosyltransferase expressed in *E. coli*. *J. Biochem.* **1993,** *113*, 747–753.
6. Boeggeman, E. E.; Balaji, P. V.; Sethi, N.; Masibay, A. S.; Qasba, P. K., Expression of deletion constructs of bovine beta-1,4-galactosyltransferase in *Escherichia coli*: importance of Cys134 for its activity. *Protein Eng.* **1993,** *6*, 779–785.
7. Ramasamy, V.; Ramakrishnan, B.; Boeggeman, E.; Ratner, D. M.; Seeberger, P. H.; Qasba, P. K., Oligosaccharide preferences of beta1,4-galactosyltransferase-I: crystal structures of Met340His mutant of human beta1,4-galactosyltransferase-I with a pentasaccharide and trisaccharides of the N-glycan moiety. *J. Mol. Biol.* **2005,** *353*, 53–67.
8. Endo, T.; Koizumi, S.; Tabata, K.; Ozaki, A., Cloning and expression of beta1,4-galactosyltransferase gene from *Helicobacter pylori*. *Glycobiology* **2000,** *10*, 809–813.
9. Namdjou, D. J.; Chen, H. M.; Vinogradov, E.; Brochu, D.; Withers, S. G.; Wakarchuk, W. W., A beta-1,4-galactosyltransferase from *Helicobacter pylori* is an efficient and versatile biocatalyst displaying a novel activity for thioglycoside synthesis. *Chembiochem* **2008,** *9*, 1632–1640.
10. Morrison, J. F.; Ebner, K. E., Studies on galactosyltransferase. Kinetic investigations with glucose as the galactosyl group acceptor. *J. Biol. Chem.* **1971,** *246*, 3985–3991.
11. Guiseley, K. B.; Ruoff, P. M., Monosaccharide sulfates. I. Glucose 6-sulfate. Preparations, characterization of the crystalline potassium salt, and kinetic studies. *J. Org. Chem.* **1961,** 26, 1248–1254.
12. Polat, T.; Wong, C. H., Anomeric reactivity-based one-pot synthesis of heparin-like oligosaccharides. *J. Am. Chem. Soc.* **2007,** 129, 12795–12800.

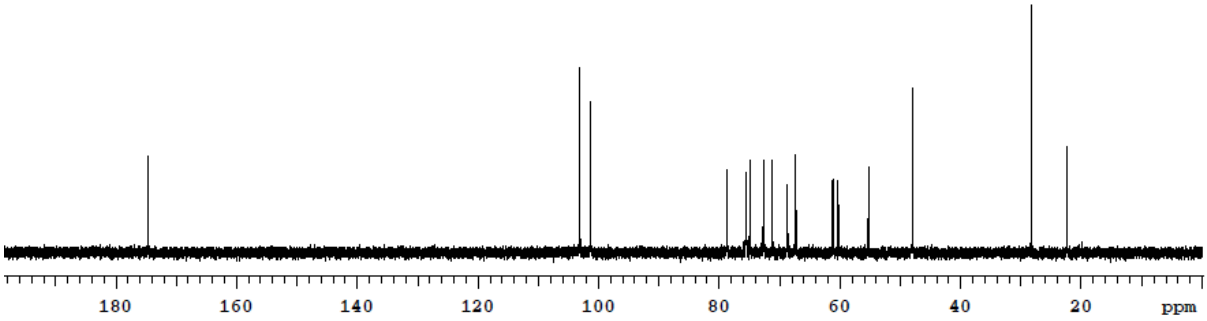
Compound **12**, *Galβ1–4GlcNAc*



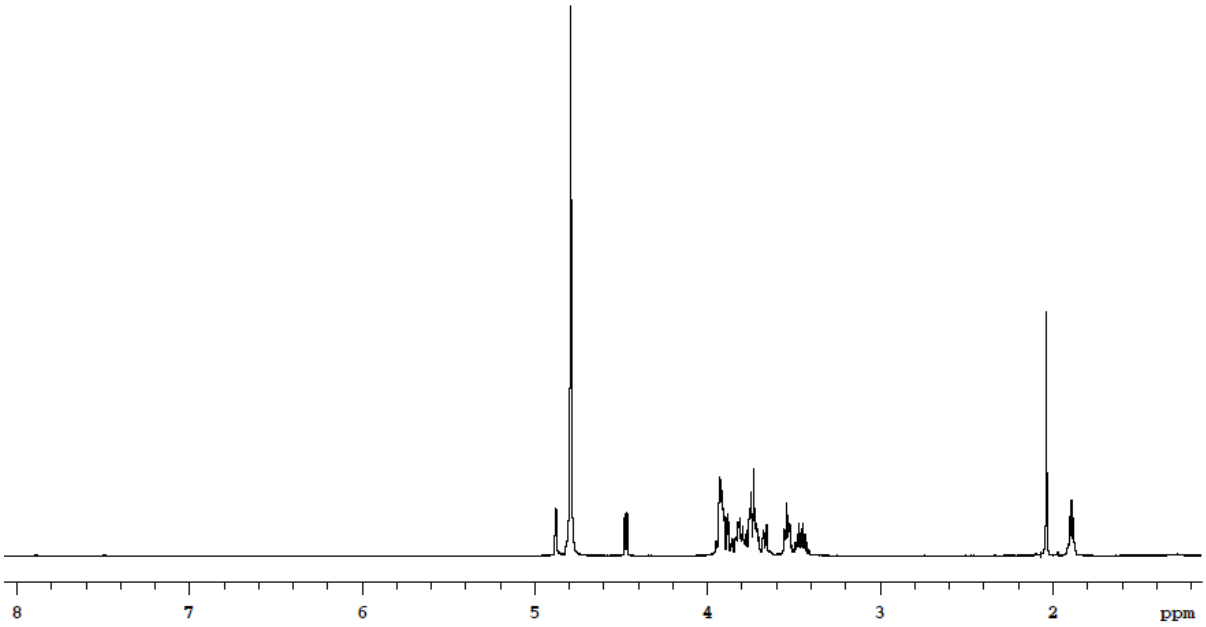


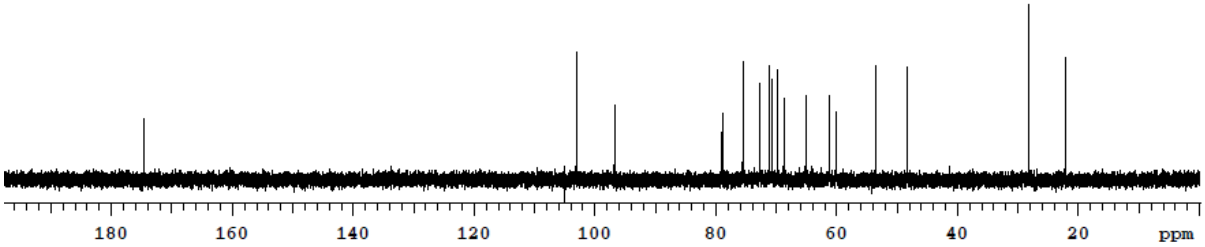
Compound **13,** *Galβ1–4GlcNAcβProN*$_3$



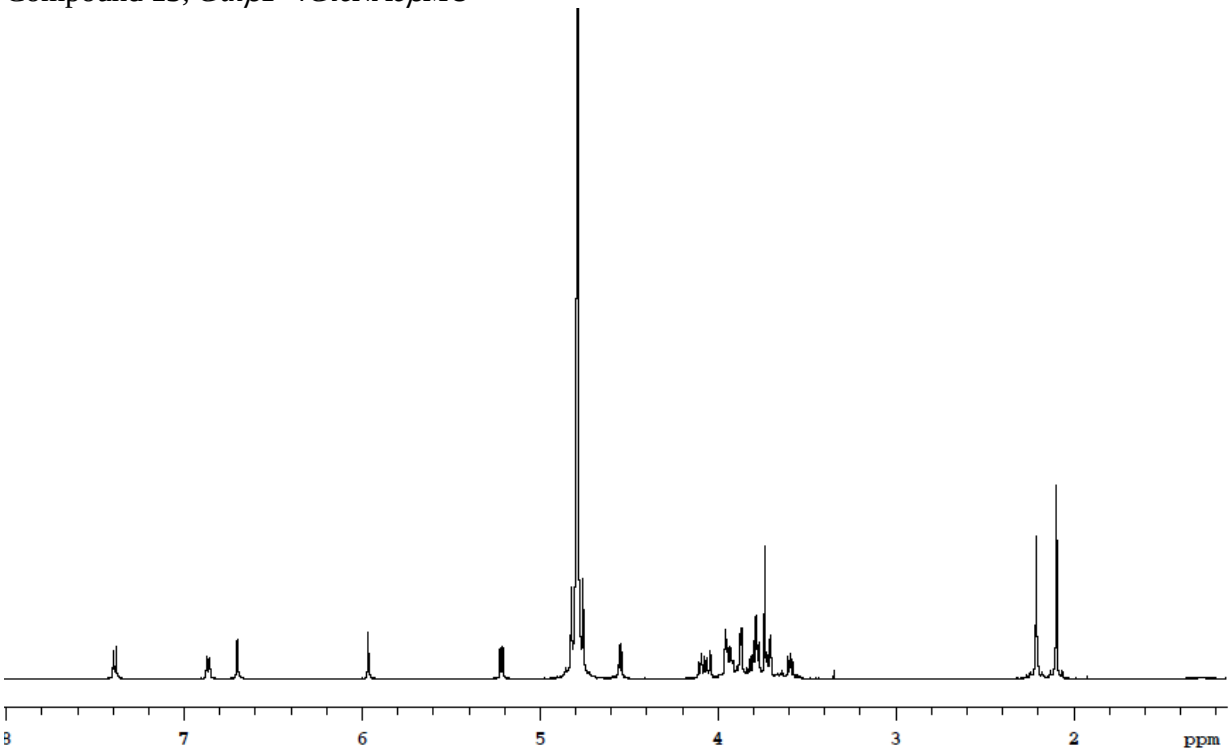


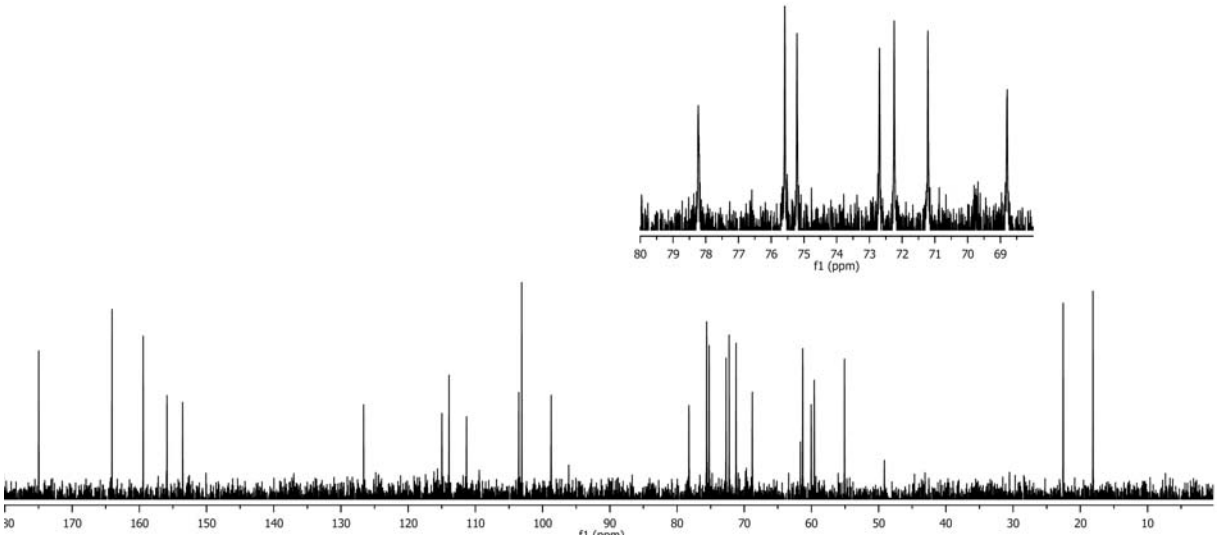
Compound **14**, *Galβ1–4GlcNAcαProN*$_3$



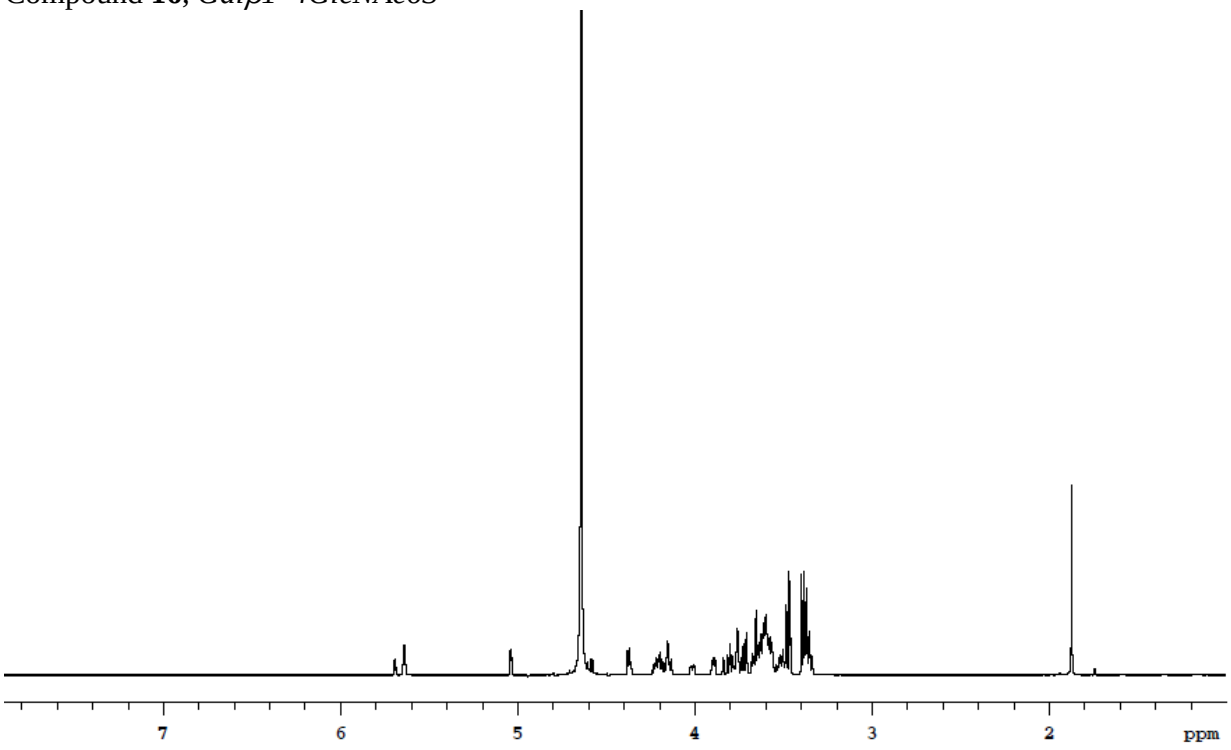


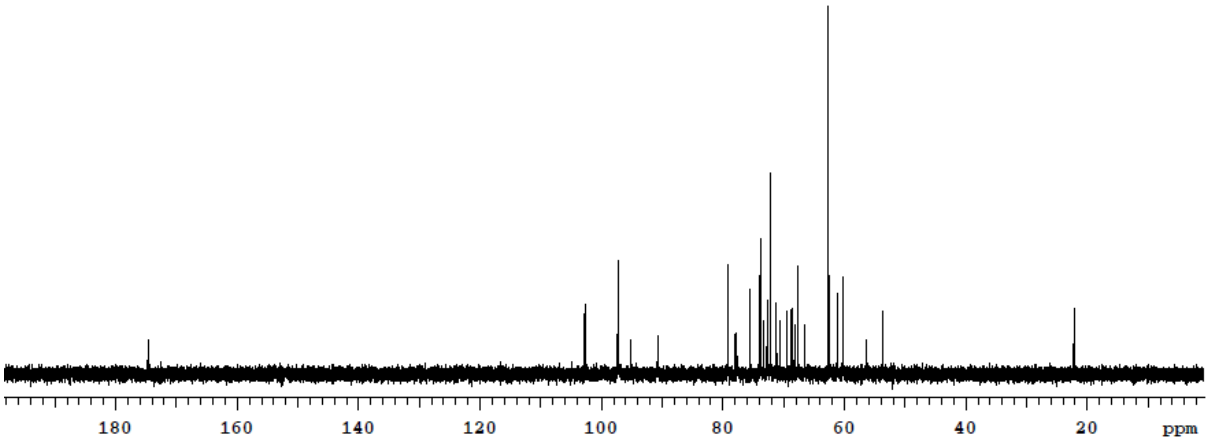
Compound **15**, *Galβ1–4GlcNAcβMU*



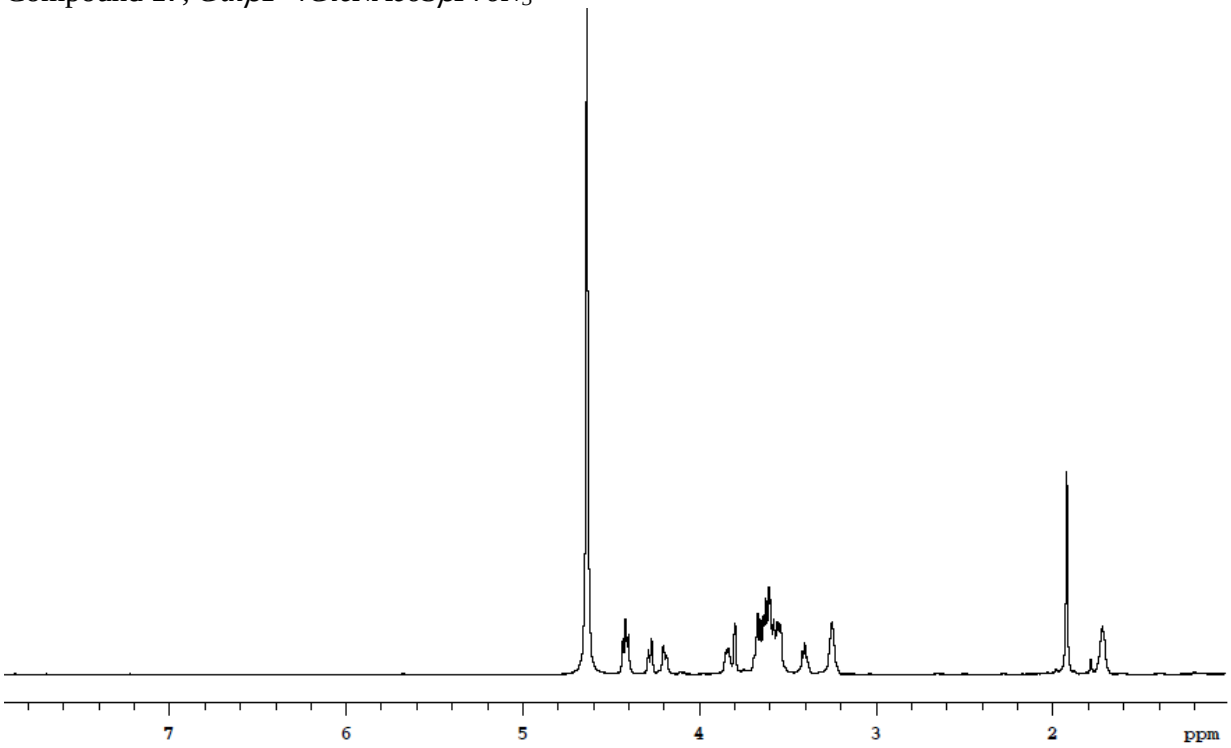


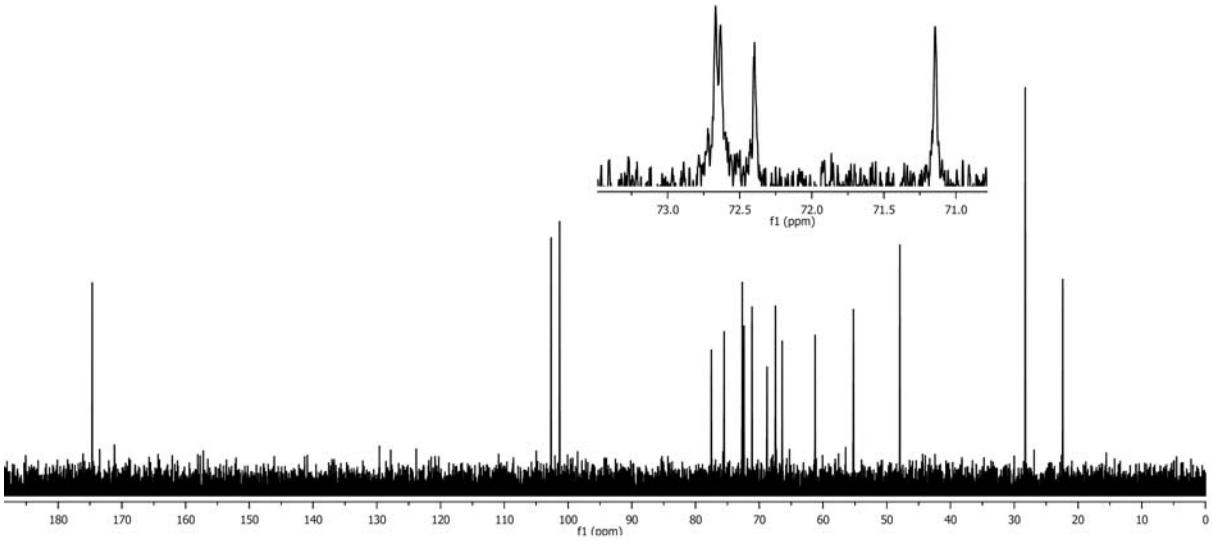
Compound **16**, *Galβ1–4GlcNAc6S*



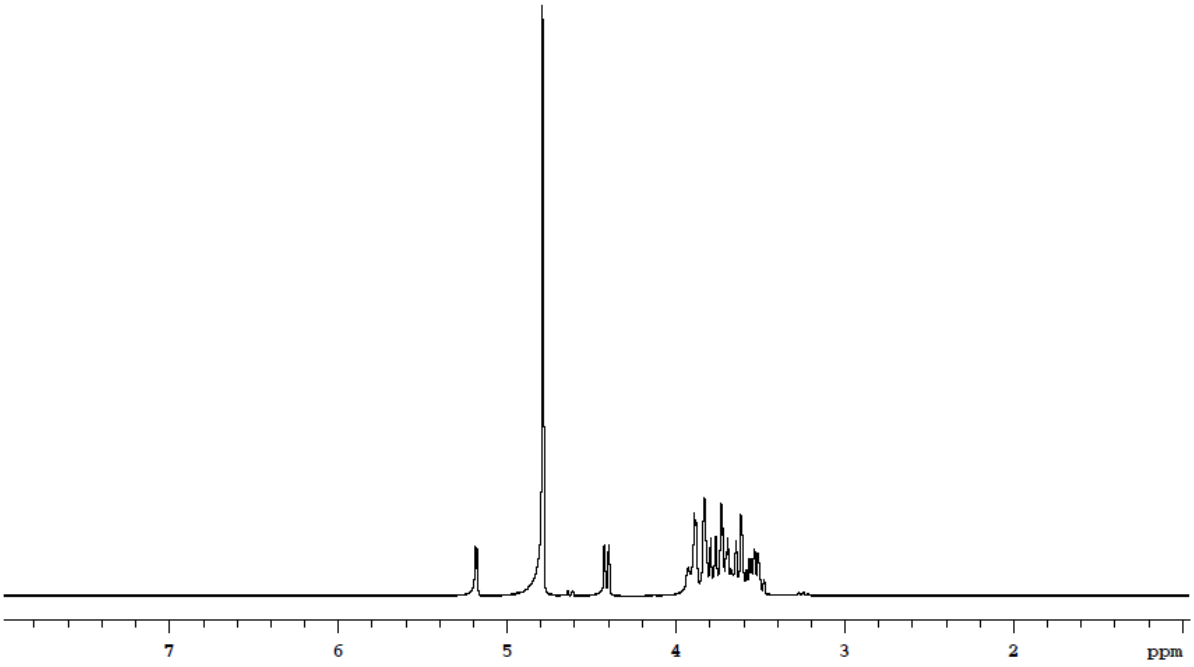


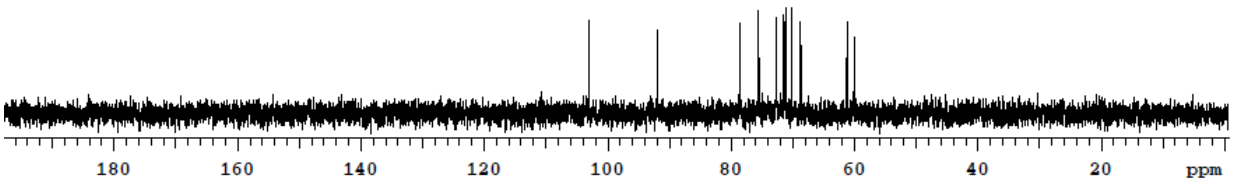
Compound **17**, *Galβ1–4GlcNAc6SβProN*$_3$



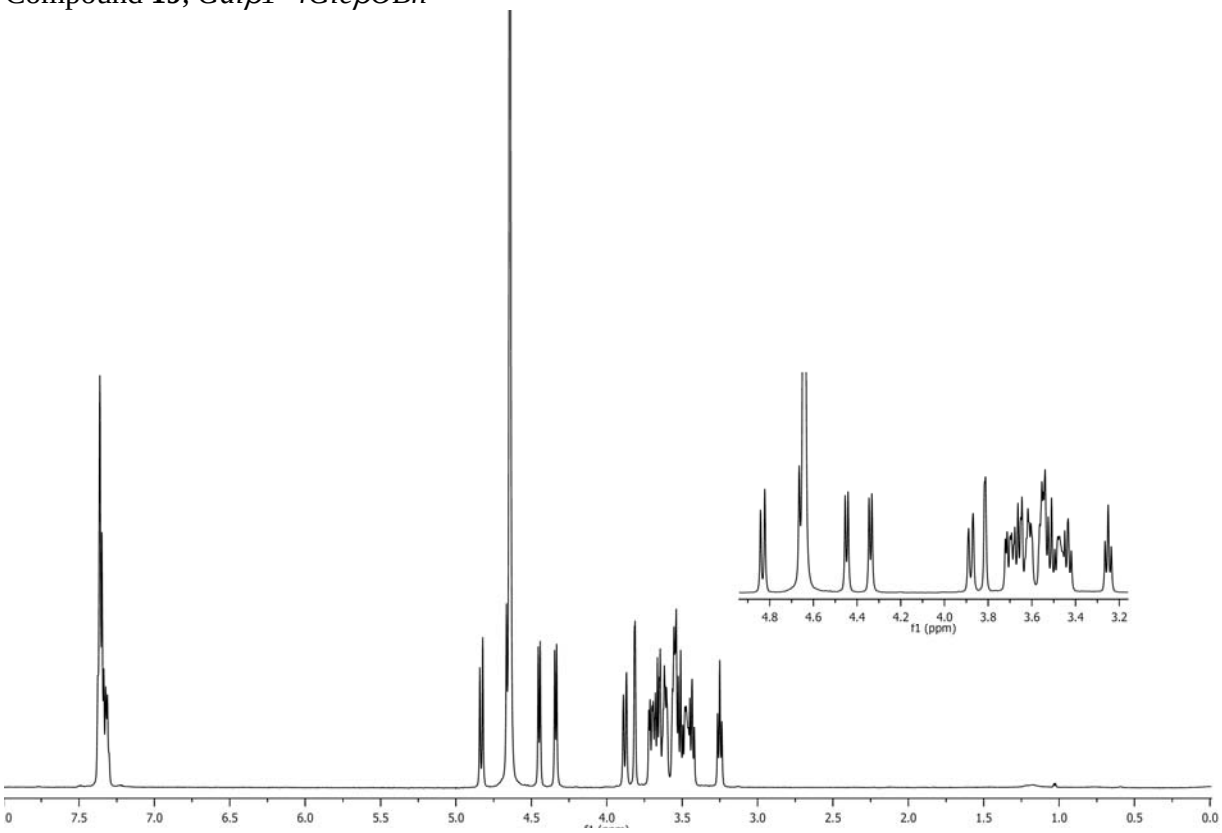


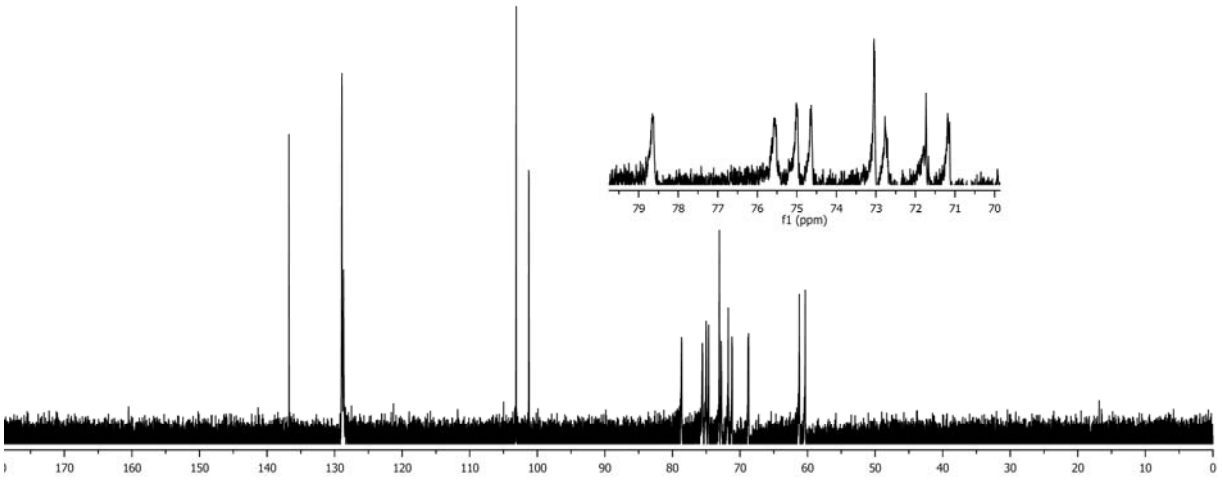
Compound **18**, *Galβ1–4Glc*



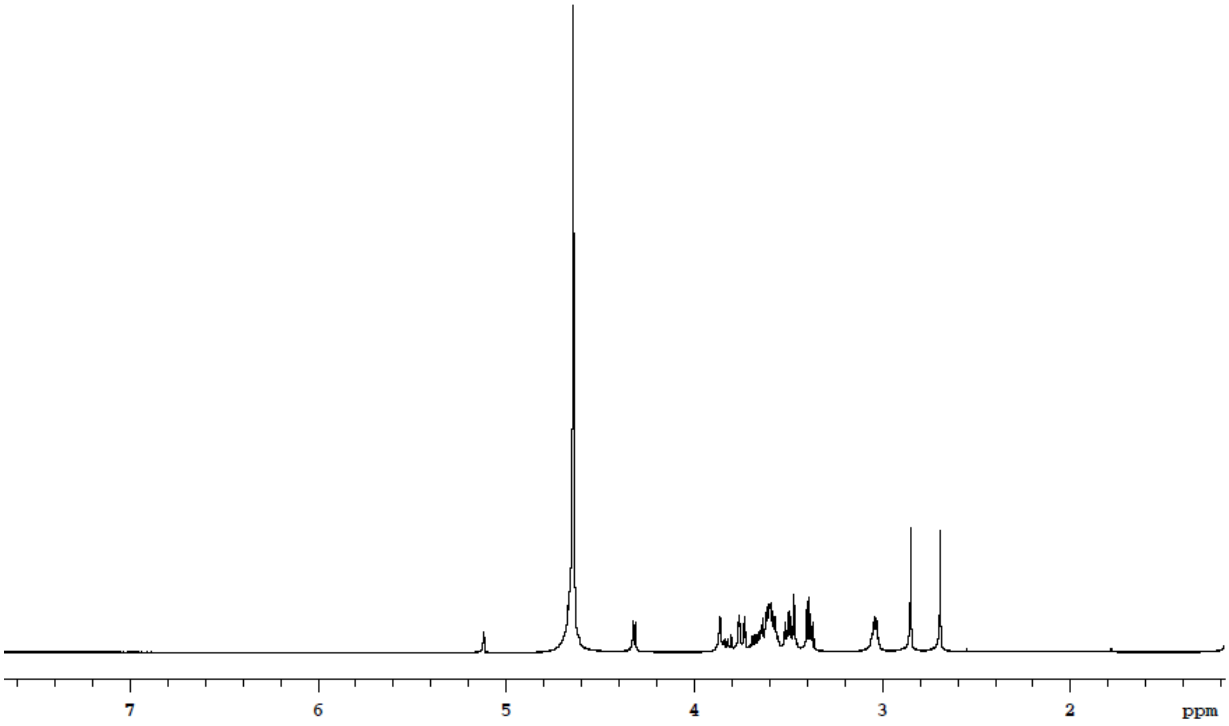


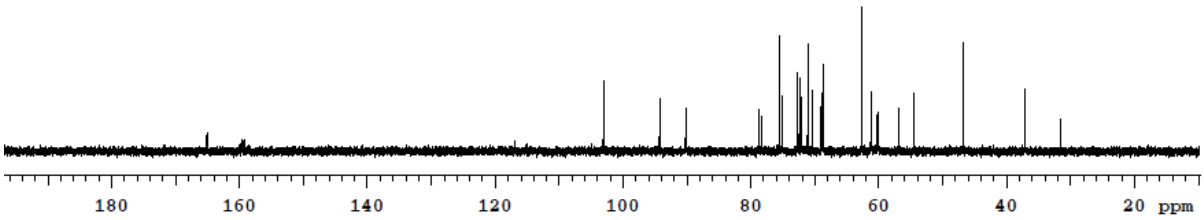
Compound **19**, *Galβ1–4GlcβOBn*



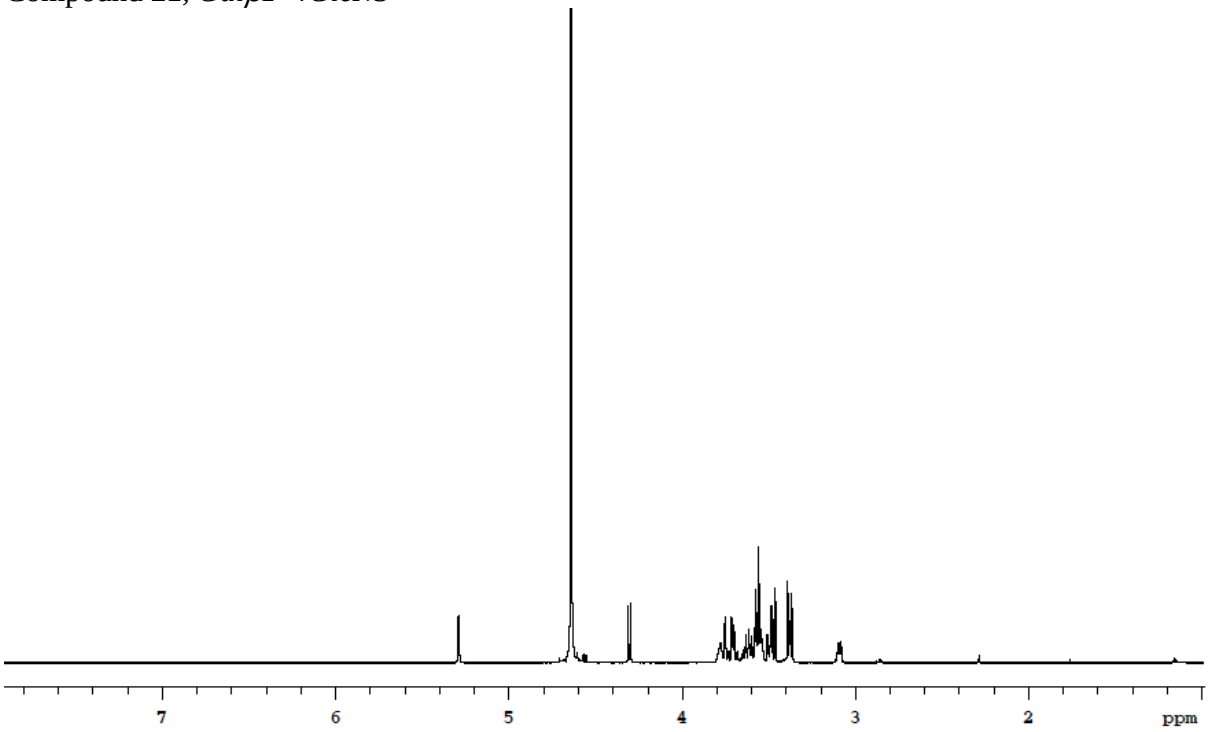


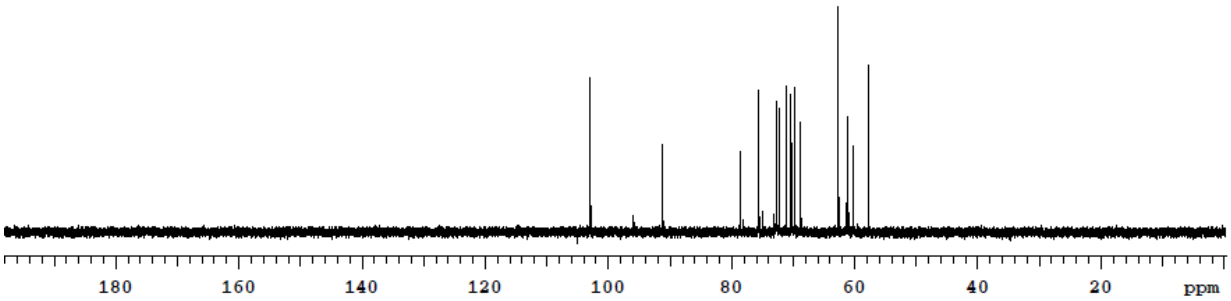
Compound **20,** *Galβ1–4GlcNHTFA*





Compound **21**, *Galβ1–4GlcNS*





Compound **22**, *Galβ1–4GlcN*$_3$