SUPPORTING INFORMATION

Convergent Chemoenzymatic Synthesis of O-GalNAc Rare Cores 5, 7, 8 and Their Sialylated Forms

Madhusudhan Reddy Gadi,[1†] Congcong Chen,[1,2†] Shumin Bao,[1†] Shuaishuai Wang,[1] Yuxi Guo,[1] Jinghua Han,[1] Weidong Xiao,[3] and Lei Li[1*]

[1]Department of Chemistry and Center for Diagnostics & Therapeutics, Georgia State University, Atlanta, GA 30303, USA

[2]Shandong Academy of Pharmaceutical Science, Key Laboratory of Biopharmaceuticals, Engineering Laboratory of Polysaccharide Drugs, National-Local Joint Engineering Laboratory of Polysaccharide Drugs, Jinan 250101, China

[3]Department of Pediatrics, Herman B Wells Center for Pediatric Research, Indiana University School of Medicine, Indianapolis, IN 46202, USA

[†]Equal Contribution

*Correspondence: lli22@gsu.edu (L.L.), ORCID: 0000-0002-1146-0761

Table of Contents

I. Materials and biochemicals……Page S2

II. Chemical synthesis of O-GalNAc cores 5, 7, 8……Page S2

III. Enzymatic assays and synthetic procedures……Page S8

IV. HPLC and MS analysis of sialyltransferase-catalyzed reactions……Page S10

V. Supporting figures ……Page S18

VI. Microarray fabrication, assay, and data……Page S20

VII. HPLC, MS, and NMR data of purified compounds……Page S23

VIII. NMR spectra……Page S35

IX. References……Page S47

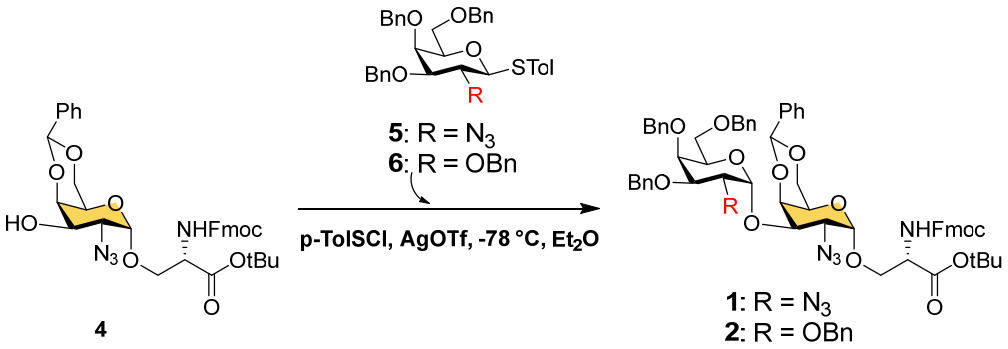
I. Materials and biochemicals

Unless otherwise stated, chemicals were purchased from Sigma-Aldrich and used without further purification. Enzymes including *Neisseria meningitidis* CMP-sialic acid synthetase (NmCSS),[1] *Photobacterium damsela* α2-6 sialyltransferase (Pd26ST)[2], and *E. coli* Aldolase[3] were expressed and purified as previously described. N-terminal GFP-tagged human ST6GalNAc1, 2, 4, 5, 6 were purchased from Glyco Expression Technologies (Athens, GA), C-terminal His_6-tagged human ST6GalNAc4 was obtained from R&D Systems (Minneapolis, MN).

II. Chemical synthesis of O-GalNAc cores 5, 7, 8

ESI-mass spectrometry was performed on an LTQ-Orbitrap Elite mass spectrometer equipped with EASY-spray source and nano-LC Ultimate 3000 high-performance liquid chromatography system (Thermo Fisher). Anhydrous dichloromethane (DCM), TMSOTf, TfOH, PTSA×H_2O, and solid sodium methoxide were purchased from Sigma Aldrich. TFA was purchased from Alfa Aesar. ^1H and ^{13}C NMR spectra were recorded on a Bruker AVANCE 400 (400 MHz) or Bruker AVANCE 600 (600 MHz) spectrometer at 25°C. All ^1H Chemical shifts (in ppm) were assigned according to $CDCl_3$ (δ= 7.24 ppm), MeOD (δ = 4.79 ppm) and all ^{13}C NMR was calibrated with $CDCl_3$ (*δ* = 77.00 ppm). 2-D NMR experiments such as HMBC, COSY, TOCSY, HSQC, HSQC-TOCSY were recorded on the Bruker AVANCE 600 (600 MHz) spectrometer at 25 °C.

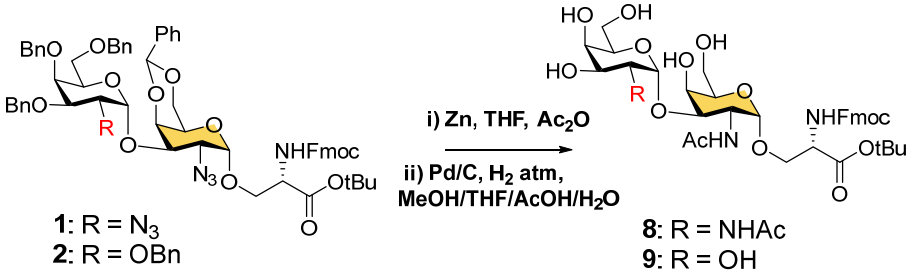
Synthesis of Compound 1 and 2:



To a solution of glycosyl acceptor **4**[4] (1.2 mmol) and AgOTf (2 mmol) in 10 mL of anhydrous diethyl ether was added 4 Å molecular sieves power and stirred at RT over 1h under argon atmosphere. The solution was cooled to -78 °C using a dry ice bath and freshly prepared *p*TolSCl (1.2 mmol) was added to maintain the anhydrous conditions. After 10 min, the respective glycosyl donor (**5**/**6,** 1 mmol) in 2 mL of anhydrous

diethyl ether was slowly added over 10 min. The reaction mixture was allowed to reach RT over 2 h and quenched with DIPEA (2 mmol). The reaction mixture was filtered over celite 545 using diethyl ether. The organic layer was washed with aq. $NaHCO_3$ followed by brine and dried over anhydrous sodium sulfate. The organic layer was filtered and concentrated under reduced pressure and purified using Biotage flash chromatography system using hexanes and ethyl acetate to obtain the respective disaccharide **1** (80% yield) and **2** (86% yield). **Compound 1:** ^{1}H NMR (600 MHz, Chloroform-*d*) δ 7.80 (d, *J* = 7.6 Hz, 2H), 7.63 (dd, *J* = 15.0, 7.4 Hz, 4H), 7.48 – 7.26 (m, 27H), 5.90 (d, *J* = 7.7 Hz, 1H), 5.60 (s, 1H), 5.25 (d, *J* = 3.7 Hz, 1H), 5.05 (d, *J* = 3.5 Hz, 1H), 4.92 (dd, *J* = 11.4, 3.1 Hz, 1H), 4.79 – 4.68 (m, 2H), 4.56 (dd, *J* = 13.4, 10.2 Hz, 2H), 4.51 – 4.36 (m, 6H), 4.28 – 4.19 (m, 4H), 4.19 – 4.10 (m, 4H), 4.01 – 3.90 (m, 4H), 3.74 – 3.67 (m, 2H), 3.63 (dd, *J* = 9.1, 5.4 Hz, 1H), 1.53 (d, *J* = 3.1 Hz, 9H). ^{13}C NMR (151 MHz, $CDCl_3$) δ 168.67, 155.77, 143.65, 141.19, 138.19, 137.70, 137.41, 137.26, 128.76, 128.40, 128.33, 128.20, 128.03, 127.86, 127.82, 127.77, 127.69, 127.62, 127.05, 127.00, 125.83, 125.03, 124.94, 119.97, 100.47, 99.45, 94.52, 82.81, 77.21, 77.00, 76.79, 76.61, 74.87, 73.36, 73.25, 72.24, 71.09, 70.60, 70.08, 69.65, 69.01, 68.06, 67.07, 62.98, 58.95, 58.56, 54.88, 46.95, 27.84. ESI-MS calcd for **1** $C_{62}H_{65}N_7O_{13}Na$ $[M + Na]^+$ *m/z* = 1138.4538, found: 1138.4520. **Compound 2:** ^{1}H NMR (600 MHz, Chloroform-*d*) δ 7.83 (d, *J* = 7.6 Hz, 2H), 7.67 (d, *J* = 7.5 Hz, 2H), 7.60 – 7.55 (m, 2H), 7.45 (td, *J* = 7.6, 2.6 Hz, 2H), 7.41 – 7.12 (m, 26H), 5.92 (d, *J* = 7.9 Hz, 1H), 5.49 (s, 1H), 5.30 (d, *J* = 3.7 Hz, 1H), 5.10 (d, *J* = 3.6 Hz, 1H), 5.01 (d, *J* = 10.9 Hz, 1H), 4.91 (d, *J* = 11.0 Hz, 1H), 4.86 (d, *J* = 10.8 Hz, 1H), 4.68 – 4.61 (m, 3H), 4.55 (d, *J* = 11.0 Hz, 1H), 4.52 – 4.44 (m, 4H), 4.41 (dd, *J* = 10.6, 7.2 Hz, 1H), 4.31 – 4.24 (m, 2H), 4.23 – 4.15 (m, 3H), 4.08 – 3.94 (m, 4H), 3.87 – 3.67 (m, 5H), 1.56 (s, 9H). ^{13}C NMR (151 MHz, $CDCl_3$) δ 168.71, 155.81, 143.70, 141.24, 141.20, 138.74, 138.40, 138.04, 137.79, 137.33, 128.98, 128.26, 128.24, 128.19, 128.13, 127.90, 127.85, 127.78, 127.70, 127.64, 127.61, 127.60, 127.46, 127.43, 127.09, 127.03, 126.12, 125.07, 125.02, 119.97, 100.85, 99.80, 93.32, 82.82, 81.69, 79.08, 77.21, 77.13, 77.00, 76.94, 76.79, 75.58, 74.77, 73.40, 72.15, 71.51, 71.11, 70.32, 69.98, 69.20, 68.11, 67.15, 63.08, 58.40, 55.00, 47.01, 27.89. ESI-MS calcd for **2** $C_{69}H_{72}N_7O_{14}Na$ $[M + Na]^+$ *m/z* = 1203.4943, found: 1203.4915.

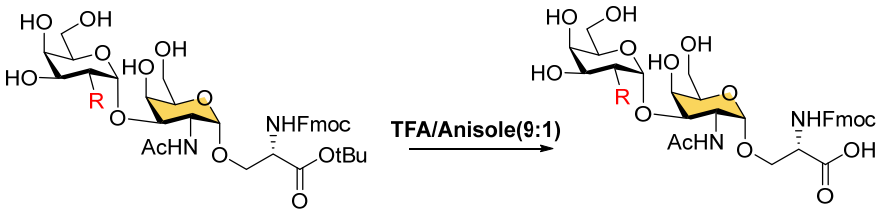
Synthesis of Compound 8 and 9:



To a solution of disaccharide **1** or **2** (0.5 mmol) in a mixture of THF:AcOH:Ac_2O (8.5:3.8:1.3 mL) was added activated solid Zn (2.7 g) and stirred at RT until completion. The reaction mixture was filtered over

celite 545 and concentrated to dryness under reduced pressure. The crude was dissolved in EtOAc and washed with aq. NaHCO3, brine and dried over anhydrous sodium sulfate. The organic layer was filtered and concentrated to dryness under reduced pressure. The crude was dissolved in THF:MeOH:H_2O:AcOH (5:2.5:2:0.5 mL) and 200 mg of 10% pd/c was added. The reaction mixture was stirred under H_2 atm over 24 hours for completion. The reaction mixture was filtered over celite 545 and concentrated under reduced pressure. The crude was purified using C18 reverse phase chromatography using Biotage flash chromatography system using water and acetonitrile to obtain compound **8** (74% yield) and compound **9** (72% yield) respectively from compound **1** and **2.** Compound **8:** ^{1}H NMR (600 MHz, Methanol-d_4) δ 7.75 – 7.71 (m, 2H), 7.61 (d, *J* = 7.4 Hz, 2H), 7.35 – 7.30 (m, 2H), 7.28 – 7.22 (m, 2H), 4.89 (d, *J* = 3.8 Hz, 1H), 4.73 (d, *J* = 3.9 Hz, 1H), 4.41 – 4.32 (m, 3H), 4.27 – 4.21 (m, 2H), 4.17 (t, *J* = 6.7 Hz, 1H), 3.94 (d, *J* = 3.1 Hz, 1H), 3.82 (ddt, *J* = 8.8, 4.1, 2.1 Hz, 2H), 3.77 (ddd, *J* = 18.1, 6.3, 3.6 Hz, 3H), 3.72 (d, *J* = 7.5 Hz, 1H), 3.68 – 3.62 (m, 4H), 3.59 (dd, *J* = 11.0, 3.1 Hz, 1H), 1.93 (s, 3H), 1.91 (s, 3H), 1.39 (s, 9H). ^{13}C NMR (101 MHz, Methanol-d_4)) δ 172.55, 171.99, 170.00, 144.58, 141.67, 128.15, 127.57, 125.51, 120.39, 117.70, 99.20, 94.98, 82.59, 73.85, 71.76, 71.37, 69.41, 69.21, 68.47, 66.97, 65.84, 62.20, 61.99, 55.60, 50.29, 47.58, 27.63, 22.65, 22.40. ESI-MS calcd for **8** $C_{38}H_{51}N_3O_{15}Na$ [M + Na]$^+$ *m/z* = 812.3218, found: 812.3195. **Compound 9**: ^{1}H NMR (600 MHz, Methanol-d_4) δ 7.68 (dd, *J* = 7.6, 2.3 Hz, 2H), 7.56 (d, *J* = 7.6 Hz, 2H), 7.28 (t, *J* = 7.5 Hz, 2H), 7.20 (t, *J* = 7.4 Hz, 2H), 4.84 (d, *J* = 3.9 Hz, 1H), 4.69 (d, *J* = 3.8 Hz, 1H), 4.32 (tdd, *J* = 15.8, 6.8, 3.8 Hz, 3H), 4.21 (t, *J* = 3.9 Hz, 1H), 4.15 – 4.10 (m, 1H), 3.97 (d, *J* = 3.0 Hz, 1H), 3.81 – 3.73 (m, 4H), 3.68 – 3.59 (m, 3H), 3.56 – 3.51 (m, 2H), 3.44 (t, *J* = 9.3 Hz, 1H), 3.27 (dd, *J* = 9.8, 3.8 Hz, 1H), 3.19 (p, *J* = 1.7 Hz, 2H), 3.09 (t, *J* = 9.1 Hz, 1H), 1.85 (s, 3H), 1.34 (s, 9H). ^{13}C NMR (151 MHz, MeOD) δ 172.17, 169.54, 157.27, 143.87, 143.82, 141.24, 127.45, 126.84, 124.77, 119.60, 98.63, 95.84, 82.00, 74.06, 73.57, 72.88, 71.91, 70.98, 70.37, 67.47, 66.59, 64.99, 61.62, 61.27, 55.23, 48.06, 48.05, 47.99, 47.92, 47.91, 47.86, 47.78, 47.76, 47.63, 47.62, 47.49, 47.48, 47.35, 47.34, 47.21, 47.19, 47.03, 26.93, 21.82. ESI-MS calcd for **9** $C_{36}H_{48}N_2O_{15}Na$ [M + Na]$^+$ *m/z* = 771.2952, found: 771.2900.

Synthesis of Compound 10 and 11:



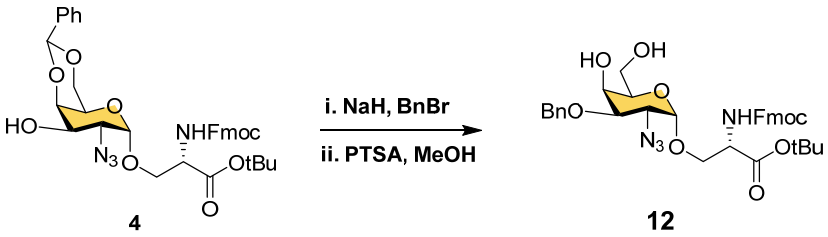
8: R = NHAc
9: R = OH

10: R = NHAc, 98%, **core 5**
11: R = OH, 95%, **core 8**

A solution of compound **8** (0.5 mmol**)** or **9** (0.5 mmol) respectively in 10 mL of TFA:Anisole (9:1) was stirred at RT for 10 min. The reaction mixture was concentrated under reduced pressure at 30 °C and

was purified using C18 reverse phase chromatography using Biotage flash chromatography system using water and acetonitrile to obtain compound **10** (98% yield) and compound **11** (95% yield) respectively from compound **8** and **9** respectively. Compound **10**: ^{1}H NMR (600 MHz, Methanol-d_4) δ 7.83 (d, *J* = 7.6 Hz, 2H), 7.71 (dd, *J* = 7.6, 3.6 Hz, 2H), 7.42 (t, *J* = 7.5 Hz, 2H), 7.34 (t, *J* = 7.4 Hz, 2H), 4.99 (d, *J* = 3.7 Hz, 1H), 4.85 (d, *J* = 3.8 Hz, 1H), 4.50 – 4.39 (m, 4H), 4.33 (dd, *J* = 11.1, 3.8 Hz, 1H), 4.27 (t, *J* = 6.7 Hz, 1H), 4.03 (d, *J* = 3.1 Hz, 1H), 4.00 – 3.81 (m, 6H), 3.81 – 3.66 (m, 5H), 2.02 (s, 6H). ^{13}C NMR (151 MHz, MeOD) δ 172.71, 172.45, 172.03, 157.27, 143.86, 141.24, 127.44, 126.83, 124.80, 119.58, 98.75, 94.77, 73.59, 71.64, 71.16, 69.00, 68.47, 67.95, 66.65, 65.18, 61.59, 61.25, 54.41, 49.72, 48.04, 47.90, 47.76, 47.61, 47.47, 47.33, 47.19, 47.02, 21.69, 21.47. ESI-MS calcd for **10** $C_{34}H_{43}N_3O_{15}Na$ $[M + Na]^+$ *m/z* = 756.2592, found: 756.2580. Compound **11:** ^{1}H NMR (600 MHz, Methanol-d_4) δ 7.82 (dd, *J* = 7.8, 2.8 Hz, 2H), 7.70 (d, *J* = 7.4 Hz, 2H), 7.42 (t, *J* = 7.5 Hz, 2H), 7.34 (t, *J* = 7.4 Hz, 2H), 4.98 (d, *J* = 3.9 Hz, 1H), 4.85 (d, *J* = 3.8 Hz, 1H), 4.45 (ttd, *J* = 11.1, 7.9, 4.0 Hz, 4H), 4.26 (q, *J* = 6.0 Hz, 1H), 4.11 (d, *J* = 3.0 Hz, 1H), 3.98 – 3.87 (m, 4H), 3.86 – 3.72 (m, 3H), 3.71 – 3.64 (m, 2H), 3.58 (t, *J* = 9.3 Hz, 1H), 3.41 (ddd, *J* = 9.8, 4.0, 1.7 Hz, 1H), 3.33 (p, *J* = 1.6 Hz, 2H), 3.22 (td, *J* = 9.3, 2.2 Hz, 1H), 2.01 (s, 3H). ^{13}C NMR (151 MHz, MeOD) δ 172.35, 172.00, 157.27, 143.85, 141.22, 127.44, 126.84, 124.80, 119.58, 98.73, 95.93, 74.08, 73.58, 72.87, 71.91, 70.93, 70.37, 67.88, 66.65, 65.08, 61.60, 61.27, 54.43, 48.05, 47.91, 47.77, 47.62, 47.48, 47.34, 47.20, 47.00, 21.70. ESI-MS calcd for **11** $C_{32}H_{40}N_2O_{15}Na$ $[M + Na]^+$ *m/z* = 715.2326, found: 715.2350.

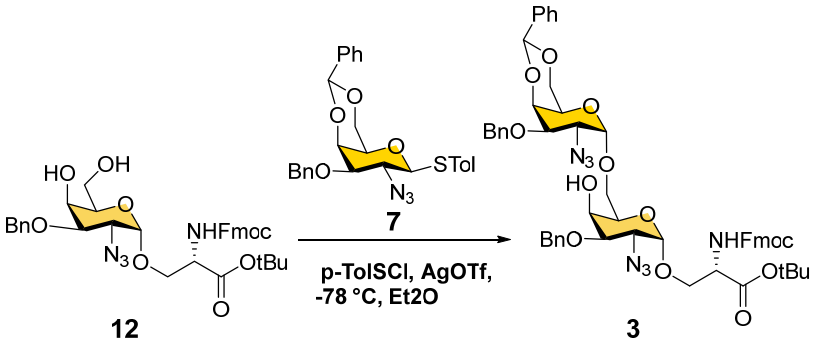
Synthesis of Compound 12:



To a solution of compound **4** (1g, 1.5 mmol) in anhydrous DMF was added NaH (1.5 mmol) at 0 °C followed by BnBr (2.0 mmol) and stirred at RT over 12h. The reaction mixture was diluted with diethyl ether and washed with aq. 1N HCl, aq. $NaHCO_3$, and brine. The organic layer was dried over anhydrous sodium sulfate, filtered, and concentrated under reduced pressure. The crude was dissolved in 20 mL of anhydrous MeOH and PTSA (0.3 mmol) was added and stirred until completion. The reaction was quenched with DIPEA (0.3 mmol) and concentrated to dryness and purified using Biotage flash chromatography system using hexanes and ethyl acetate to obtain compound **12** (0.72 g, 72% yield). ^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.79 (d, *J* = 7.5 Hz, 2H), 7.65 (dd, *J* = 7.5, 2.8 Hz, 2H), 7.45 – 7.31 (m, 12H), 6.08 (d, *J* = 8.2 Hz, 1H), 4.88 (d, *J* = 3.6 Hz, 1H), 4.70 (d, *J* = 3.3 Hz, 2H), 4.45 (d, *J* = 6.9 Hz, 2H), 4.25 (t, *J* = 6.8 Hz, 1H), 4.16 (dd, *J* = 11.3, 3.4 Hz, 1H), 4.09 (d, *J* = 3.2 Hz, 1H), 3.94 – 3.86 (m, 2H),

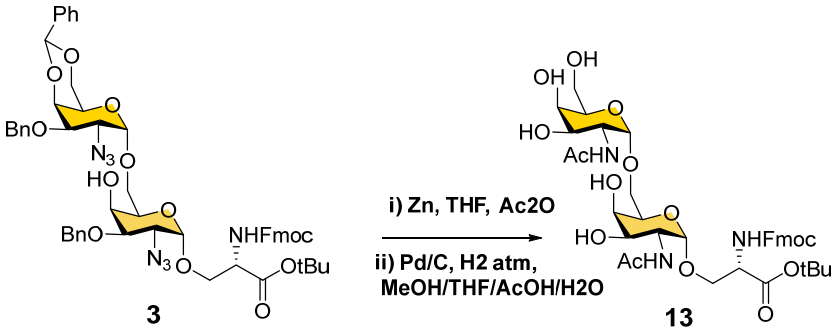
3.86 – 3.73 (m, 5H), 3.70 (dd, *J* = 10.4, 3.6 Hz, 2H), 1.50 (s, 9H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 169.03, 155.98, 143.85, 141.34, 136.99, 128.73, 128.36, 128.07, 127.75, 127.13, 125.16, 120.04, 99.73, 82.97, 77.41, 77.09, 76.78, 75.54, 71.85, 70.67, 70.32, 67.15, 66.96, 62.83, 59.00, 55.13, 47.17, 27.95. ESI-MS calcd for **12** $C_{35}H_{38}N_4O_9Na$ $[M + Na]^+$ *m/z* = 681.2536, found: 681.2400.

Synthesis of Compound 3:



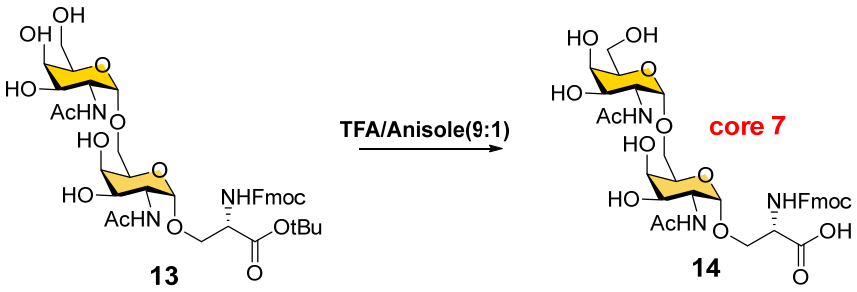
Compound **3** was obtained in 75% yield by following the procedures employed for the synthesis of compound **1.** ^{1}H NMR (600 MHz, Chloroform-*d*) δ 7.81 (d, *J* = 7.6 Hz, 2H), 7.67 (t, *J* = 6.7 Hz, 2H), 7.54 (dd, *J* = 7.5, 2.1 Hz, 2H), 7.47 – 7.32 (m, 18H), 5.84 (d, *J* = 8.2 Hz, 1H), 5.46 (s, 1H), 5.09 (d, *J* = 3.3 Hz, 1H), 4.92 (d, *J* = 3.6 Hz, 1H), 4.78 – 4.67 (m, 4H), 4.51 – 4.41 (m, 3H), 4.29 (t, *J* = 7.4 Hz, 1H), 4.23 – 4.19 (m, 2H), 4.07 – 3.89 (m, 8H), 3.77 (ddd, *J* = 21.1, 10.2, 4.5 Hz, 2H), 3.68 (s, 1H), 1.55 (s, 9H). ^{13}C NMR (151 MHz, $CDCl_3$) δ 168.76, 155.78, 143.75, 143.68, 141.21, 137.85, 137.52, 136.94, 128.87, 128.62, 128.58, 128.32, 128.27, 128.24, 128.08, 128.03, 127.96, 127.76, 127.73, 127.62, 127.05, 126.09, 126.07, 125.10, 119.98, 100.75, 99.15, 98.67, 82.85, 77.21, 77.14, 77.00, 76.94, 76.79, 75.47, 74.32, 72.78, 71.74, 70.95, 69.18, 69.10, 68.87, 67.15, 66.74, 65.83, 62.82, 58.89, 58.66, 54.84, 47.08, 27.87. ESI-MS calcd for **3** $C_{55}H_{59}N_7O_{13}Na$ $[M + Na]^+$ *m/z* = 1048.4069, found: 1048.4025.

Synthesis of Compound 13:



Compound **13** was obtained in 70% yield from compound **3** following the same procedure employed for compound **8**. ^{1}H NMR (600 MHz, Methanol-d_4) δ 7.83 (d, *J* = 7.5 Hz, 2H), 7.69 (t, *J* = 7.0 Hz, 2H), 7.42 (t, *J* = 7.5 Hz, 2H), 7.35 (dt, *J* = 8.3, 4.1 Hz, 2H), 4.84 (t, *J* = 3.7 Hz, 2H), 4.46 (dd, *J* = 6.6, 3.7 Hz, 2H), 4.39 (t, *J* = 4.5 Hz, 1H), 4.34 – 4.26 (m, 3H), 3.97 (ddd, *J* = 7.5, 4.2, 1.2 Hz, 1H), 3.95 – 3.89 (m, 4H), 3.88 – 3.83 (m, 2H), 3.80 (dd, *J* = 10.9, 3.1 Hz, 1H), 3.78 – 3.75 (m, 1H), 3.72 (dd, *J* = 6.1, 3.5 Hz, 2H), 3.55 (dd, *J* = 10.3, 4.2 Hz, 1H), 2.00 (s, 3H), 1.98 (s, 3H), 1.49 (s, 9H). ^{13}C NMR (151 MHz, MeOD) δ 172.66, 172.40, 169.70, 157.03, 143.84, 141.25, 127.49, 126.85, 124.72, 119.63, 98.92, 97.43, 82.20, 70.92, 69.45, 69.08, 68.91, 68.62, 68.15, 67.90, 66.66, 61.27, 55.41, 50.13, 49.95, 48.02, 47.88, 47.74, 47.60, 47.46, 47.31, 47.17, 46.99, 26.93, 21.51. ESI-MS calcd for **13** $C_{38}H_{51}N_3O_{15}Na$ $[M + Na]^+$ *m/z* = 812.3218, found: 812.3215.

Synthesis of Compound 14



Compound **14** was obtained in 92% yield from compound **13** following the same procedure employed for compound **10.** ^{1}H NMR (600 MHz, Methanol-d4) δ 7.82 (d, J = 7.5 Hz, 2H), 7.69 (dd, J = 7.4, 4.8 Hz, 2H), 7.42 (t, J = 7.5 Hz, 2H), 7.34 (t, J = 7.4 Hz, 2H), 4.84 (d, J = 3.2 Hz, 1H), 4.44 (dtd, J = 17.3, 11.7, 10.6, 6.2 Hz, 2H), 4.29 (dt, J = 18.5, 6.8 Hz, 3H), 4.03 (dd, J = 8.0, 3.8 Hz, 1H), 3.99 – 3.91 (m, 4H), 3.87 (d, J = 3.1 Hz, 2H), 3.83 – 3.73 (m, 3H), 3.55 (dd, J = 10.5, 3.6 Hz, 1H), 2.03 (s, 3H), 2.00 (s, 3H). ^{13}C NMR (151 MHz, CD3OD) δ 143.81, 143.78, 141.23, 127.47, 126.83, 124.73, 119.61, 99.03, 97.52, 70.32, 69.52, 69.47, 69.23, 68.71, 68.20, 66.77, 61.63, 50.08, 50.06, 48.05, 47.91, 47.77, 47.62, 47.48, 47.34, 47.28, 47.20, 47.03, 21.53, 21.42. ESI-MS calcd for 14 C34H43N3O15Na [M + Na]+ m/z 756.2592, found: 756.2590.

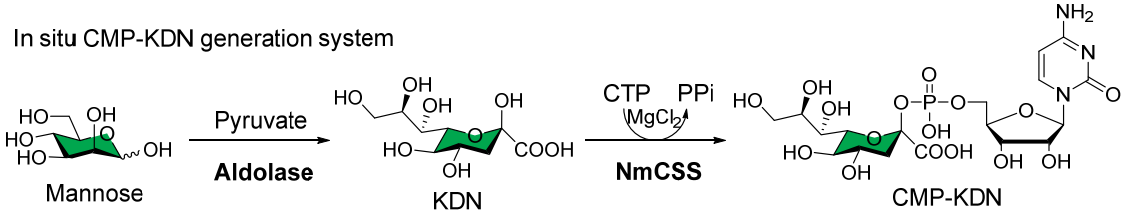
III. Enzymatic assays and synthetic procedures

Activity assays of Pd26ST towards rare core acceptors using CMP-Neu5Ac/Gc donors

A reaction system of 20 μL contains Tris-HCl buffer (100 mM, pH 8.0), one of Fmoc-protected core acceptor (1, 5, or 10 mM), CMP-Neu5Ac or CMP-Neu5Gc (2 equivalents of acceptor), $MgCl_2$ (10 mM), and 80 μg of purified Pd26ST. The reaction was incubated at 37 °C for 1 h (or overnight), quenched by the addition of 1 equivalent of ice-cold ethanol, centrifuged and monitored by HPLC and/or MALDI-TOF or ESI-mass spectrometry (MS). Conversion rates were calculated according to integrated areas of the acceptor and sialyl-core product peak.

Activity assays of Pd26ST towards rare core acceptors using a CMP-KDN in situ generation system

A reaction system of 20 μL contains Tris-HCl buffer (100 mM, pH 8.0), one of Fmoc-protected core acceptor (1, 5, or 10 mM), 2 equivalents of CTP, 2 equivalents of pyruvate, $MgCl_2$ (10 mM), excess amounts NmCSS[1] and *E. coli* Aldolase,[3] and 80 μg of purified Pd26ST. The reaction was incubated at 37 °C for 1 h (or overnight), quenched by addition of 1 equivalent of ice-cold ethanol, centrifuged and monitored by HPLC and/or MALDI-TOF or ESI-mass spectrometry (MS). Conversion rate was calculated according to integrated areas of the acceptor and sialyl-core product peak.



Synthetic scale synthesis sialyl-rare cores using Pd26ST

Reaction systems contain Tris-HCl buffer (100 mM, pH 8.0), one of Fmoc-protected core acceptor (20 mM), 2 equivalents of CMP-Neu5Ac/Neu5Gc or the CMP-KDN in situ generation system, $MgCl_2$ (20 mM), and excess amounts of purified Pd26ST. The reactions were incubated at 30 °C for up to 4 hours, and purified by semi-preparative HPLC, followed with Bio-gel P2 chromatography (50 mM NH_4HCO_3 as running buffer) to remove residue trifluoroacetic acid (TFA) from HPLC elution. TFA could cause significant amount of desialylation during NMR characterization if not removed.

Activity assays of human ST6GalNAcs towards rare core acceptors using CMP-Neu5Ac as the donor

A reaction system of 10 μL contains MES buffer (100 mM, pH 7.0), 0.1 mM of a Fmoc-protected core acceptor (core 1, 5, 7, or 8, or a glycopeptide with core 1), CMP-Neu5Ac (10 mM), and 2 μg of ST6GalNAcs. The reaction was incubated at 37 °C for 48 h, quenched and analyzed by HPLC as mentioned above.

Analytic HPLC method for analyzing reactions

HPLC was performed on a Shimadzu LC20D analytical HPLC system, using a GL Science Inertsil ODS-4 column (100 Å, 5 μm, 4.6 mm × 250 mm), and monitored by a UV detector (260 nm) and/or fluorescent detector (E_x 260nm, E_m 310 nm).[5] The running solvents were solvent A (0.1% TFA in ddH_2O) and solvent B (0.1% TFA in acetonitrile). The running condition was gradient elution with solvent B% linear increased from 20% to 40% within 25 mins, with a total flow rate of 1 mL/min.

Separation of synthesized sialyl-rare core separation

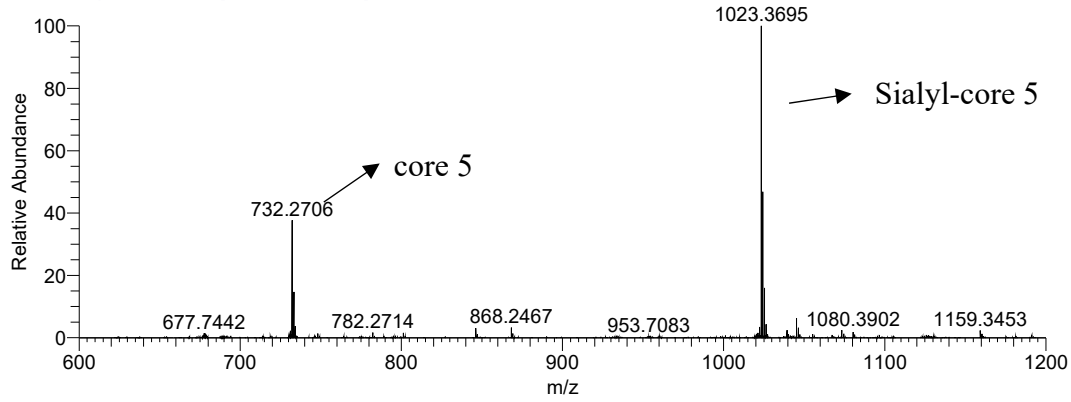
HPLC was performed on a Shimadzu LC20 analytical HPLC system, using a GL Science Inertsil ODS-4 column (100 Å, 5 μm, 10 mm × 250 mm), and monitored by a UV detector (260 nm). The running solvents were solvent A (0.1% TFA in ddH_2O) and solvent B (0.1% TFA in acetonitrile). The running condition was gradient elution with solvent B% linear increased from 20% to 40% within 25 mins, with a total flow rate of 4.5 mL/min. Product-containing fractions were concentrated and loaded onto a Bio-gel P2 column (2.5 x 120 cm) that was preequilibrated with 50 mM NH_4HCO_3 to remove residue TFA. Product-containing fractions were then concentrated and lyophilized twice for NMR characterization and long-term storage.

Analytic HPLC method for analyzing purified rare cores and sialylated rare cores to check the purity

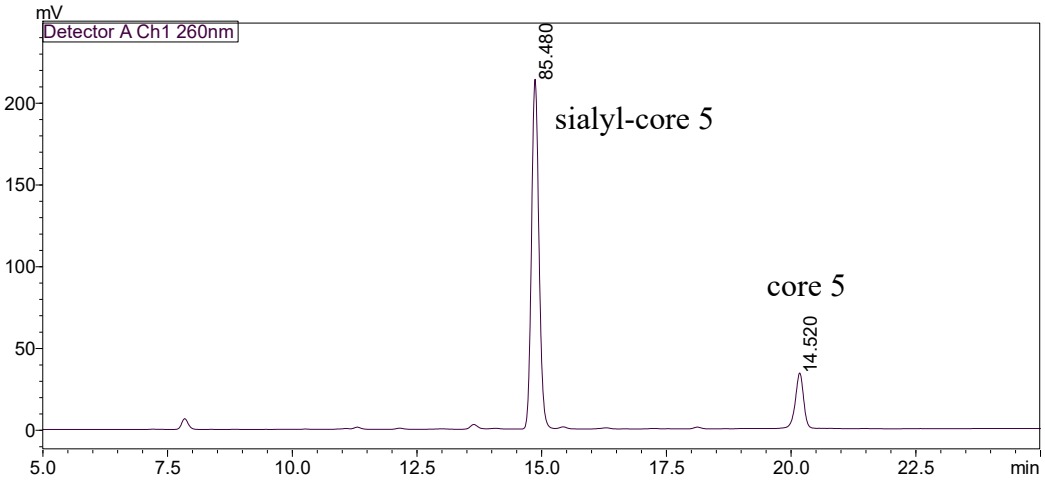
HPLC was performed on a Shimadzu Nexera analytical HPLC system, using a Waters BEH C18 column (100 Å, 5 μm, 4.6 mm × 250 mm), and monitored by a UV detector (260 nm) and/or fluorescent detector (E_x 260nm, E_m 310 nm).[5] The running solvents were solvent A (0.1% TFA in ddH_2O) and solvent B (0.1% TFA in acetonitrile). The running condition was gradient elution with solvent B% linear increased from 20% to 40% within 30 mins, with a total flow rate of 1 mL/min.

IV. HPLC and MS analysis of sialyltransferase-catalyzed reactions

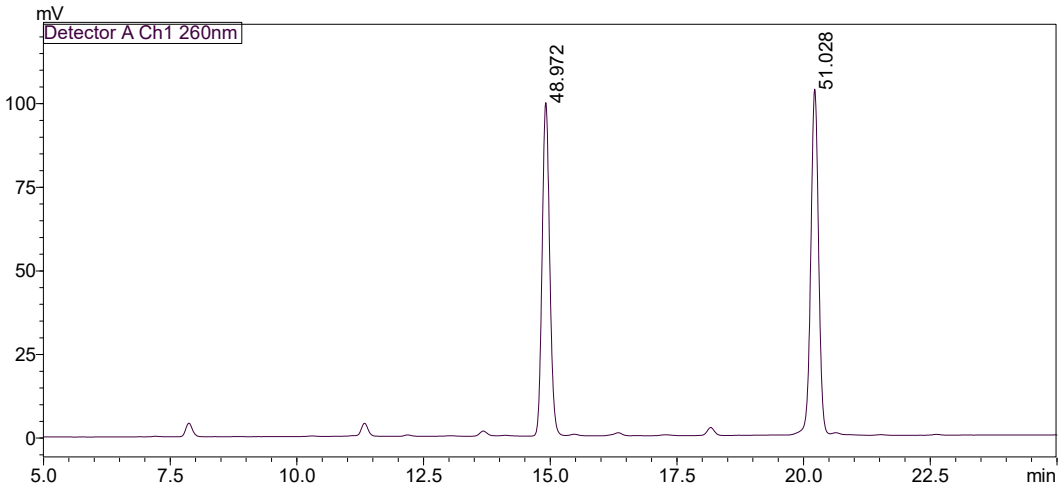
ESI-MS: Pd26ST-catazlyed reaction using CMP-Neu5Ac, 10m core 5, 1 h



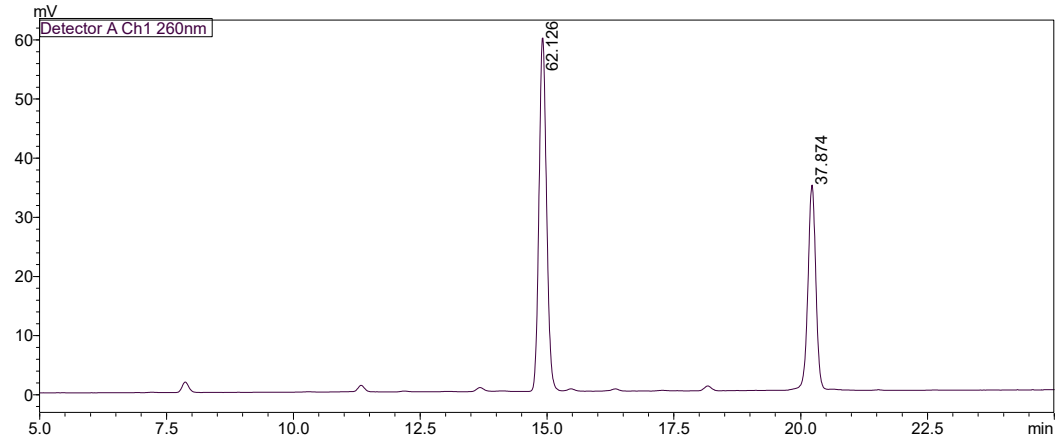
HPLC: Pd26ST-catazlyed reaction using CMP-Neu5Ac, 10m core 5, 1 h: core 5 (20.21 min), sialyl-core 5 (14.87 min), 85.5% conversion



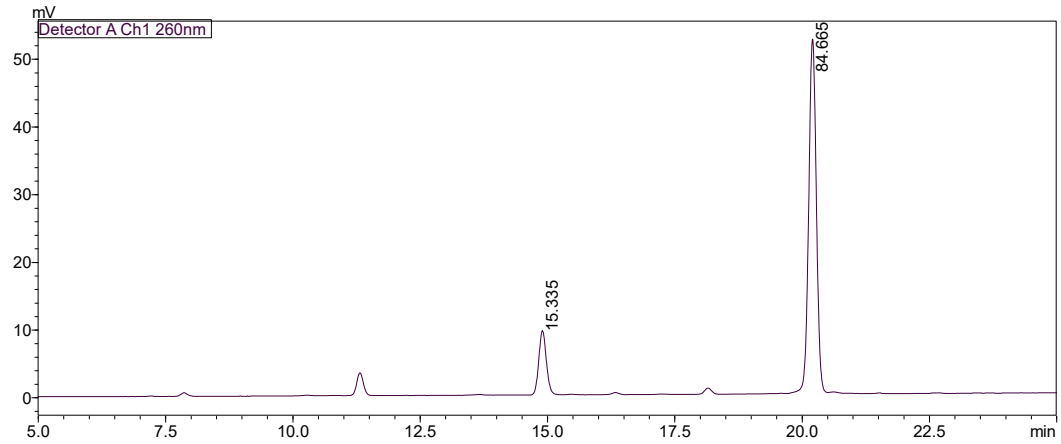
HPLC: Pd26ST-catazlyed reaction using CMP-Neu5Ac, 10 mM core 5, overnight: 49% conversion



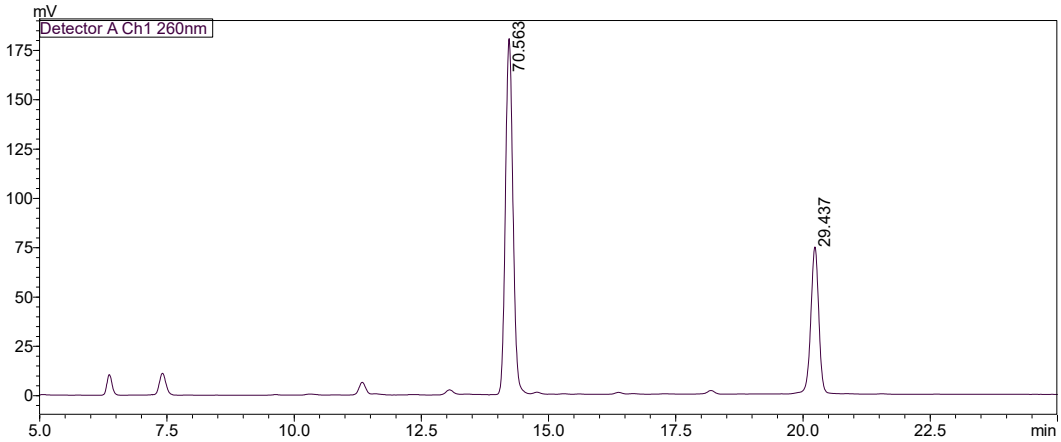
HPLC: Pd26ST-catazlyed reaction using CMP-Neu5Ac, 5 mM core 5, 1 h: 62.1% conversion



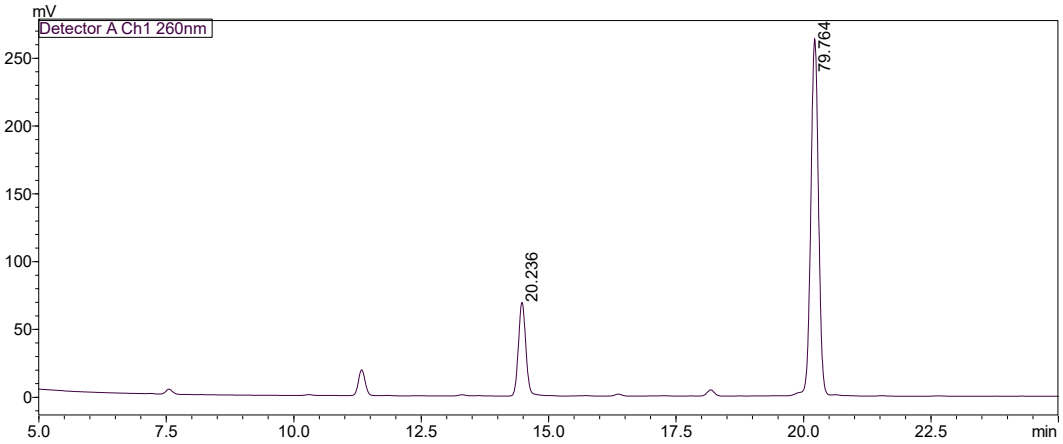
HPLC: Pd26ST-catazlyed reaction using CMP-Neu5Ac, 1 mM core 5, 1 h: 15.3% conversion



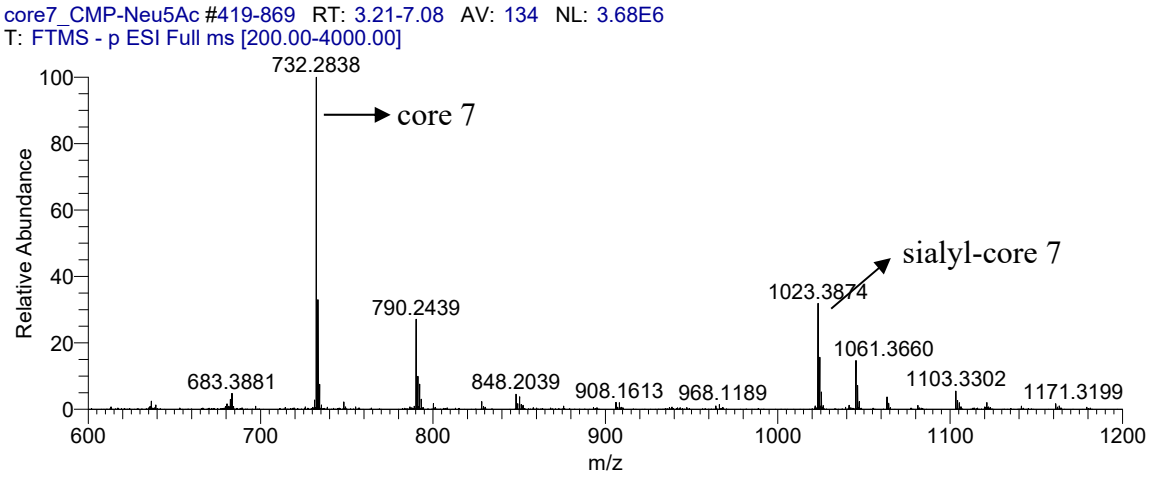
HPLC: Pd26ST-catazlyed reaction using CMP-Neu5Gc, 1 mM core 5, 1 h. core 5 (20.21 min), sialyl(neu5Gc)-core 5 (14.22 min), 70.6% conversion



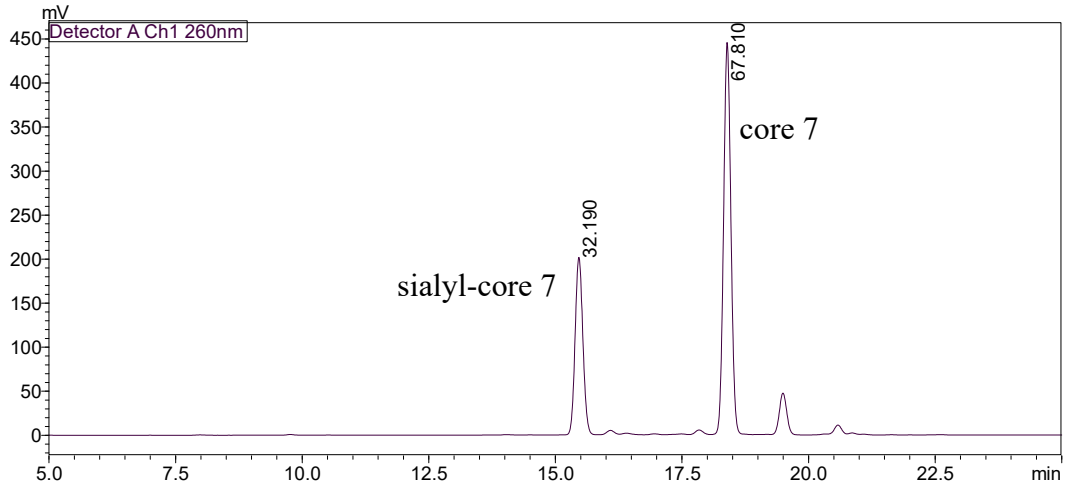
HPLC: Pd26ST-catazlyed reaction using CMP-KDN (one-pot system), 1 mM core 5, 1 h. core 5 (20.21 min), sialyl(KDN)-core 5 (14.47 min), 20.2% conversion



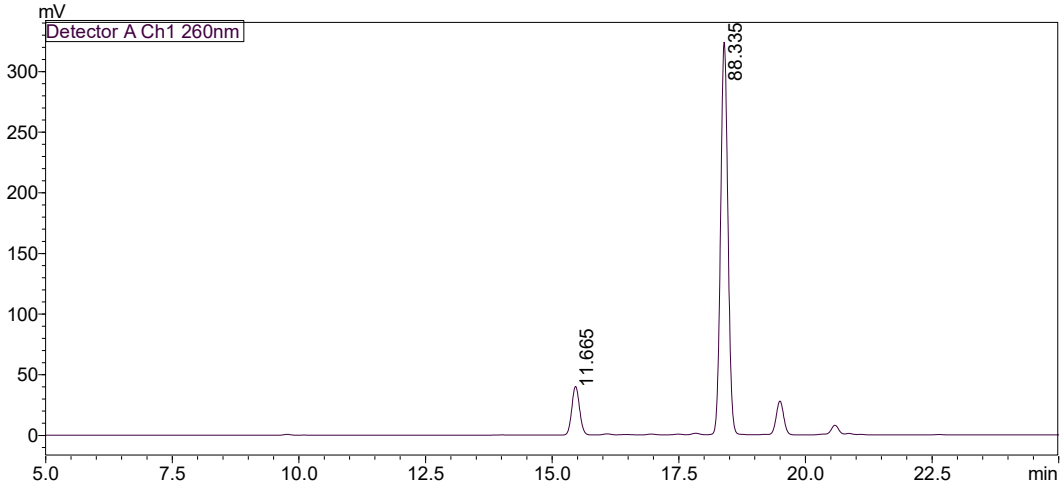
ESI-MS: Pd26ST-catazlyed reaction using CMP-Neu5Ac, 10m core 7, 1 h



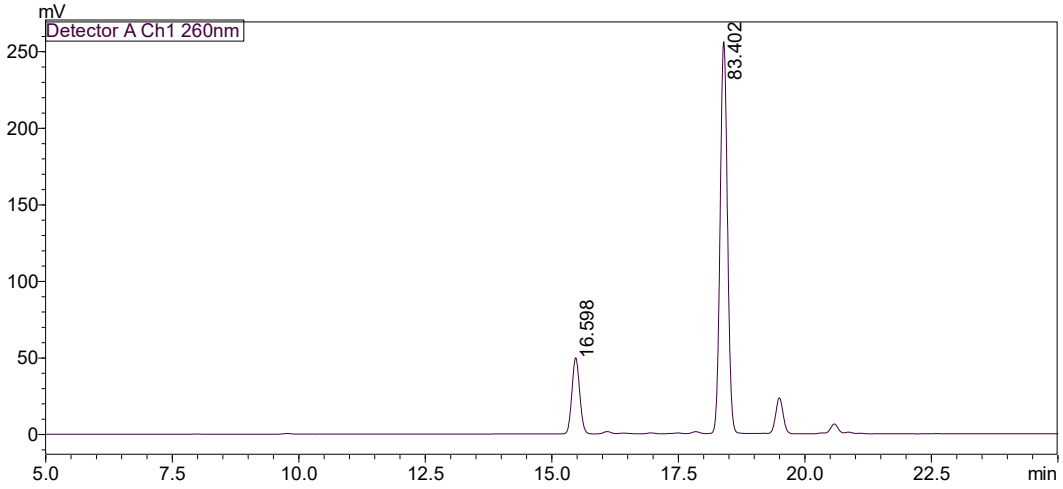
HPLC: Pd26ST-catazlyed reaction using CMP-Neu5Ac, 10m core 7, 1 h: core 7 (18.42 min), sialyl-core 7 (15.46 min), 32.2% conversion



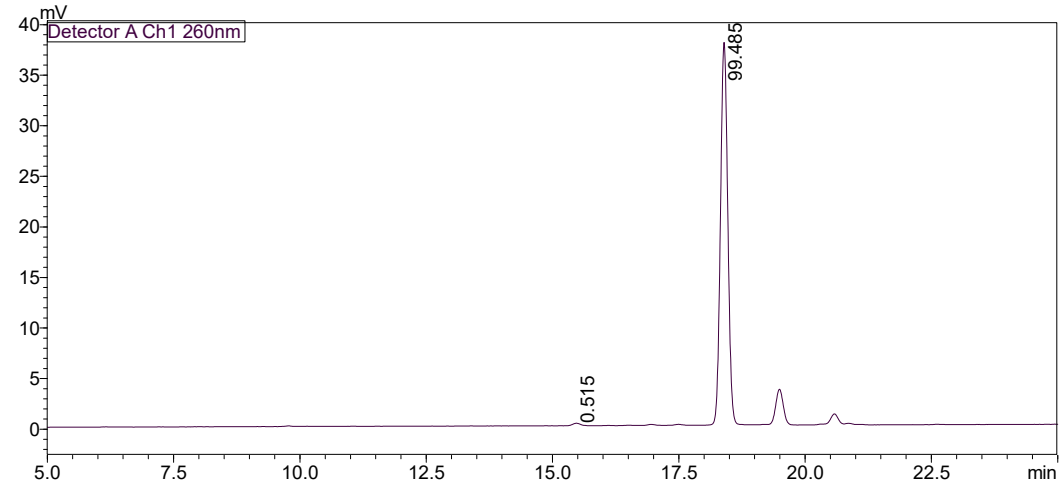
HPLC: Pd26ST-catazlyed reaction using CMP-Neu5Ac, 10 mM core 7, overnight:11.7% conversion



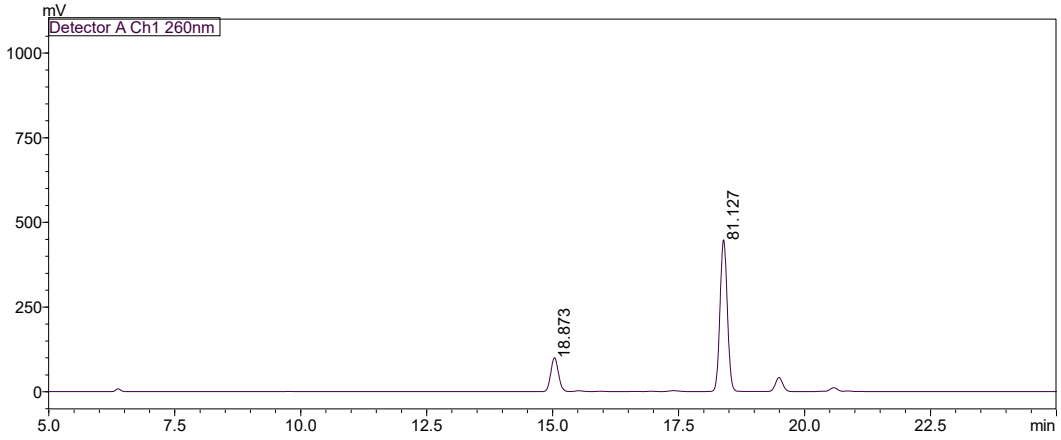
HPLC: Pd26ST-catazlyed reactions using CMP-Neu5Ac, 5 mM core 7, 1 h: 16.6% conversion



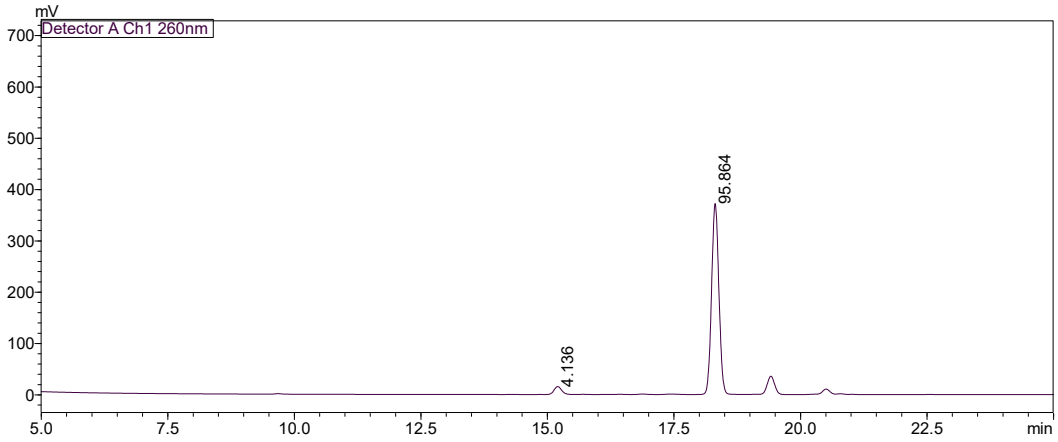
HPLC: Pd26ST-catazlyed reactions using CMP-Neu5Ac, 1 mM core 7, 1 h: <1% conversion



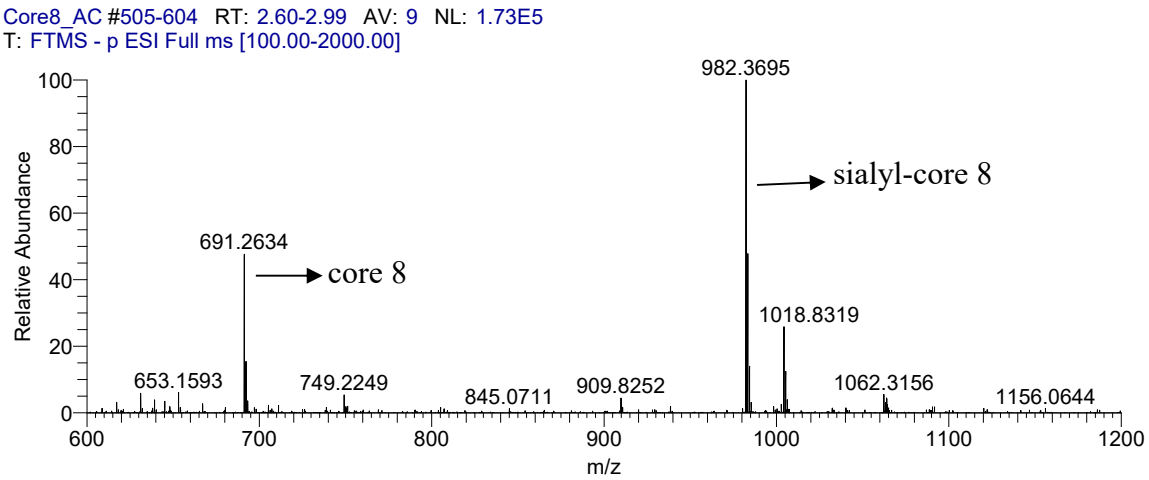
HPLC: Pd26ST-catazlyed reaction using CMP-Neu5Gc, 10m core 7, 1 h: core 7 (18.42 min), sialyl(Neu5Gc)-core 7 (15.09 min), 18.9% conversion



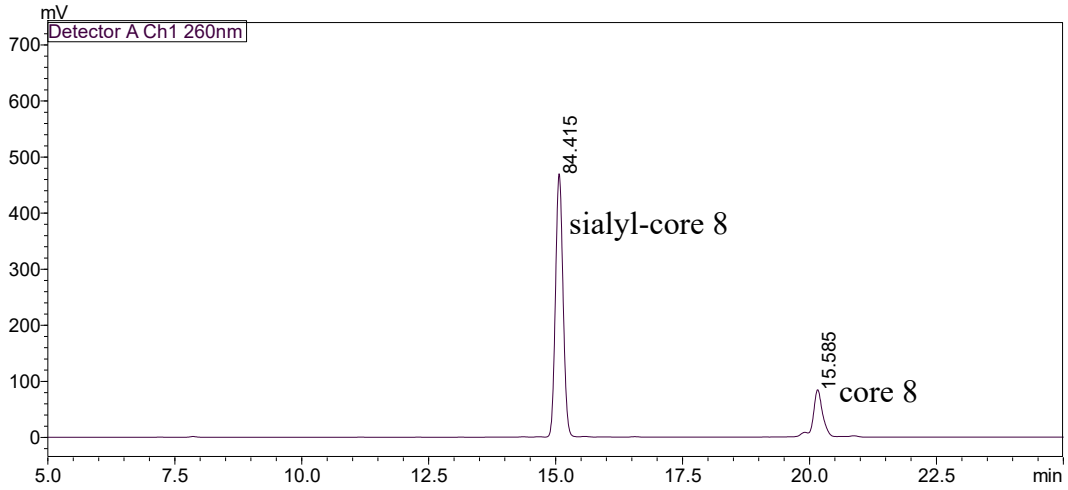
HPLC: Pd26ST-catazlyed reaction using CMP-KDN (one pot system), 10m core 7, 1 h: core 7 (18.42 min), sialyl(KDN)-core 7 (15.21 min), 4.1% conversion



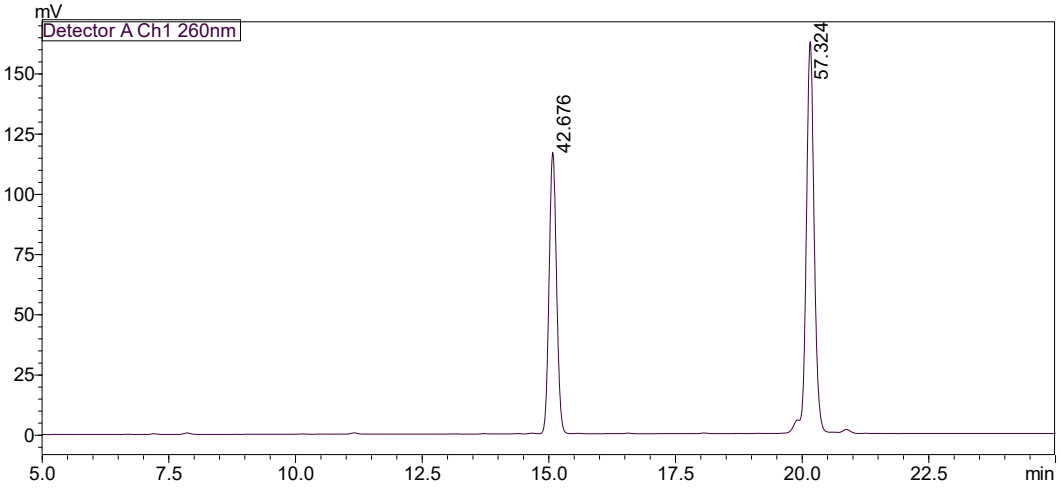
ESI-MS: Pd26ST-catazlyed reaction using CMP-Neu5Ac, 10m core 8, 1 h



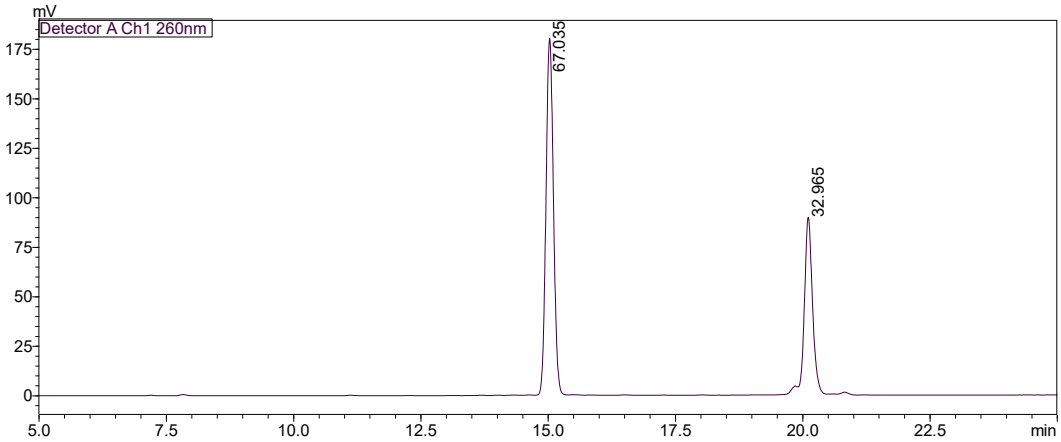
HPLC: Pd26ST-catazlyed reaction using CMP-Neu5Ac, 10m core 8, 1 h: core 8 (20.19 min), sialyl-core 8 (15.07 min), 84.4% conversion



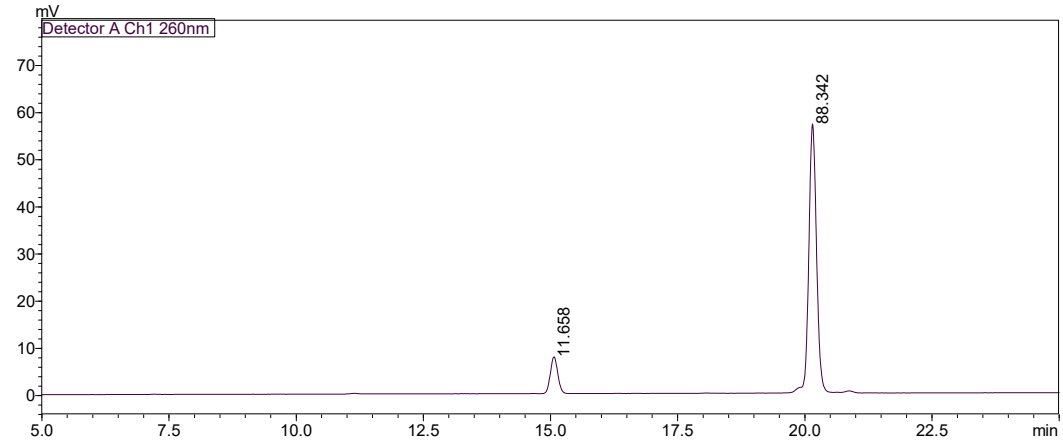
HPLC: Pd26ST-catazlyed reaction using CMP-Neu5Ac, 10 mM core 8, overnight:42.7% conversion



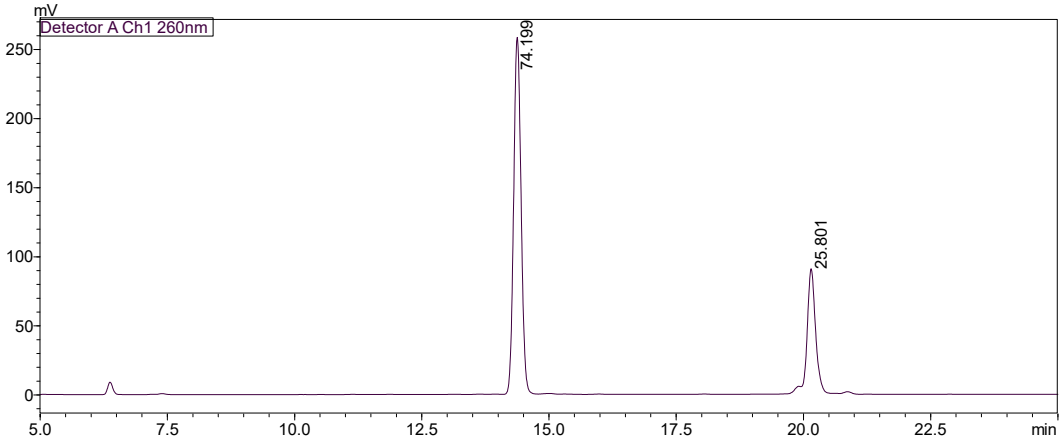
HPLC: Pd26ST-catazlyed reactions using CMP-Neu5Ac, 5 mM core 8, 1 h: 67.0% conversion



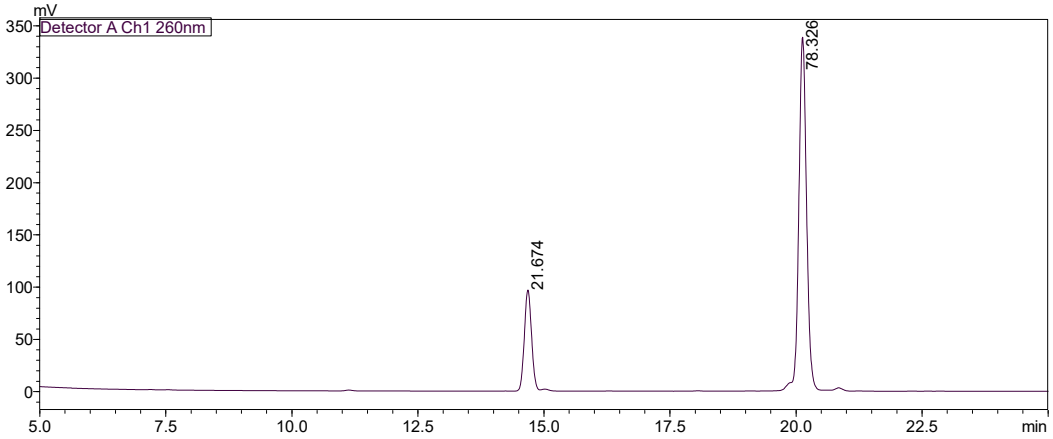
HPLC: Pd26ST-catazlyed reactions using CMP-Neu5Ac, 1 mM core 8, 1 h: 11.7% conversion



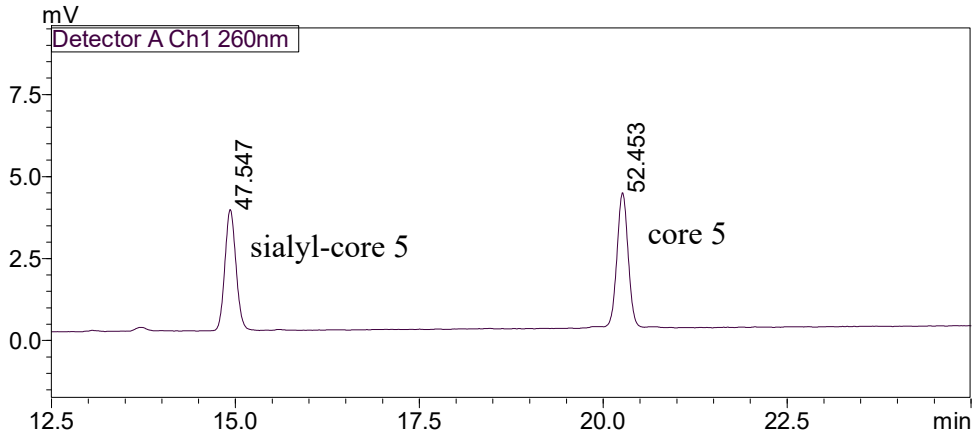
HPLC: Pd26ST-catazlyed reaction using CMP-Neu5Gc, 10m core 8, 1 h: core 8 (20.19 min), sialyl(Neu5Gc)-core 8 (14.38 min), 74.2% conversion



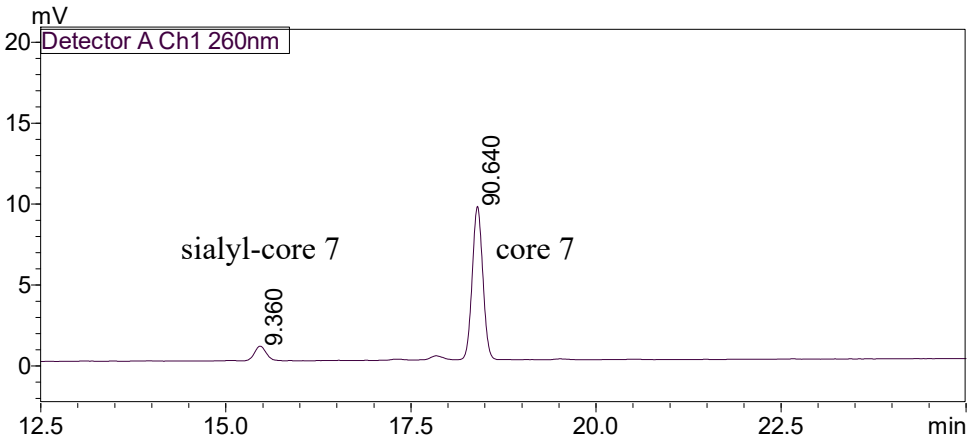
HPLC: Pd26ST-catazlyed reaction using CMP-KDN (one pot system), 10m core 8, 1 h: core 8 (20.19 min), sialyl(KDN)-core 8 (15.64 min), 21.7% conversion



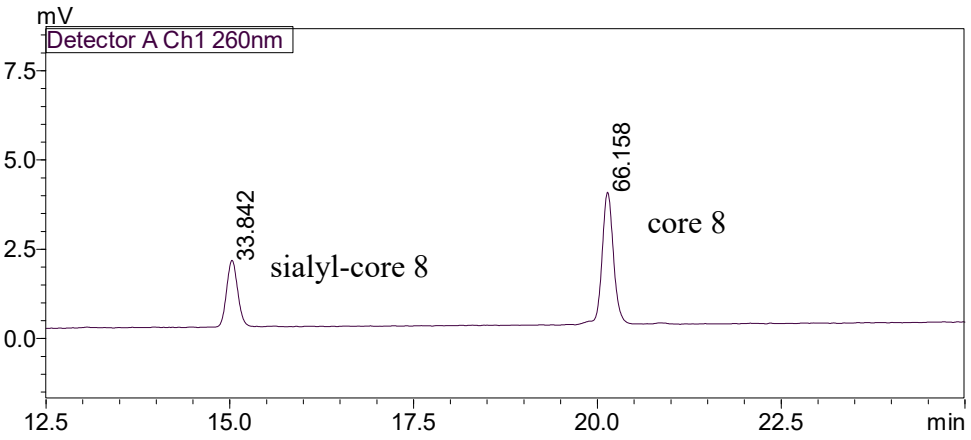
HPLC: ST6GalNAc1-catazlyed reaction using core 5 as acceptor, CMP-Neu5Ac as donor: 47.5% conversion



HPLC: ST6GalNAc1-catazlyed reaction using core 7 as acceptor, CMP-Neu5Ac as donor: 9.4% conversion



HPLC: ST6GalNAc1-catazlyed reaction using core 8 as acceptor, CMP-Neu5Ac as donor: 33.8% conversion



V. Supporting figures

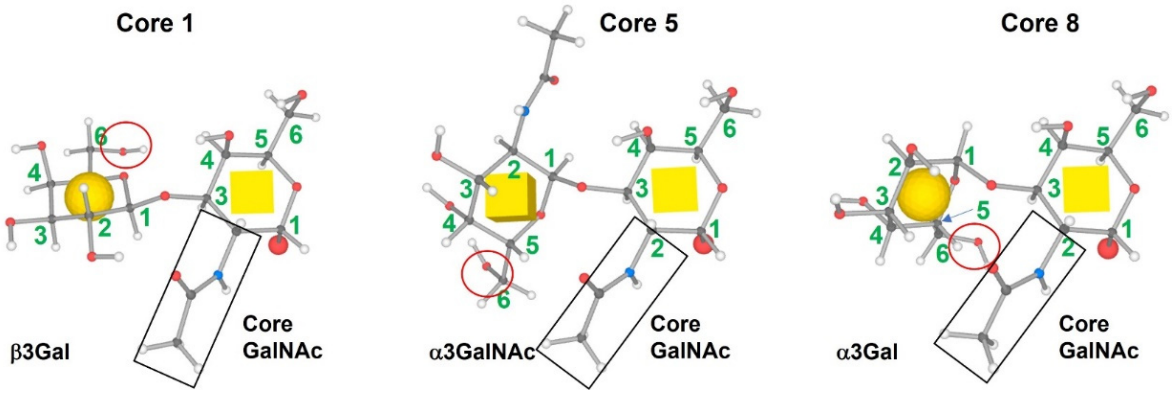


Figure S1. 3D structure of O-GalNAc core 1, core 5, and core 8 disaccharides modeled by GLYCAM (www.glycam.org). Black rectangle, C2-N-acetyl group of the core GalNAc; red oval, C6-OH group of distal Gal/GalNAc. In core 5 and core 8, C2-N-acetyl of the initial GalNAc residue are spatially close to C6-OH of distal α-linked GalNAc/Gal residues, thus may preventing Pd26ST to sialylate the C6-OH of distal GalNAc/Gal. C6-OH of the β3Gal on core 1 is far away from the C2-N-acetyl of the initial GalNAc, thus can be sialylated.

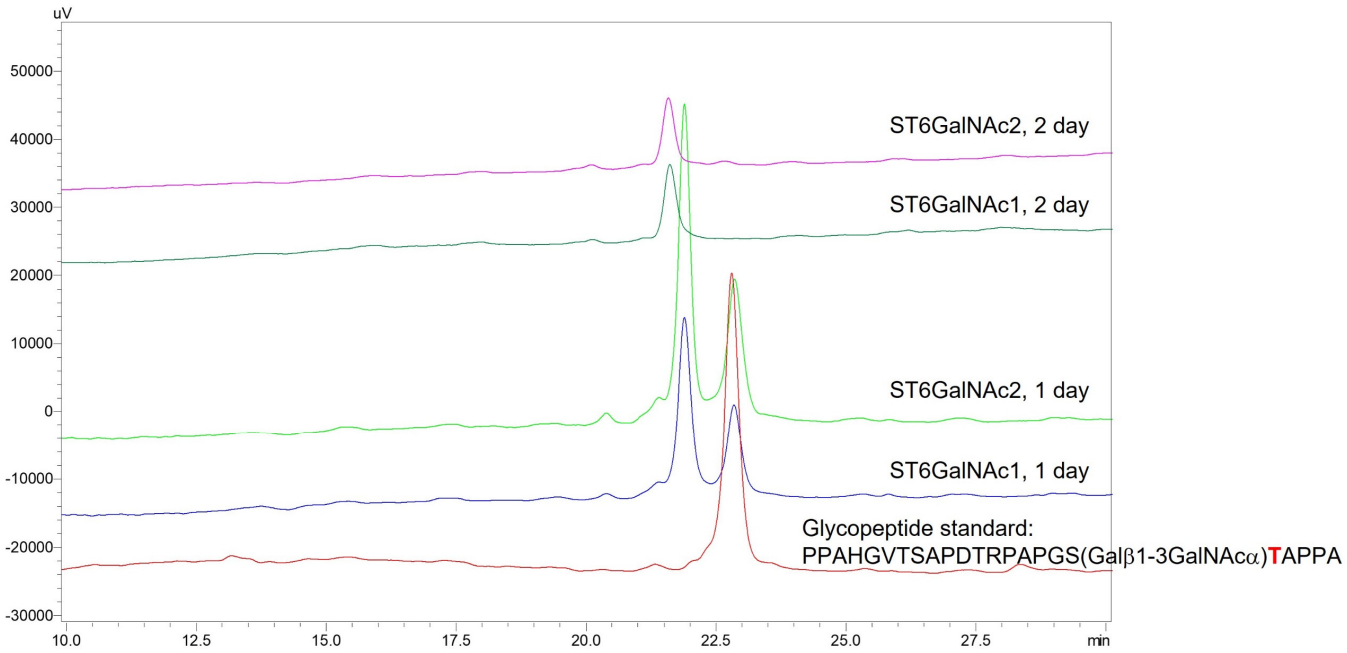


Figure S2. ST2GalNAc1 and ST2GalNAc2-catalyzed reaction using core1-Muc1 glycopeptides (the synthesis of the peptide is unpublished data) as acceptor.

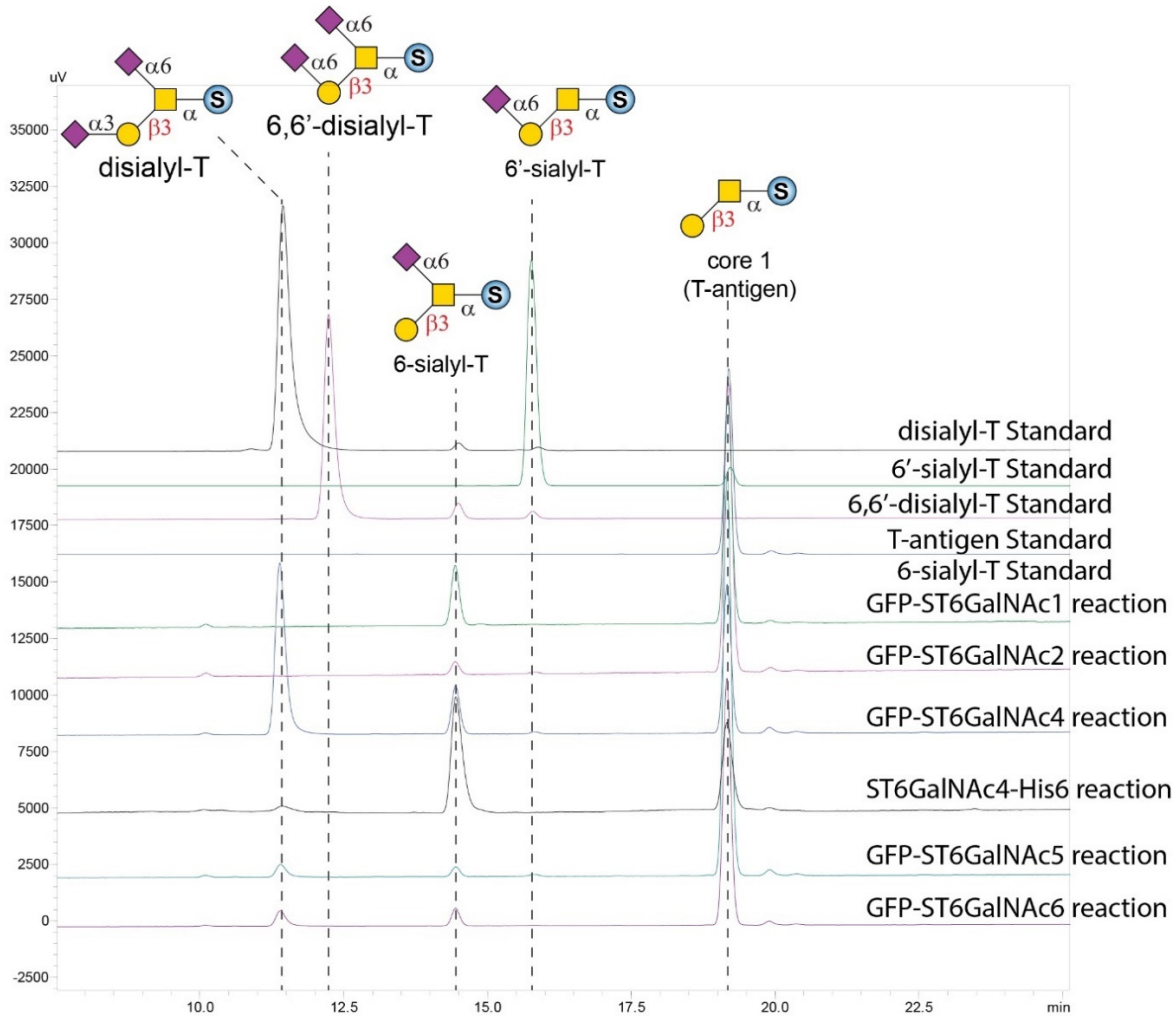


Figure S3: HPLC analysis of Human ST6GalNAcs-catazlyed reaction.

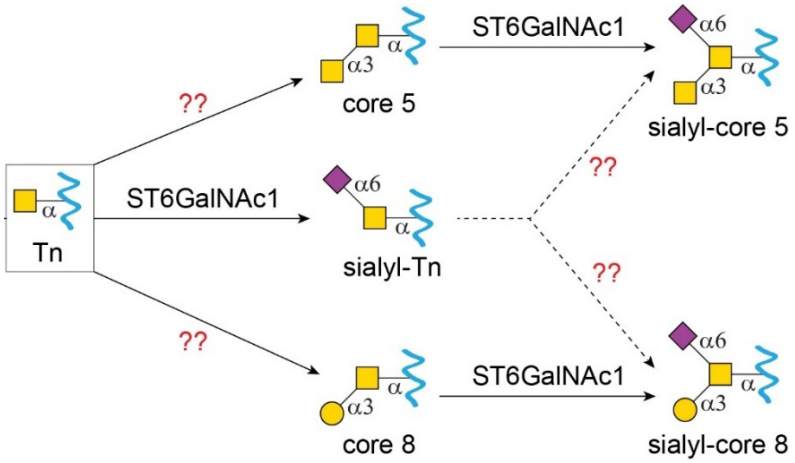


Figure S4: Proposed biosynthetic pathway of sialylated core 5 and core 8. One possible pathway is first the synthesis of core 5 and 8 by unknown GalNAc transferase or Gal transferase followed by ST6GalNAc1-catalyzed regioselective α2-6sialylation of the initial GalNAc residue. Alternatively, the two structures may be synthesized directly from sialyl-Tn by unknown glycosyltransferases. In both cases, ST6GalNAc1 is involved.

VI. Microarray fabrication, assay, and data

Fmoc removal

O-GalNAc glycans (50 μg) were dissolved in 200 μL H_2O, and 30 μL triethylamine was added to remove the Fmoc group at room temperature for 4 h. The reactions were then lyophilized, and hexane extraction was applied to remove free Fmoc.

Method for microarray fabrication

The O-GalNAc microarray was printed according to the guidelines of MIRAGE as summarized in Table S1. O-GalNAc glycans were prepared at a concentration of 100 μM in the printing buffer (150 mM phosphate, pH 8.5), and printed on Nexterion slide H-3D hydrogel coated glass microarray slides (Applied Microarrays Inc), each for replicates of six as described previously.[6] In addition, printing buffer was printed as a negative control; anti-human IgG conjugate with Cy3 (0.01 mg/mL) and anti-human IgG conjugate with Alexa 647 (0.01 mg/mL) were printed as a marker. Non-contact printing was performed at room temperature with a humidity of 60% by a sciFLEXARRAYER S3 spotter (Scienion) with one PDC 80 Piezo Dispense Capillary, and 16 subarrays were printed on each slide. After overnight dehumidification under room temperature, the slides were washed with MilliQ water and subsequently blocked with 50 mM ethanolamine in 100 mM Tris-HCl buffer (pH 9.0) for 2 hours. The blocked slides were then washed with MilliQ water twice, dried, and stored desiccated at -20 °C until use.

Method for microarray assay

All biotinylated lectins were purchased from Vector Lab (Burlingame, CA), including two T antigen specific lectins (*Arachis hypogaea* lectin, PNA; *Artocarpus integrifolia* lectin, Jacalin), three Tn antigen specific lectins (Wisteria Floribunda Lectin, WFL; *Soybean* agglutinin, SBA; *Vicia villosa* lectin, VVL; *Dolichos biflorus* agglutinin, DBA). Tn antigen monoclonal antibody (Tn 218, mouse IgM) and sialosyl-Tn antigen monoclonal antibody (STn 219, mouse IgG1) were purchased from ThermoFisher. Human PSGL-1/CD162 monoclonal antibody (Clone # CHO131, mouse IgM) and human MUC-1 polyclonal antibody (sheep IgG) were purchased from R&D Systems. Alexa Fluor™ 633 labeled goat anti-mouse IgM (heavy chain) cross-adsorbed secondary antibody and Alexa Fluor™ 647 labeled goat anti-mouse IgG (H+L) cross-adsorbed secondary antibody were from ThermoFisher. CF™ 633 labeled donkey anti-sheep IgG (H+L) highly cross-adsorbed secondary antibody was produced Milliper-Sigma.

Microarray slides were rehydrated for 30 min in blocking buffer (50 mM ethanolamine in 50 mM sodium borate, pH 9.2) and washed with H_2O before assay. All assays were performed as previously reported.[7] Biotinylated lectins were detected by Cy5-streptavidin (1 μg/mL). The concentration of secondary antibody in detecting anti-glycan antibodies is 5 μg/mL. After binding, the slides were scanned with a GenePix 4000B scanner and the collected data was analyzed with GenePix Pro. No bindings were observed for Tn218, PSGL-1 antibody, and MUC-1 antibody.

Table S1. Glycan microarray information based on MIRAGE.

Classification	Guidelines
Sample: Glycan Binding Sample	
Description of Sample	All lectins and antibodies were purchased from Vector Lab.
Sample modifications	Not applicable.
Assay protocol	Microarray analyses were performed essentially as described in NCFG website.
2. Glycan Library	
Glycan description for defined glycans	All glycans were synthesized as described in body text.
Glycan description for undefined glycans	Not applicable.
Glycan modifications	Glycans were linked with Ser.
3. Printing Surface; e.g., Microarray Slide	
Description of surface	Nexterion slide H-3D hydrogel coated glass microarray slides.
Manufacturer	Applied Microarrays Inc (Tempe, AZ, USA)
Custom preparation of surface	Not applicable.
Covalent Immobilization	Glycans were linked with Ser for robotically arraying and the amine group could covalently be immobilized on NHS ester coated glass slide.
4. Arrayer (Printer)	
Description of Arrayer	sciFLEXARRAYER S3 spotter (Scienion) with one PDC 80 Piezo Dispense Capillary, and 16 subarrays were printed on each slide
Dispensing mechanism	Non-contact liquid delivery.
Glycan deposition	Each glycan probe was printed at 1 deposit in 6 replicates.
Printing conditions	Samples were prepared at a concentration of 100 μM in the printing buffer (150 mM phosphate, pH 8.5), printing was performed at room temperature and relative humidity of 60%.
5. Glycan Microarray with “Map”	
Array layout	Each array slide contained 16 identical subarrays (pads). Each subarray contained 25 unique samples.
Glycan identification and QC	Quality control included analyses with plant lectins.
6. Detector and Data Processing	
Scanning hardware	GenePix 4000B Microarray Scanner (Molecular Devices, LLC)
Scanner settings	Laser channel: wavelength 635 nm or 535 nm PMT gain: 600 for 535 nm and 800 for 635 nm Scan power:100%
Image analysis software	GnePix Pro (Molecular Devices, LLC)
Data processing	Data was processed by using Excel. No particular normalization method or statistical analysis was used.
7. Glycan Microarray Data Presentation	
Data presentation	The microarray binding results are in Figure 3, Table S2. Binding results are presented as relative fluorescence intensity units (RFU) of binding in mean and S.D.
8. Interpretation and Conclusion from Microarray Data	
Data interpretation	No software or algorithms were used to interpret processed data.
Conclusions	Described in Results parts.

Table S2. Glycan microarray data of lectins.

Glycan No.	WFA		Jacalin		SBA		DBA		PNA		VVL		STn219	
	Average	SD	Average	SD	Average	SD	Average	SD	Average	SD	Average	SD	Average	SD
24	40877	234	65386	5	30208	374	1872	55	11	3	48015	290	0	0
25	9	5	-11	5	5	6	7	3	9	2	14	4	4893	826
26	7	2	65437	1	-12	6	9	1	48137	798	11	5	0	0
27	8	5	46588	6379	2	1	7	2	7	2	6	2	0	0
28	6	3	42	14	-12	8	12	3	36977	388	8	5	0	0
29	8	3	20	32	6	2	6	2	10	4	11	1	0	0
30	6	2	-9	7	-8	5	13	2	39986	76	9	1	0	0
31	8	4	65402	5	4	6	8	1	10	3	9	3	0	0
32	8	2	784	26	-12	8	11	1	10	3	6	3	0	0
33	7	2	-17	9	6	3	11	4	10	2	8	7	0	0
10	29631	4559	1188	482	30493	816	43374	313	9	4	29907	328	14403	2276
15	39164	1368	-5	9	7099	143	35418	651	13	4	33816	536	0	0
18	28477	809	33	18	1612	81	48160	914	12	5	16745	1823	0	0
21	26875	1256	-11	9	2461	58	42427	138	3	3	16072	974	0	0
34	5	1	-2	3	1	2	12	2	6	3	6	2	0	0
14	39932	3843	65273	18	50873	236	16	6	1	8	43034	200	0	0
16	13	7	25	34	4	3	8	4	8	4	8	3	0	0
19	12	9	16	16	7	7	14	2	7	3	12	6	0	0
22	117	8	1746	238	61	12	10	2	8	3	416	60	0	0
11	12	3	65353	7	14	3	10	5	10	3	12	2	0	0
17	13	2	-20	11	-2	8	6	2	8	1	13	3	0	0
20	11	5	457	84	14	5	10	1	11	6	6	1	0	0
23	8	3	64	36	7	3	13	4	5	1	9	2	0	0
NC	8	1	11	21	8	6	9	3	10	2	8	2	0	0
M	17739	737	25518	2081	21548	398	23042	321	22591	613	22902	408	21893	1255

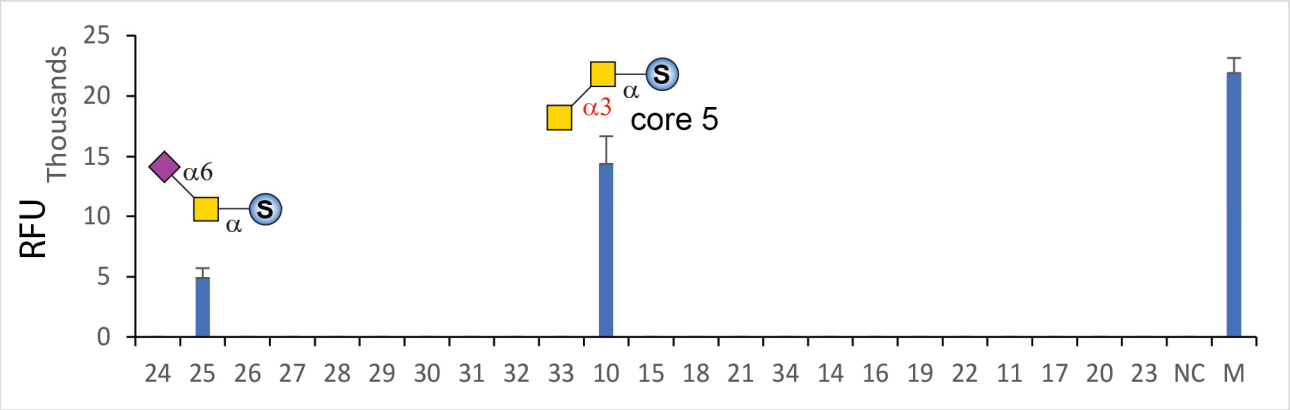
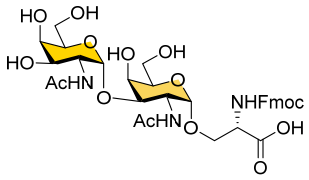


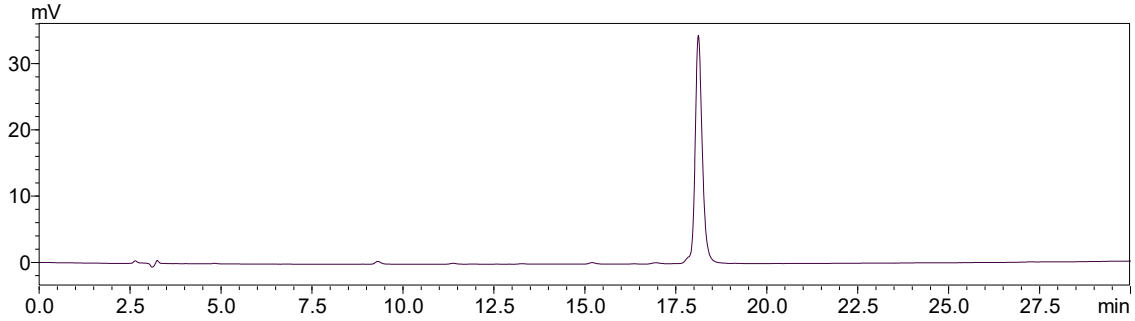
Figure S5: Glycan microarray of sialosyl-Tn antigen antibody STn 219 (10 mg/μL). As expected, the antibody binds STn-Ser, but surprisingly, the antibody also binds core 5 disaccharide with a higher RFU value. The results suggested that the monoclonal antibody STn 219 is not specific and may be used with caution.

VII. HPLC, MS, and NMR data of purified compounds

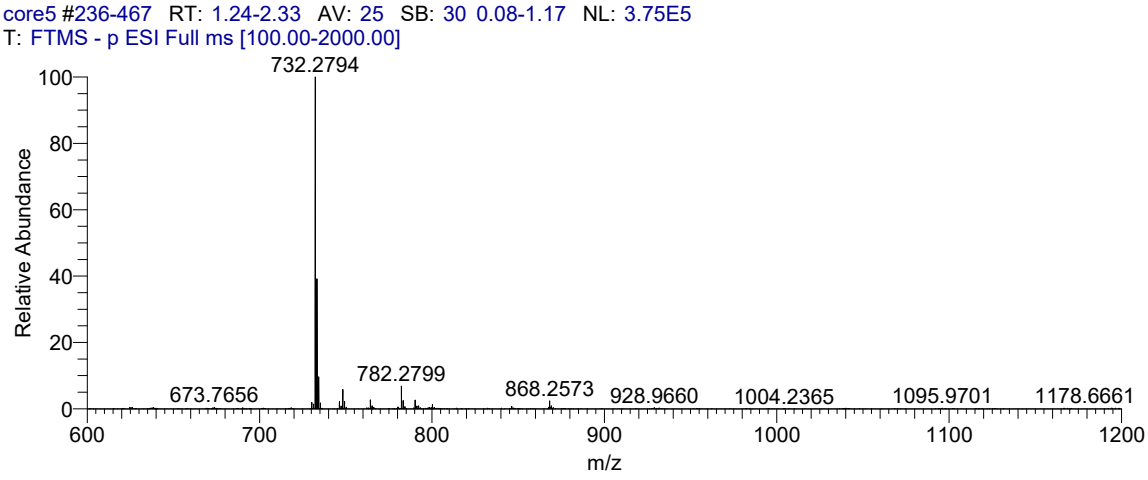
Core 5 (compound **10**)



HPLC, retention time = 18.121 min

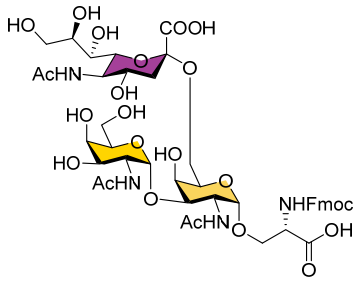


HR-MS, found 732.2794 $[M-H]^-$

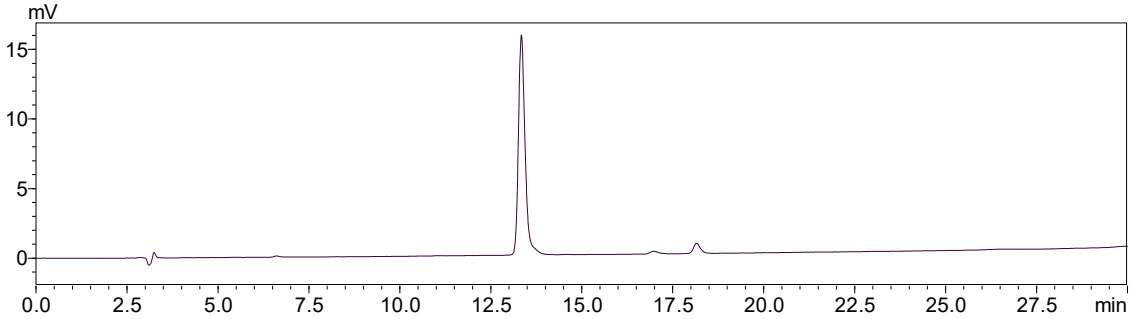


NMR data, see part II.

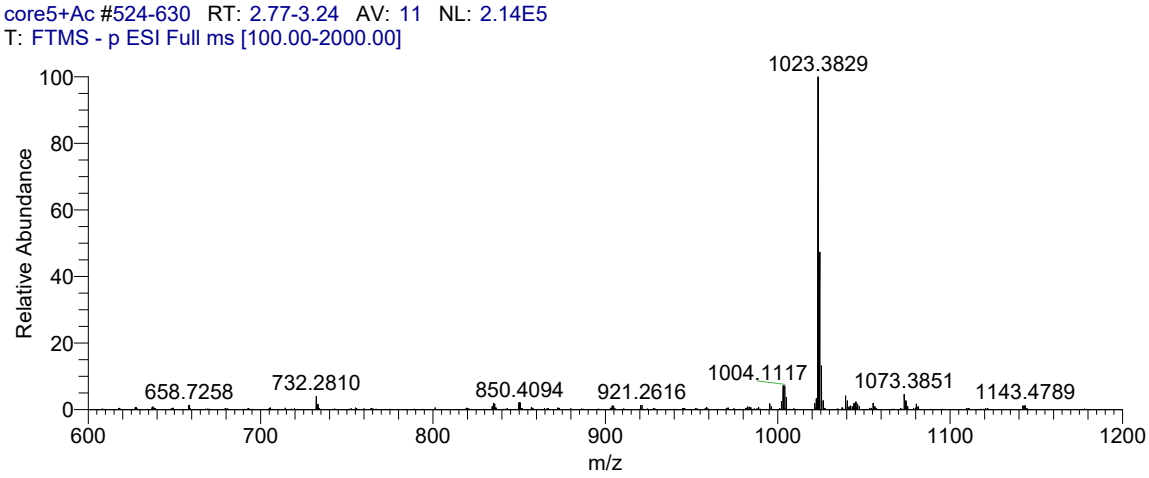
Sialyl-core 5 (Neu5Ac) (compound **15**), 7 mg.



HPLC, retention time = 13.339 min

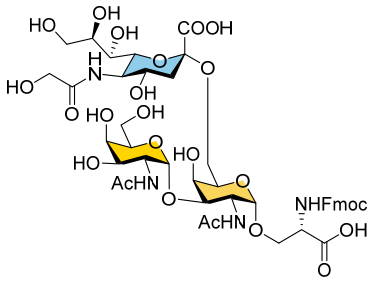


HR-MS, found 1023.3829 $[M-H]^-$

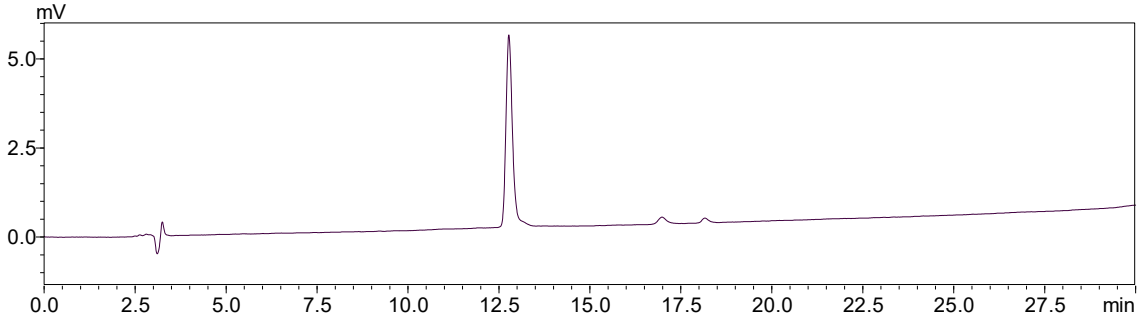


NMR: ^{1}H NMR (600 MHz, Deuterium Oxide) δ 7.68 (s, 2H), 7.51 (s, 2H), 7.29 (d, *J* = 34.6 Hz, 4H), 4.82 (s, 1H), 4.66 (d, *J* = 19.5 Hz, 1H), 4.43 (s, 2H), 4.18 – 4.03 (m, 4H), 3.87 (d, *J* = 3.3 Hz, 2H), 3.82 – 3.38 (m, 18H), 2.62 (dd, *J* = 12.5, 4.7 Hz, 1H), 1.94 (d, *J* = 5.5 Hz, 6H), 1.89 – 1.75 (m, 3H), 1.54 (t, *J* = 12.3 Hz, 1H). ^{13}C NMR (151 MHz, D_2O) δ 174.98, 174.52, 174.12, 173.06, 157.45, 143.70, 140.92, 140.85, 127.88, 127.42, 124.82, 120.09, 100.14, 97.57, 93.33, 72.55, 71.68, 71.60, 71.14, 69.32, 68.23, 68.19, 68.11, 67.79, 67.58, 66.20, 64.14, 63.33, 62.56, 60.81, 54.63, 51.81, 49.28, 47.69, 46.76, 40.18, 21.98.

Sialyl-core 5 (Neu5Gc) (compound **18**), 1.3 mg

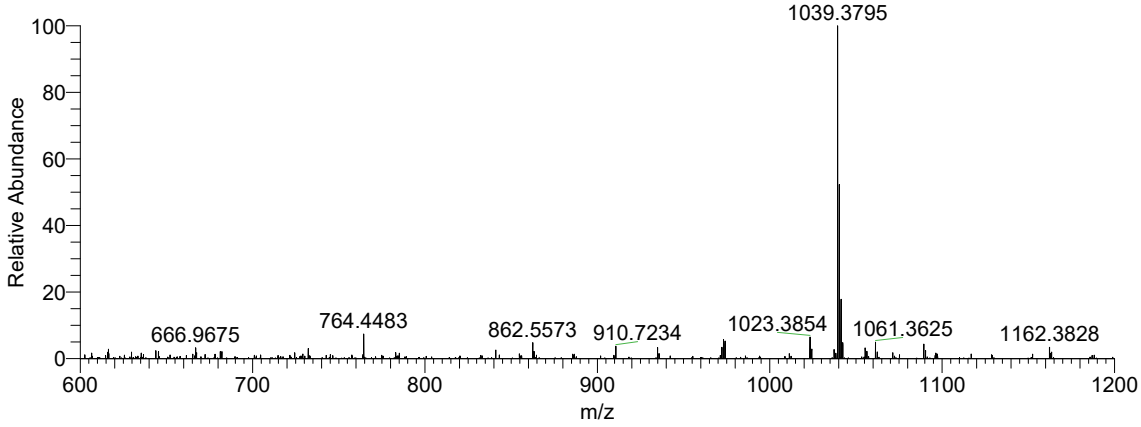


HPLC, retention time = 12.772 min

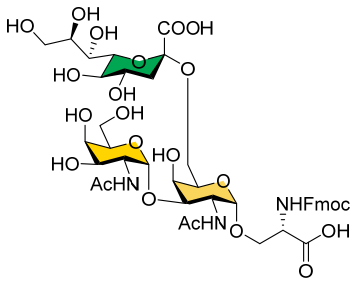


HRMS, found 1039.3795 $[M-H]^-$

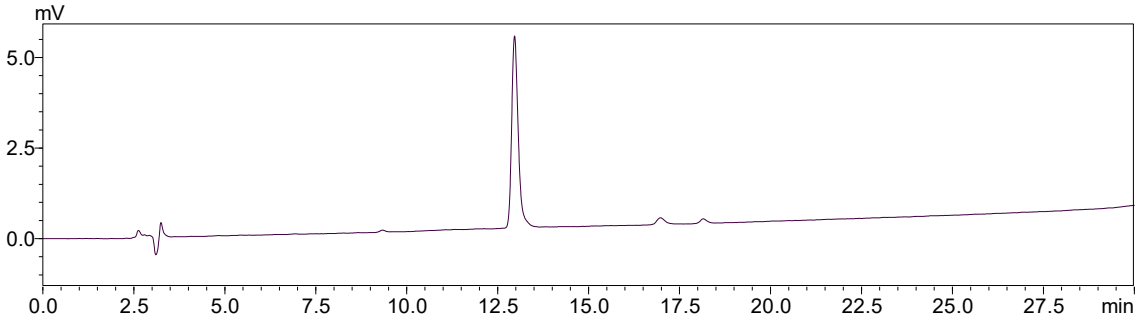
core5+Gc #523-547 RT: 2.90-3.00 AV: 3 NL: 8.24E4
T: FTMS - p ESI Full ms [100.00-2000.00]



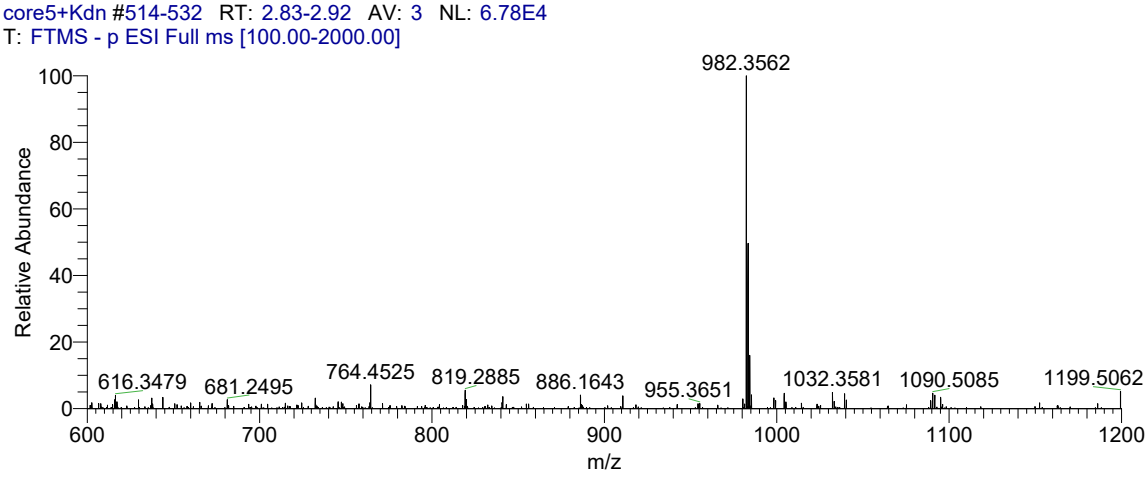
Sialyl-core 5 (KDN) (compound **21**), 1.1mg



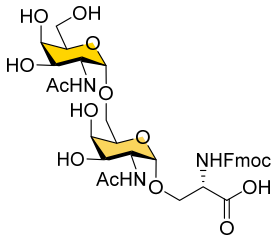
HPLC, retention time = 12.964 min



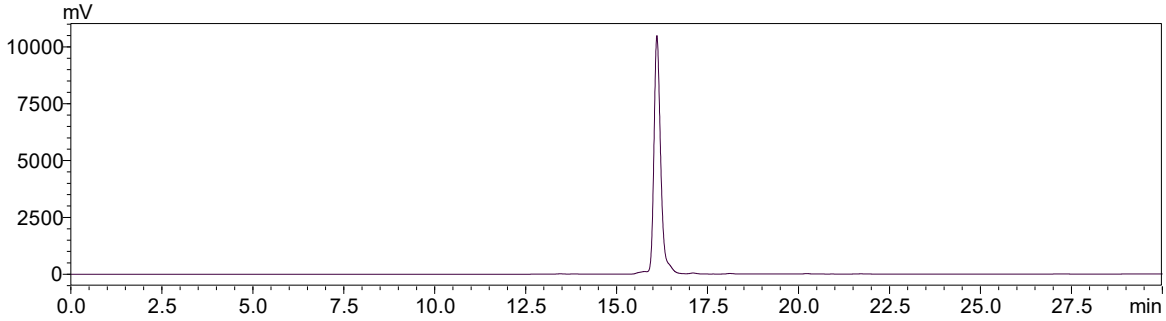
HRMS, found 982.3562 $[M-H]^-$



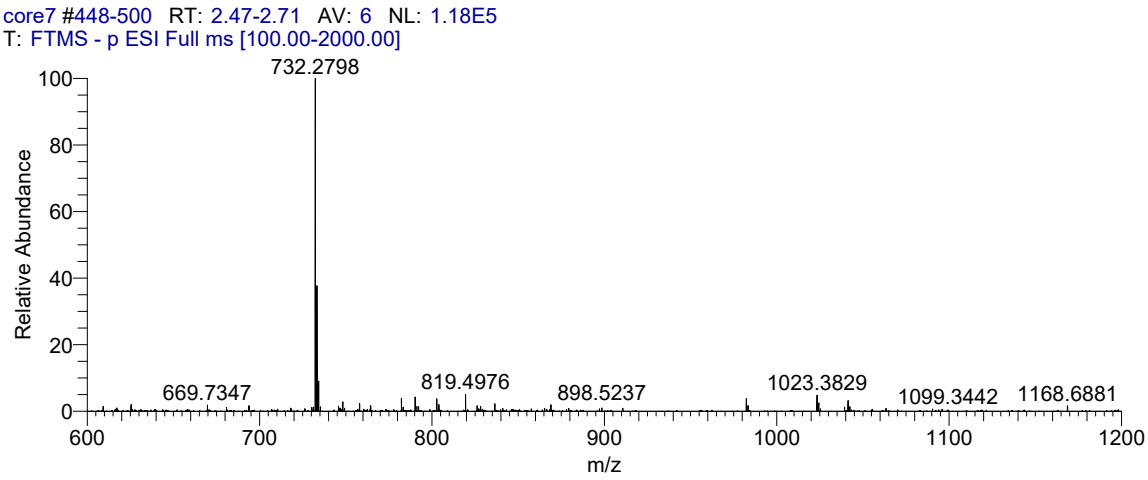
Core 7 (compound **14**)



HPLC, retention time = 16.108 min

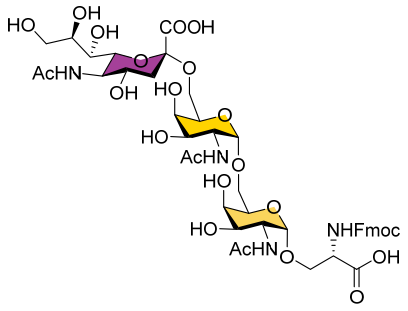


HRMS, found 732.2798 $[M-H]^-$

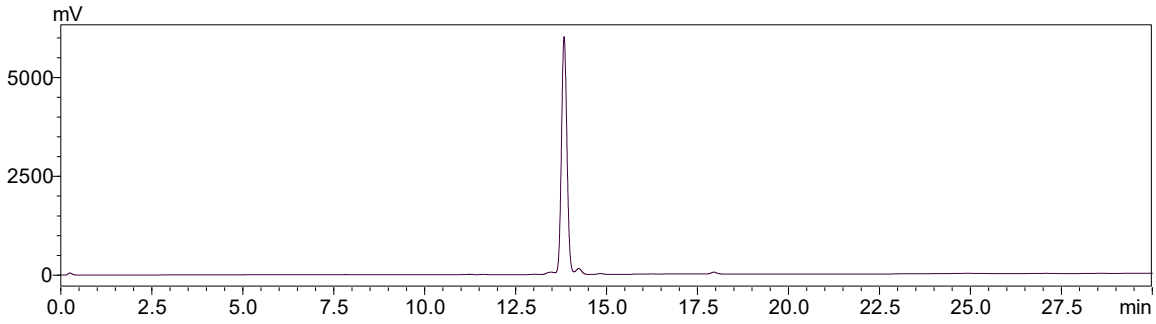


NMR data: see part II.

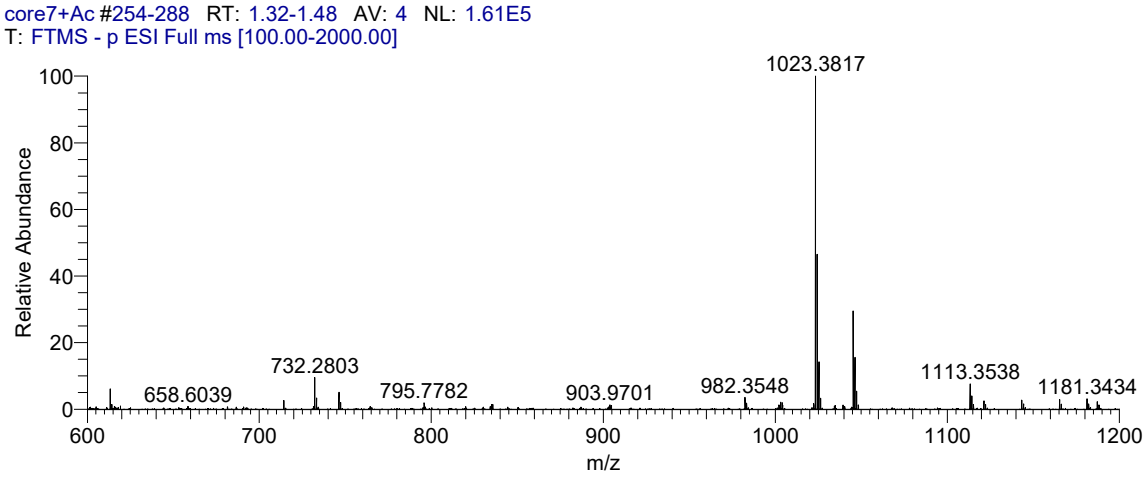
Sialyl-core 7 (Neu5Ac) (compound **16**), 1.6 mg



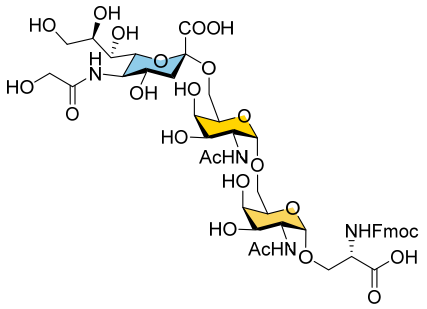
HPLC, retention time = 13.832 min



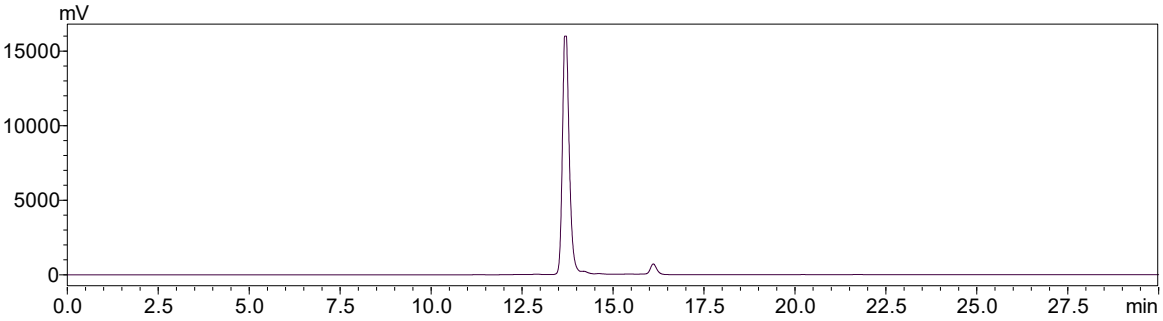
HRMS, found 1023.3817 [M-H]$^-$; 1045.3635 [[M+Na-2H]$^-$



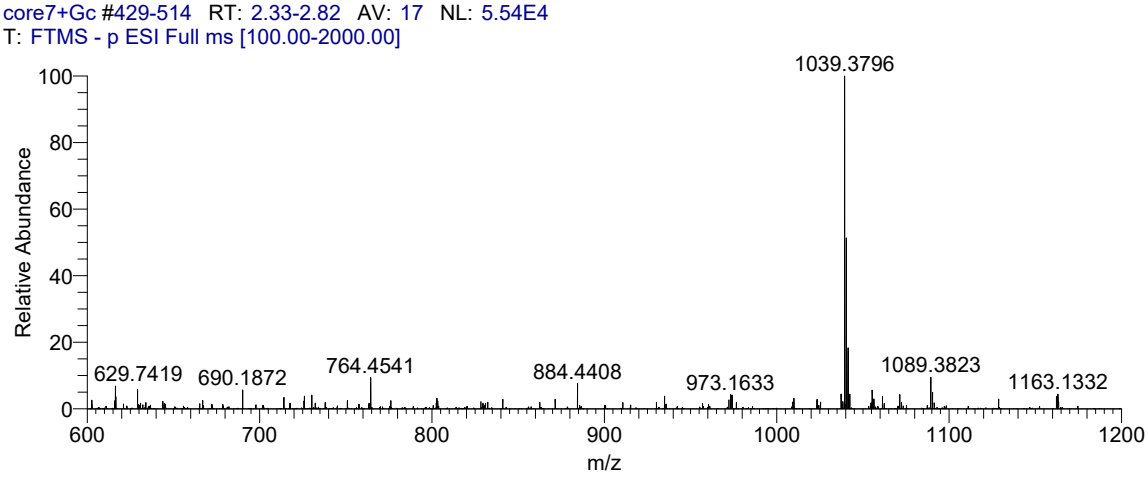
Sialyl-core 7 (Neu5Gc) (compound **19**)， 0.8 mg



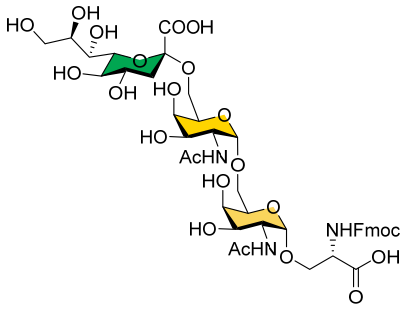
HPLC, retention time = 13.528 min



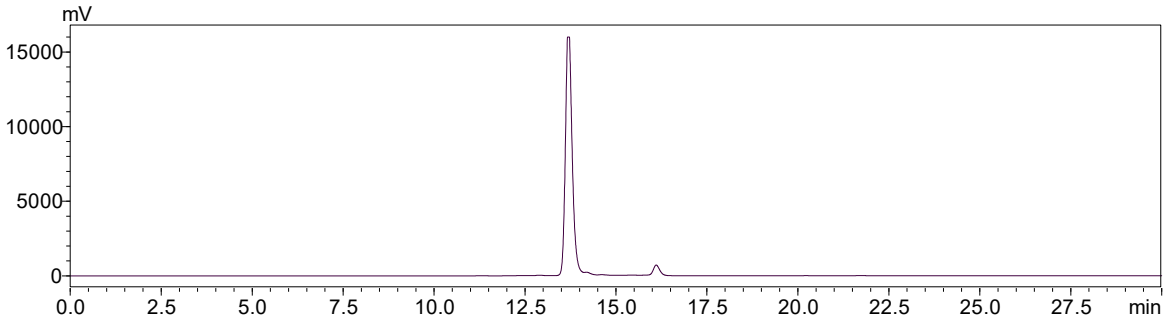
HRMS, found 1039.3796 $[M-H]^-$;



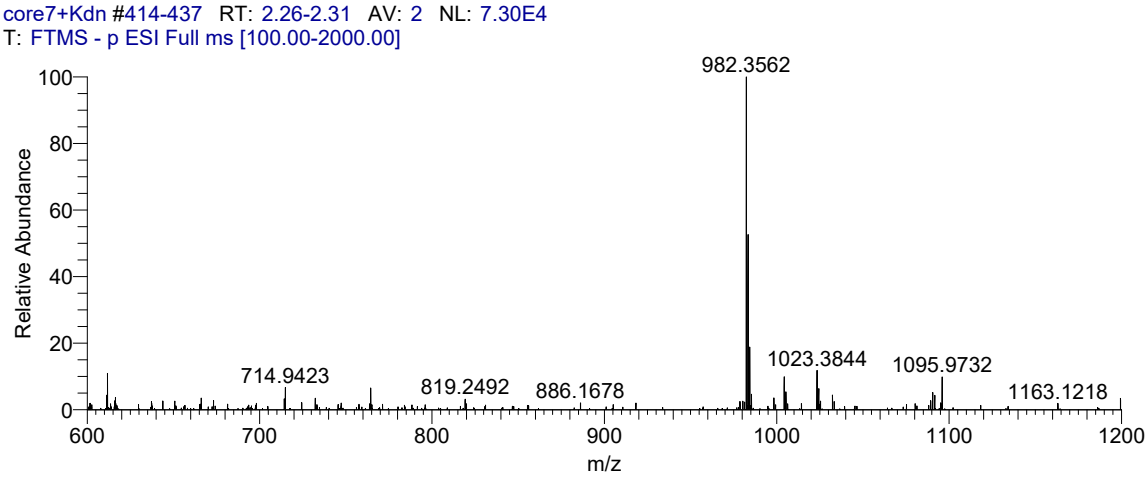
Sialyl-core 7 (KDN) (compound **22**), 0.9 mg



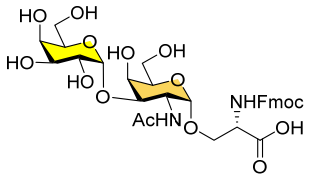
HPLC, retention time = 13.715 min



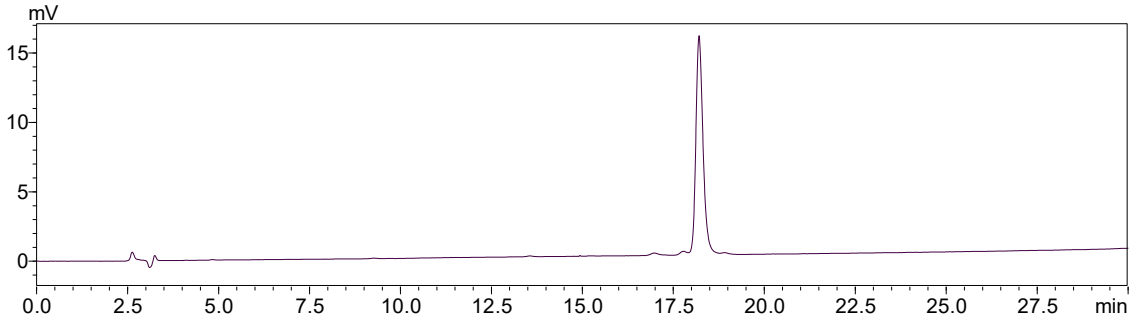
HRMS, found 982.3562 $[M-H]^-$;



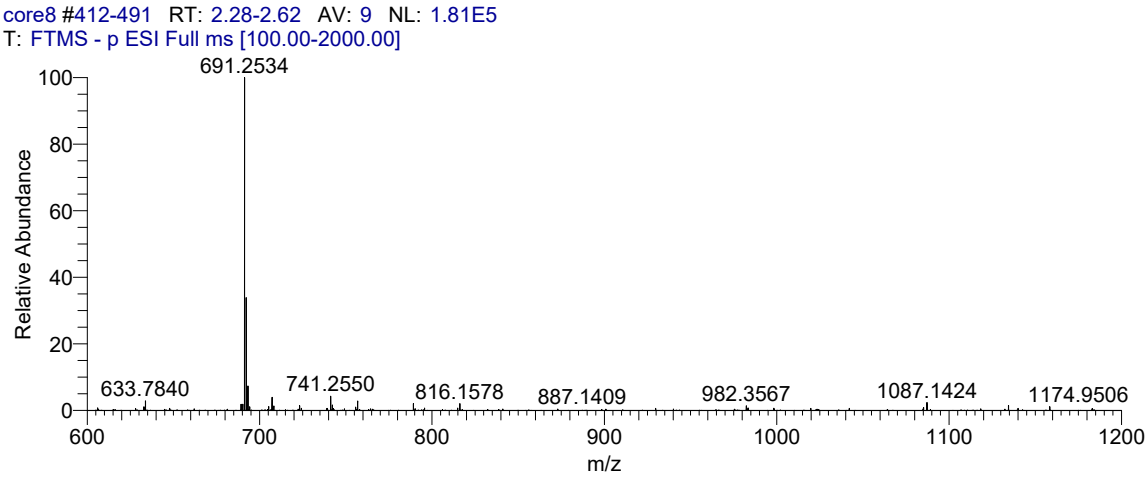
Core 8 (compound **11**)



HPLC, retention time = 18.207 min

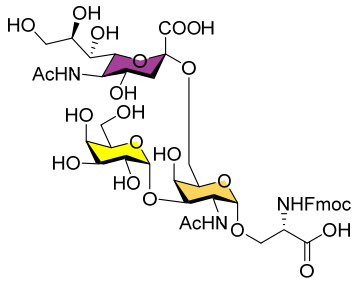


HRMS, found 691.2534 $[M-H]^-$

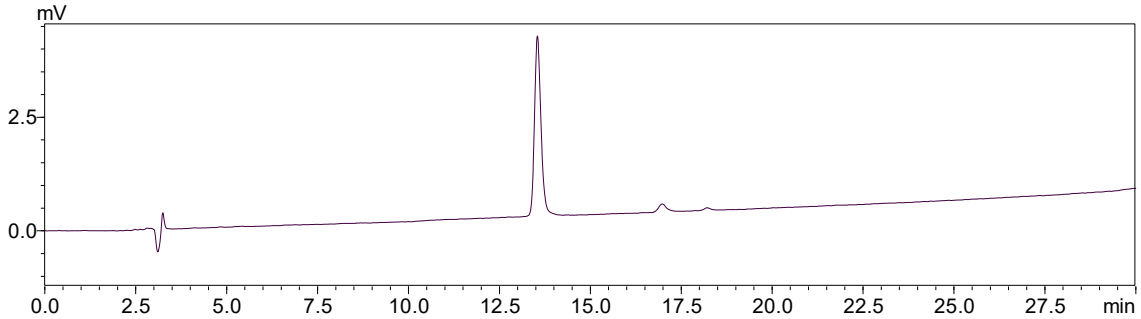


NMR data: see part II.

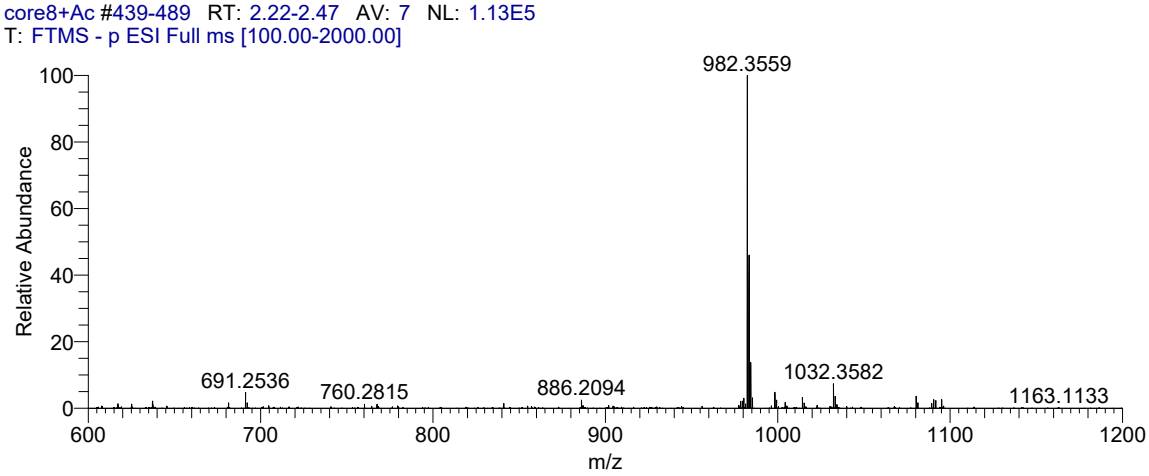
Sialyl-core 8 (Neu5Ac) (compound **17**), 6 mg



HPLC, retention time = 13.545 min

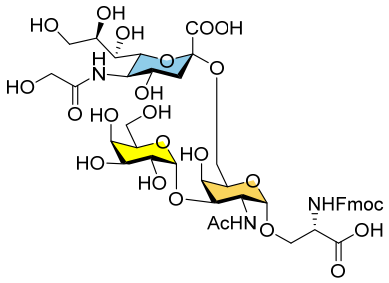


HRMS, found 982.3559 $[M-H]^-$;

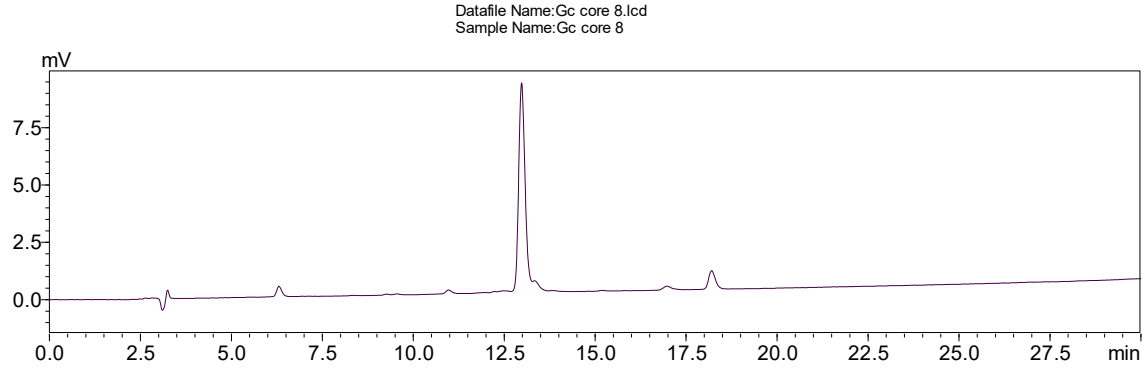


NMR: ^{1}H NMR (600 MHz, D_2O) δ 7.83 – 7.69 (m, 2H), 7.58 (d, *J* = 7.6 Hz, 2H), 7.34 (dt, *J* = 37.1, 8.1 Hz, 4H), 4.88 (d, *J* = 3.7 Hz, 1H), 4.49 (d, *J* = 27.5 Hz, 2H), 4.19 (dt, *J* = 10.8, 3.1 Hz, 1H), 4.12 (s, 1H), 4.00 (d, *J* = 13.1 Hz, 1H), 3.85 – 3.39 (m, 20H), 3.38 – 3.29 (m, 1H), 2.67 (dd, *J* = 12.4, 4.6 Hz, 1H), 1.90 (d, *J* = 75.5 Hz, 6H), 1.57 (t, *J* = 12.2 Hz, 1H). ^{13}C NMR (151 MHz, D_2O) δ 176.00, 175.02, 174.18, 173.27, 157.46, 143.87, 140.96, 127.98, 127.51, 124.93, 124.83, 120.14, 100.37, 97.69, 95.01, 72.77, 72.55, 72.25, 72.05, 71.71, 71.18, 69.24, 69.10, 68.27, 66.15, 64.56, 63.42, 62.60, 62.48, 60.31, 59.31, 55.72, 51.87, 47.75, 46.86, 40.29, 22.10, 22.03.

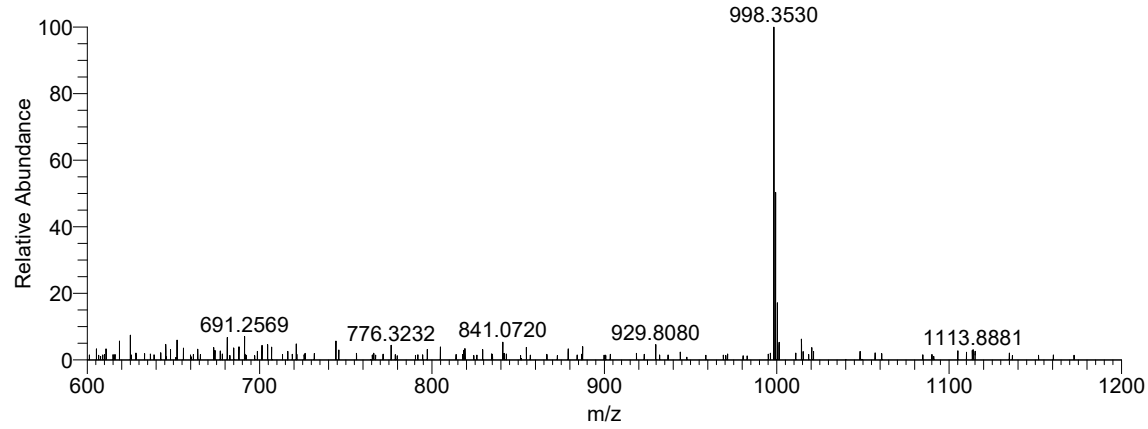
Sialyl-core 8 (Neu5Gc) (compound **20**), 1.4 mg



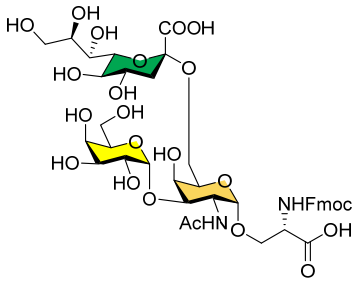
HPLC, retention time = 12.977 min



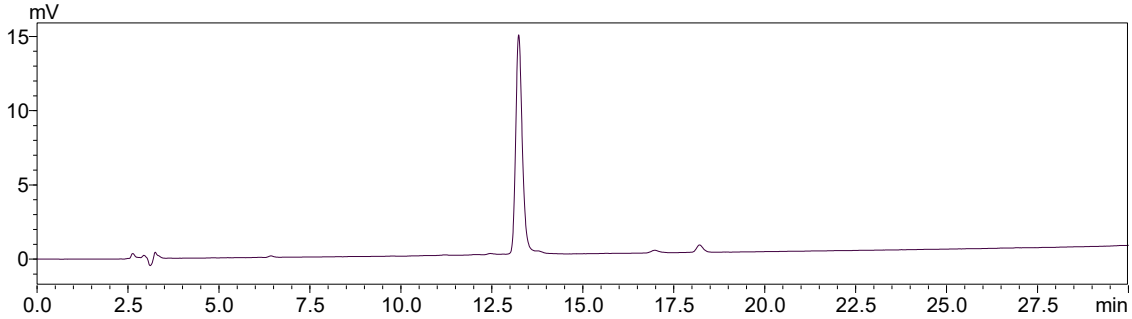
HRMS, found 998.3530 $[M-H]^-$



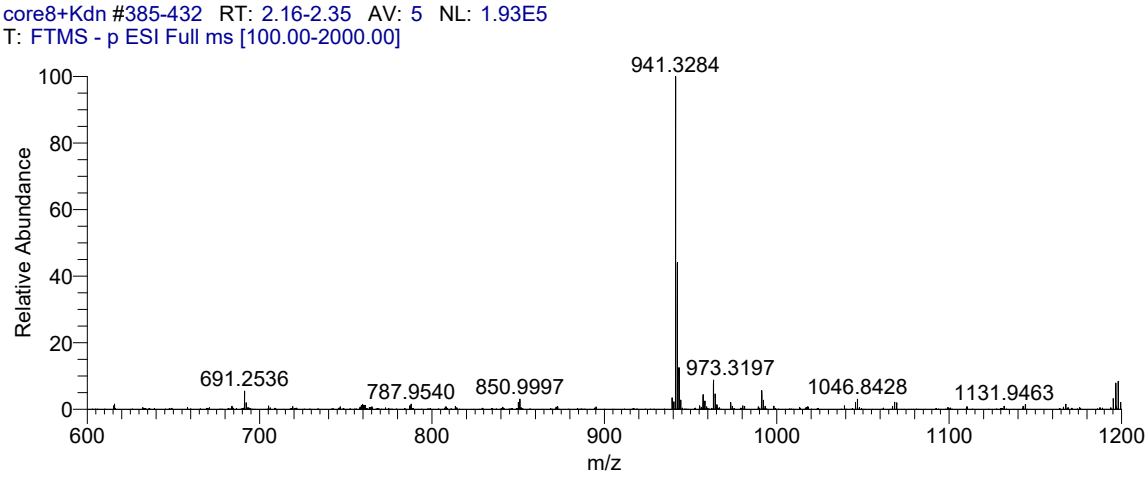
Sialyl-core 8 (KDN) (compound **23**), 0.8 mg



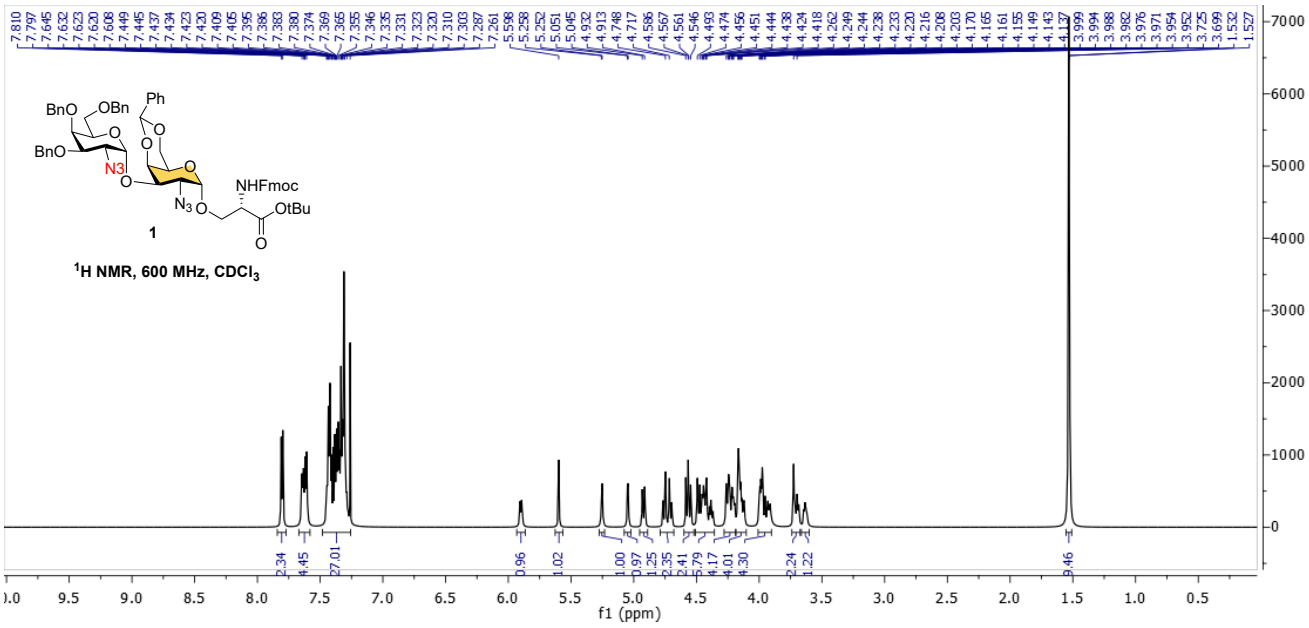
HPLC, retention time = 13.238

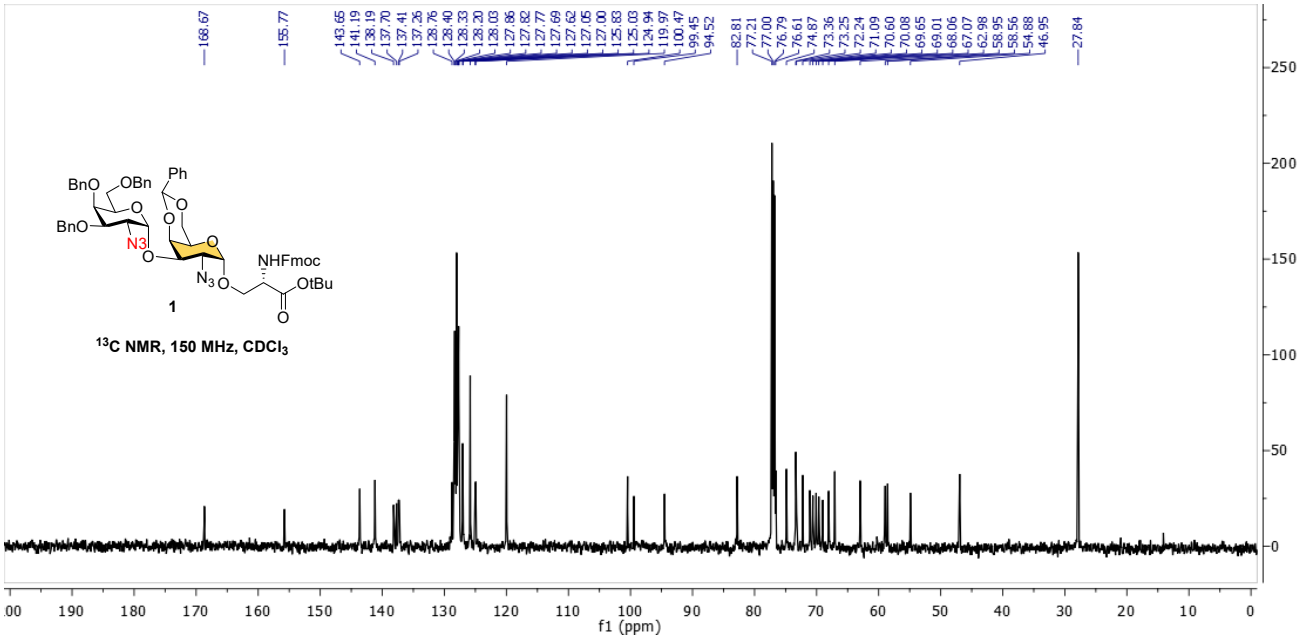


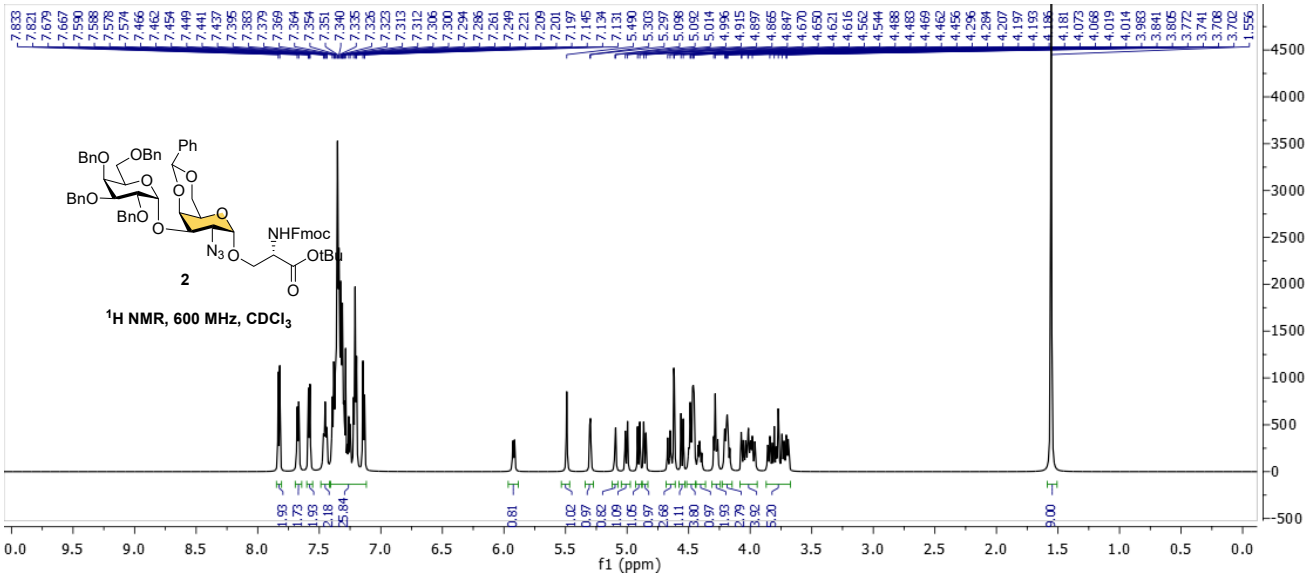
HRMS, found 941.3284 $[M-H]^-$

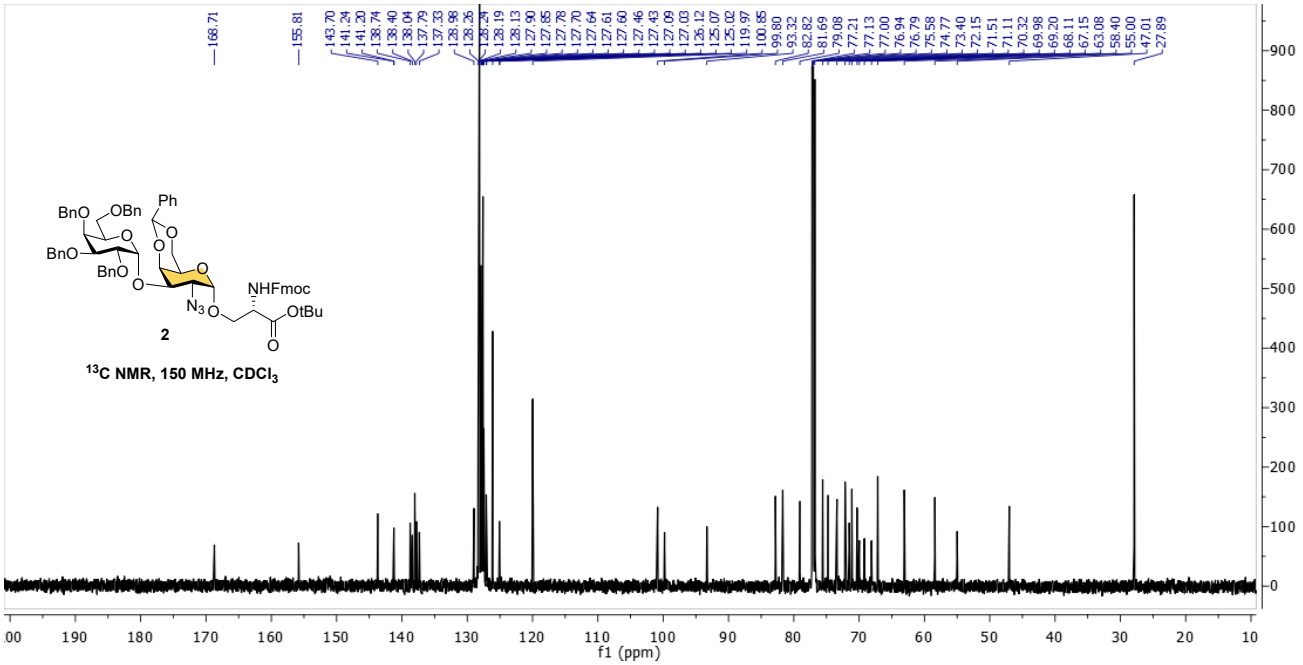


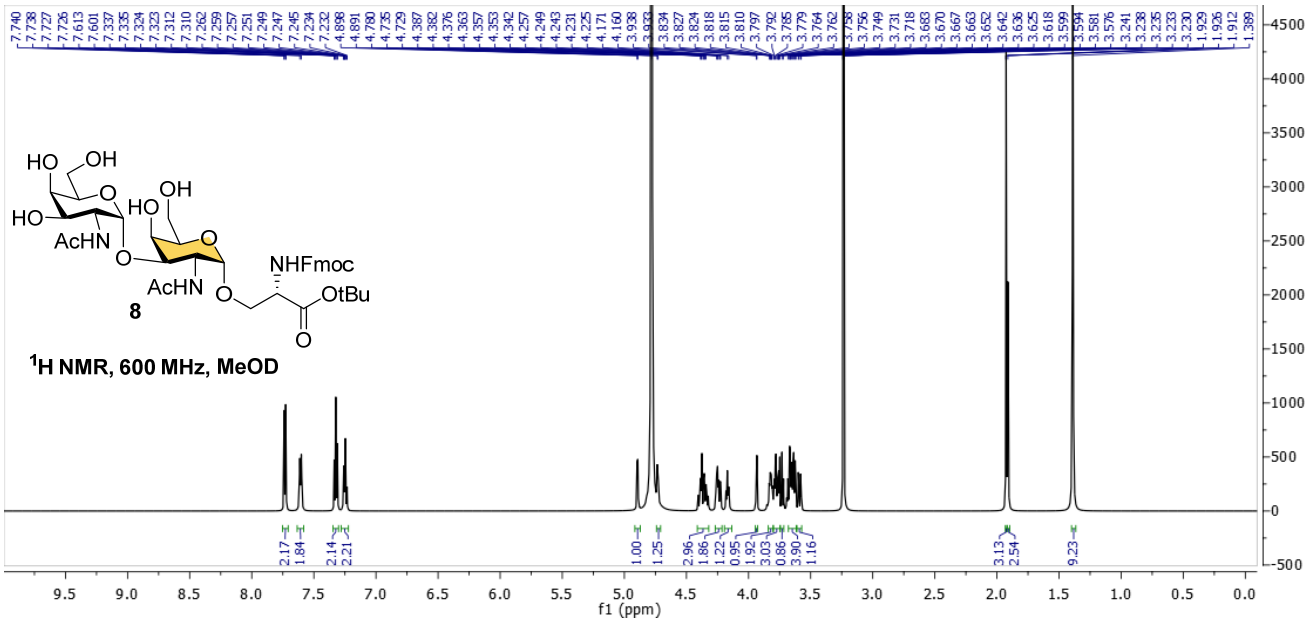
VIII. NMR Spectra

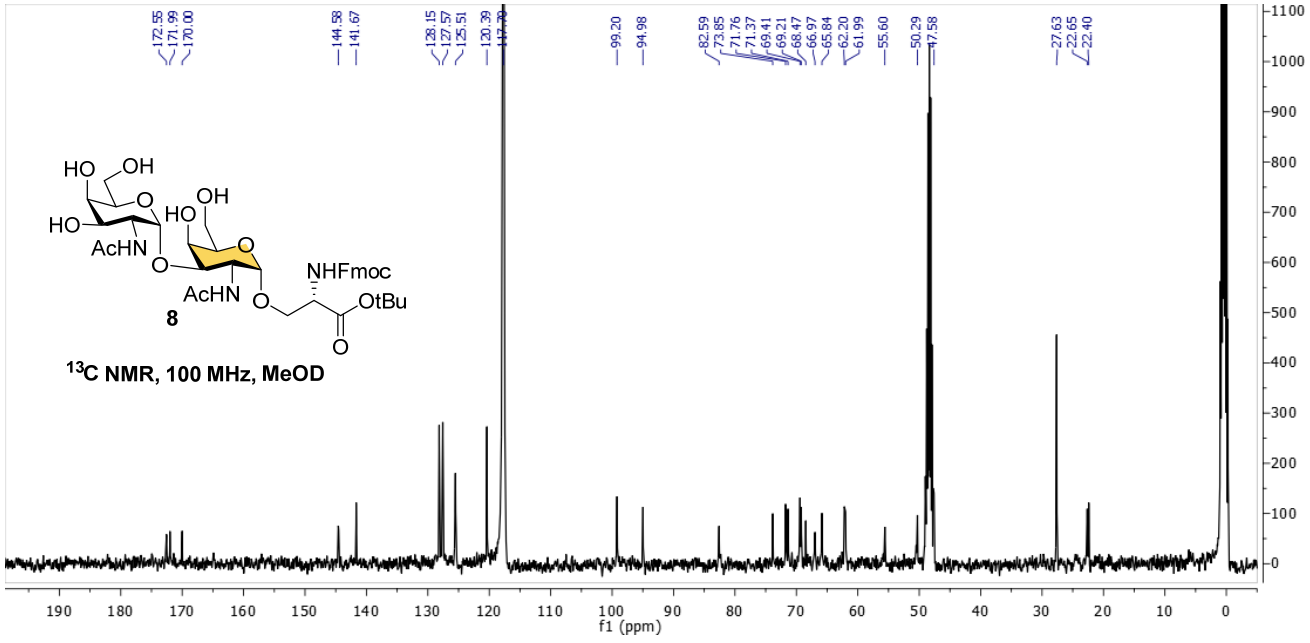


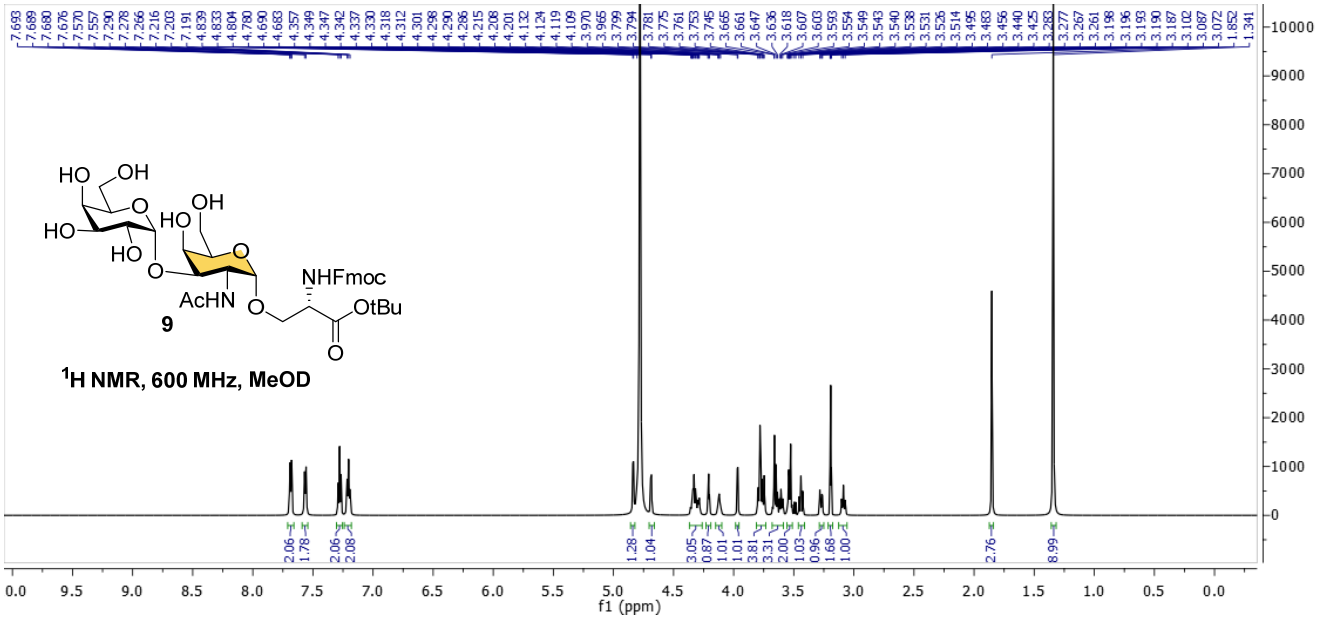


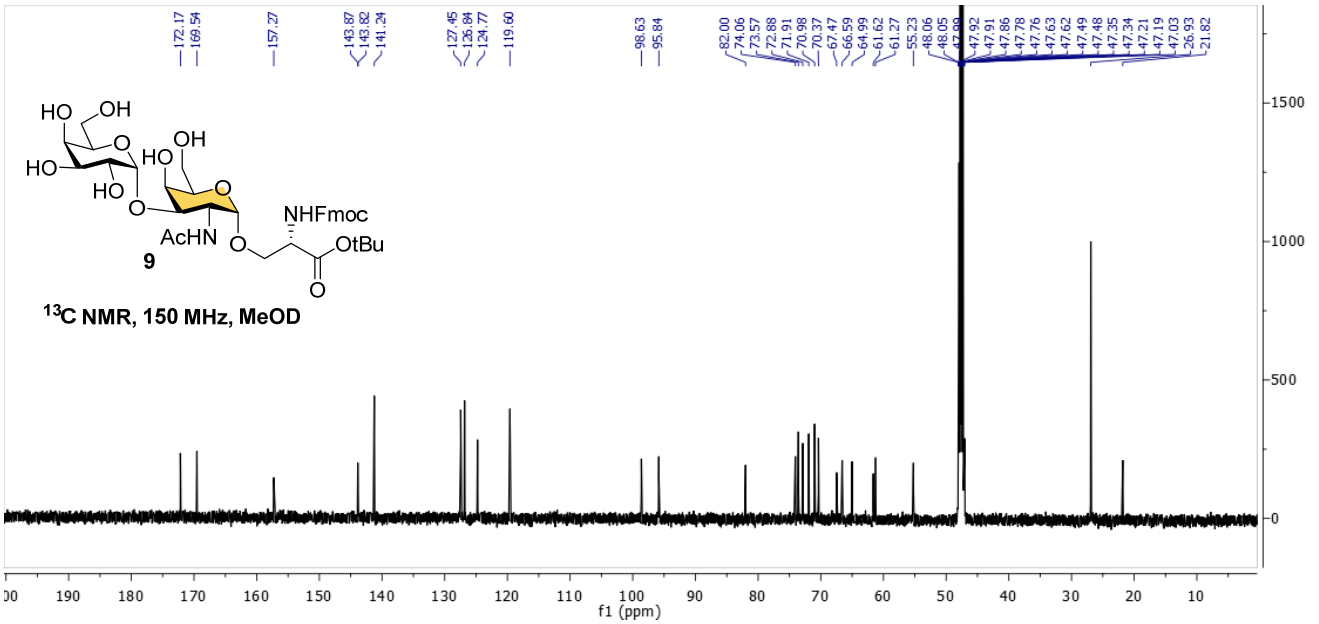

BnO
OBn
Ph
BnO
BnO
NHFmoc
OtBu
2
1H NMR, 600 MHz, CDCl3
f1 (ppm)

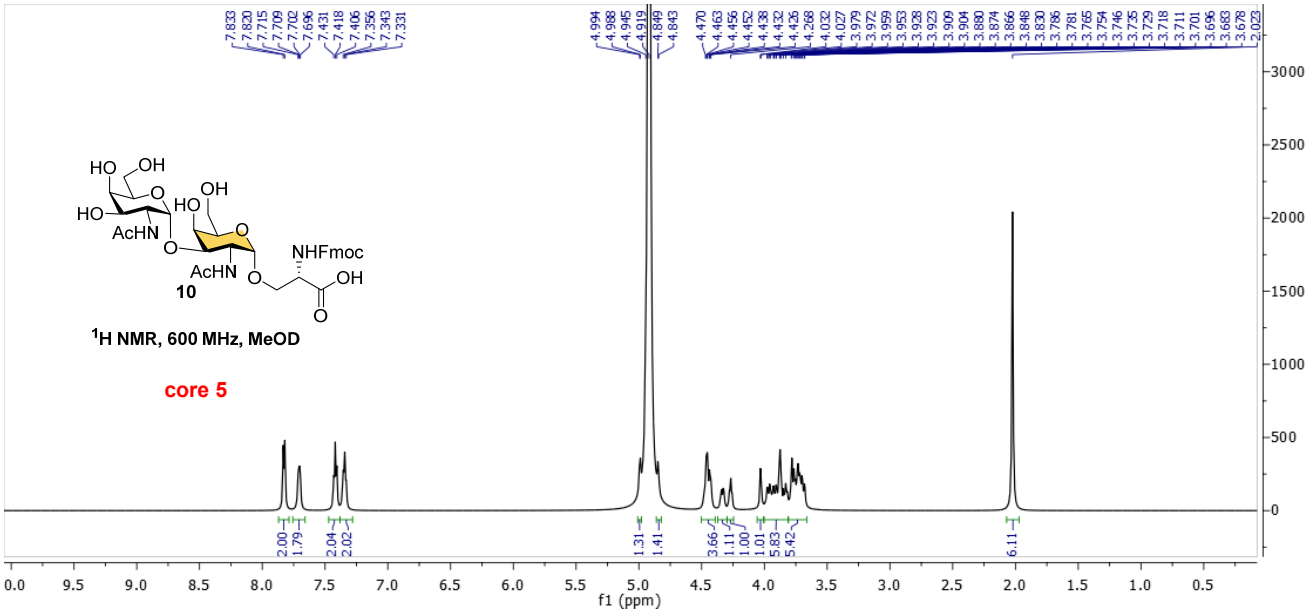

BnO
OBn
Ph
BnO
BnO
NHFmoc
OtBu
2
13C NMR, 150 MHz, CDCl3
f1 (ppm)

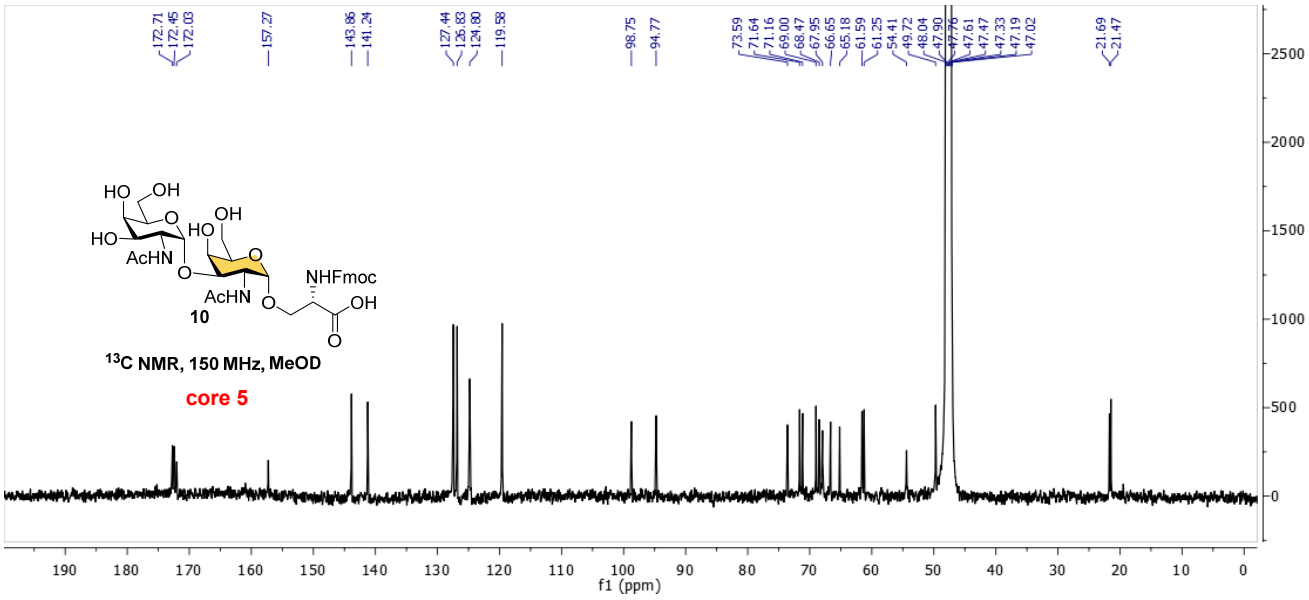

HO
OH
O
HO
OH
HO
AcHN
O
O
NHFmoc
AcHN
O
OtBu
O
8
1H NMR, 600 MHz, MeOD
f1 (ppm)

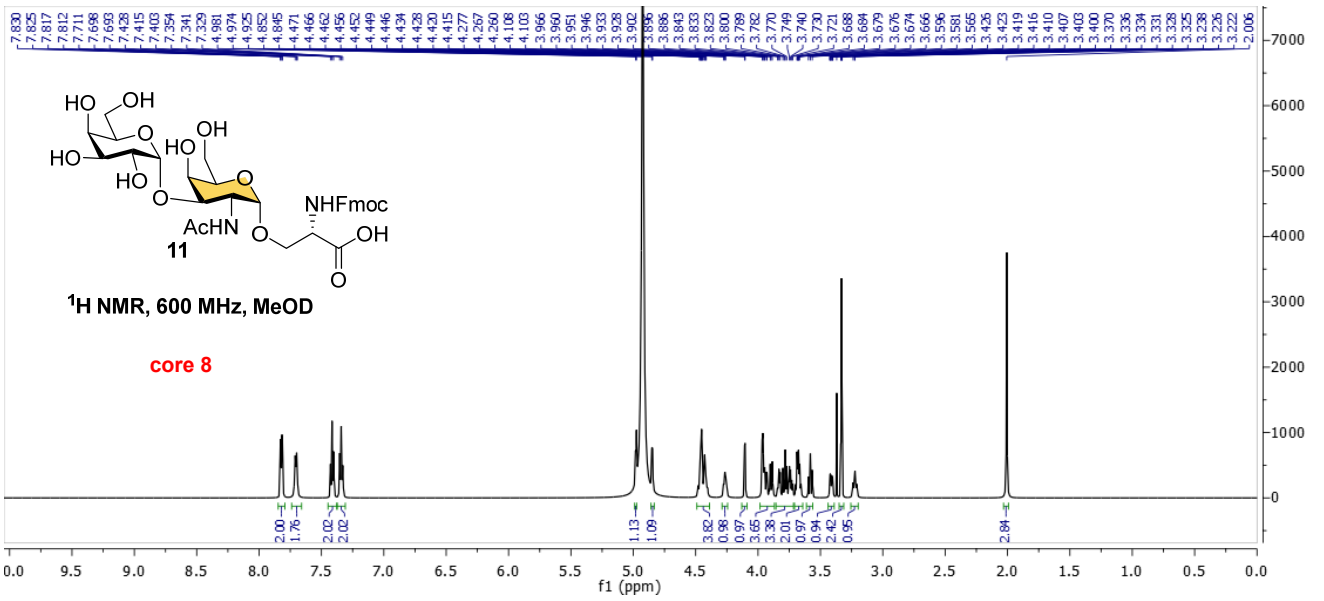

HO
OH
O
HO
OH
HO
AcHN
O
O
NHFmoc
AcHN
O
OtBu
O
8
13C NMR, 100 MHz, MeOD
f1 (ppm)

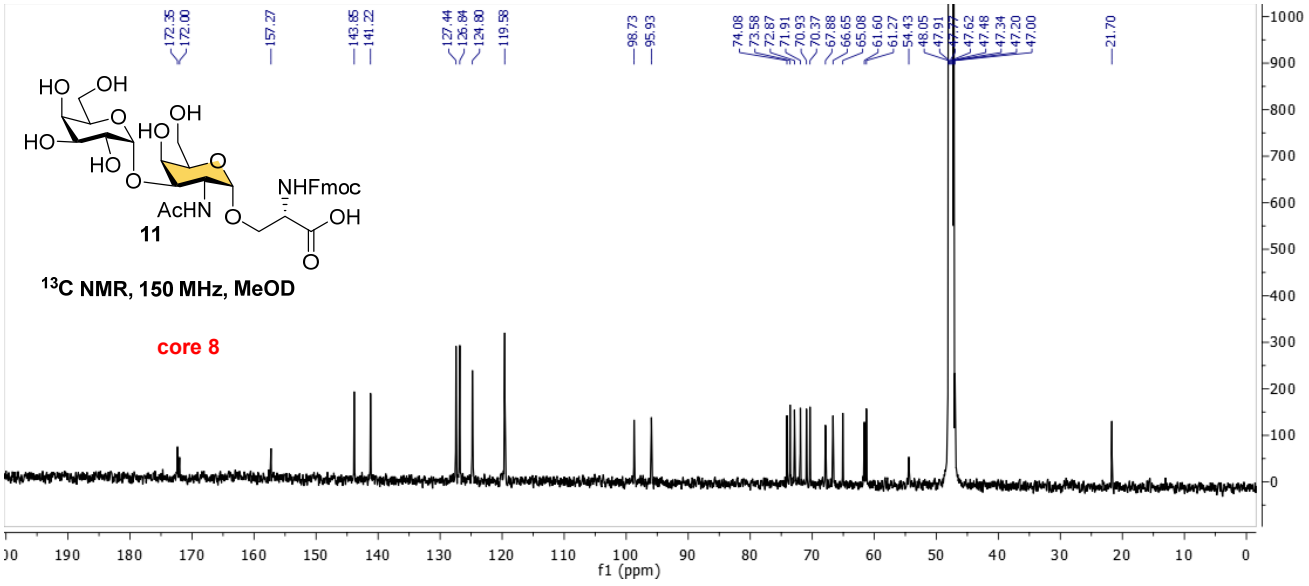

HO
OH
HO
HO
HO
OH
O
O
AcHN
NHFmoc
O
OtBu
O
9
1H NMR, 600 MHz, MeOD
f1 (ppm)

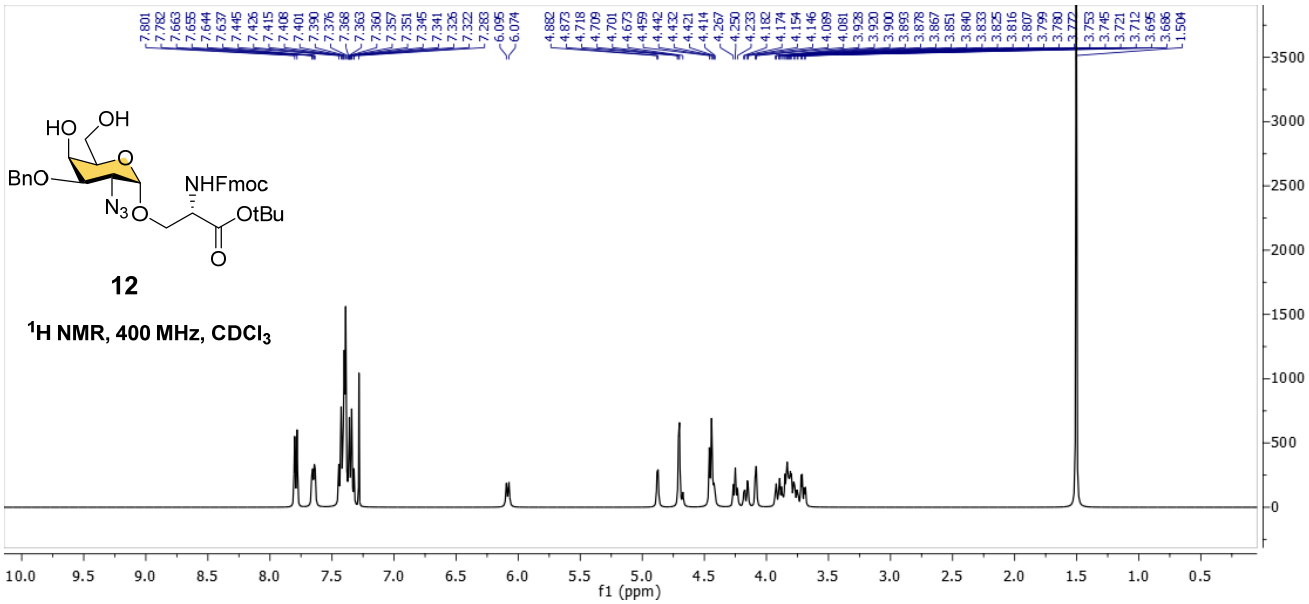

HO
OH
HO
HO
HO
OH
O
O
AcHN
NHFmoc
O
OtBu
O
9
13C NMR, 150 MHz, MeOD
f1 (ppm)

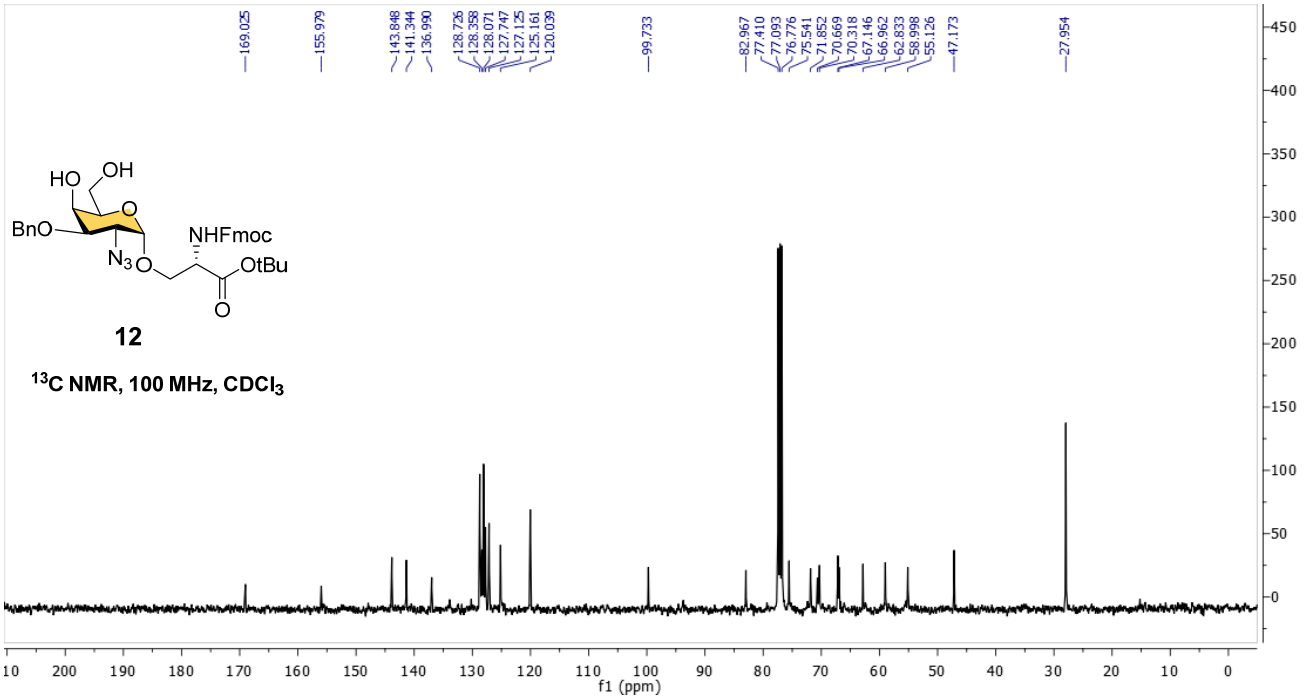

HO
OH
O
HO
OH
HO
O
AcHN
O
AcHN
NHFmoc
O
OH
O
10
^{1}H NMR, 600 MHz, MeOD
core 5
f1 (ppm)

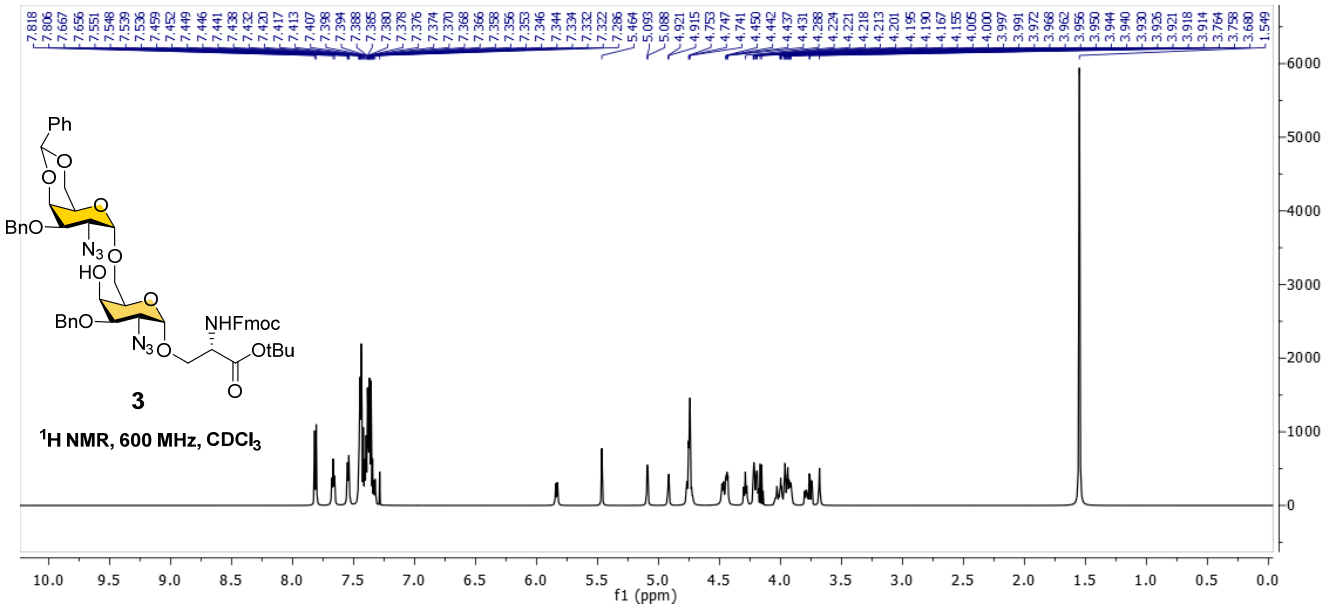

HO
OH
O
HO
OH
HO
O
AcHN
O
AcHN
NHFmoc
O
OH
O
10
^{13}C NMR, 150 MHz, MeOD
core 5
f1 (ppm)

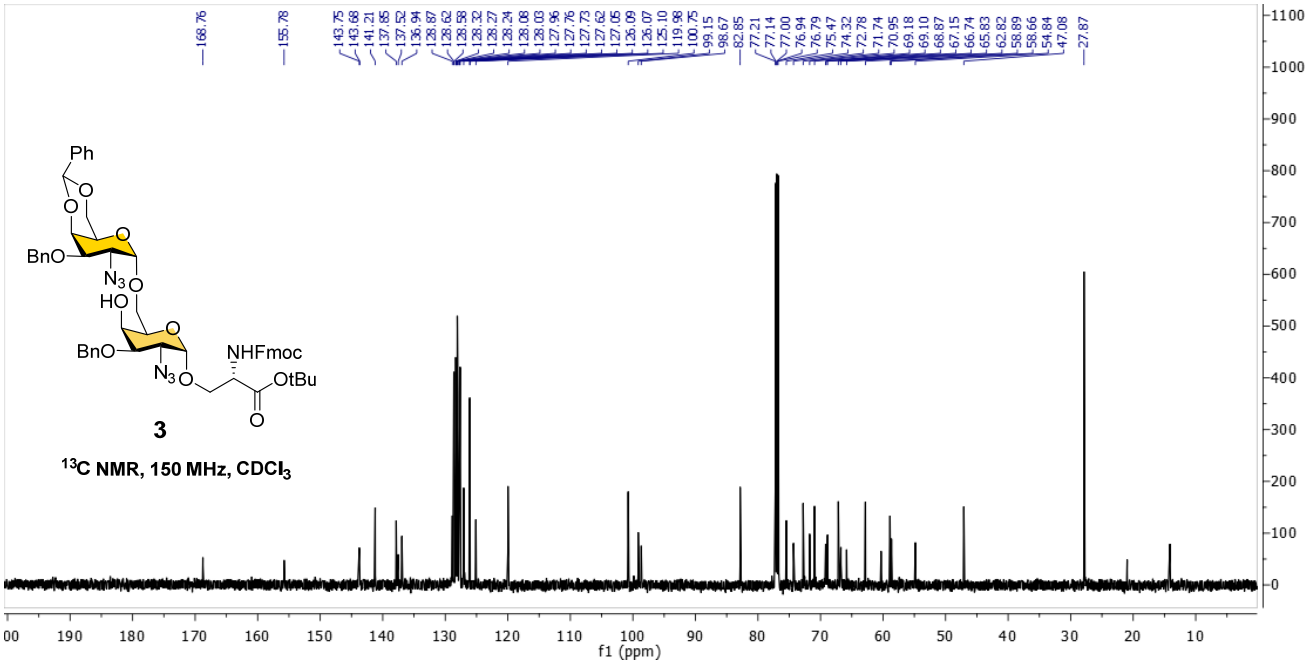

HO
OH
O
HO
OH
HO
HO
O
O
AcHN
NHFmoc
O
OH
O
11
1H NMR, 600 MHz, MeOD
core 8
f1 (ppm)

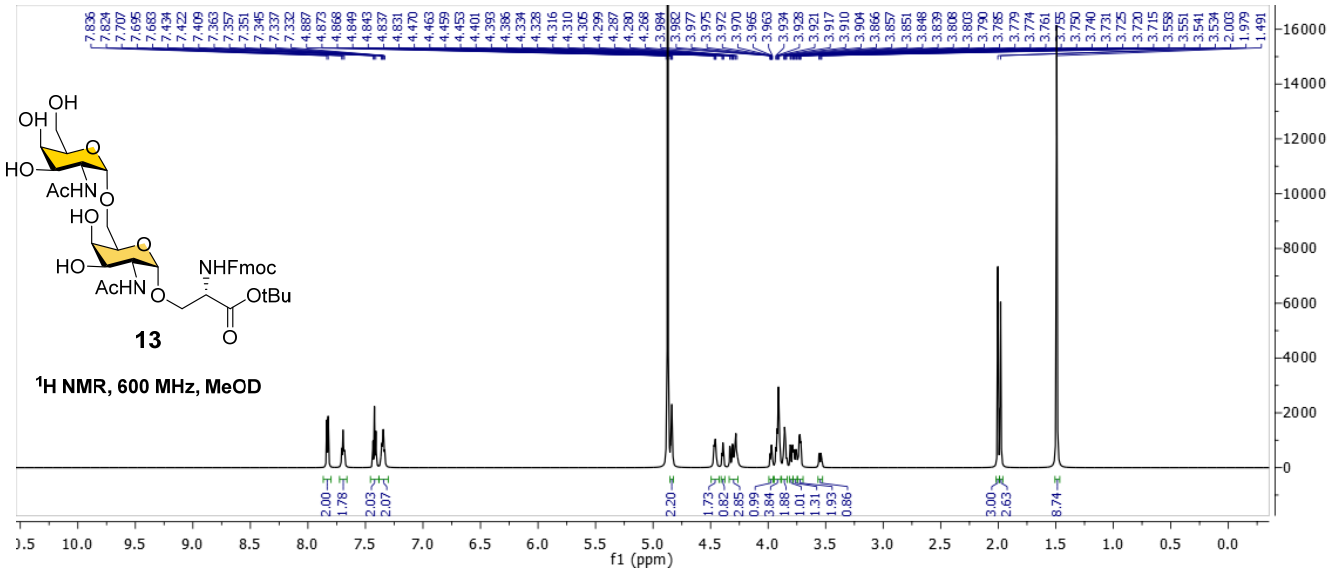

HO
OH
O
HO
OH
HO
HO
O
O
AcHN
NHFmoc
O
OH
O
11
13C NMR, 150 MHz, MeOD
core 8
f1 (ppm)

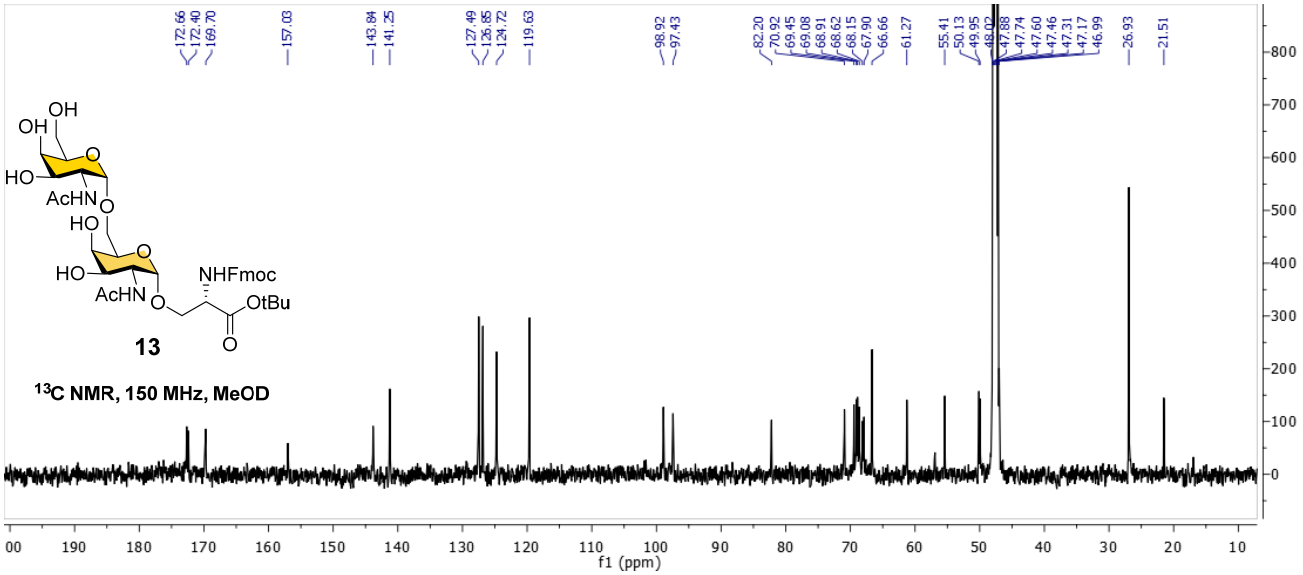

HO
OH
O
BnO
NHFmoc
N3
O
OtBu
O
12
1H NMR, 400 MHz, CDCl3
f1 (ppm)

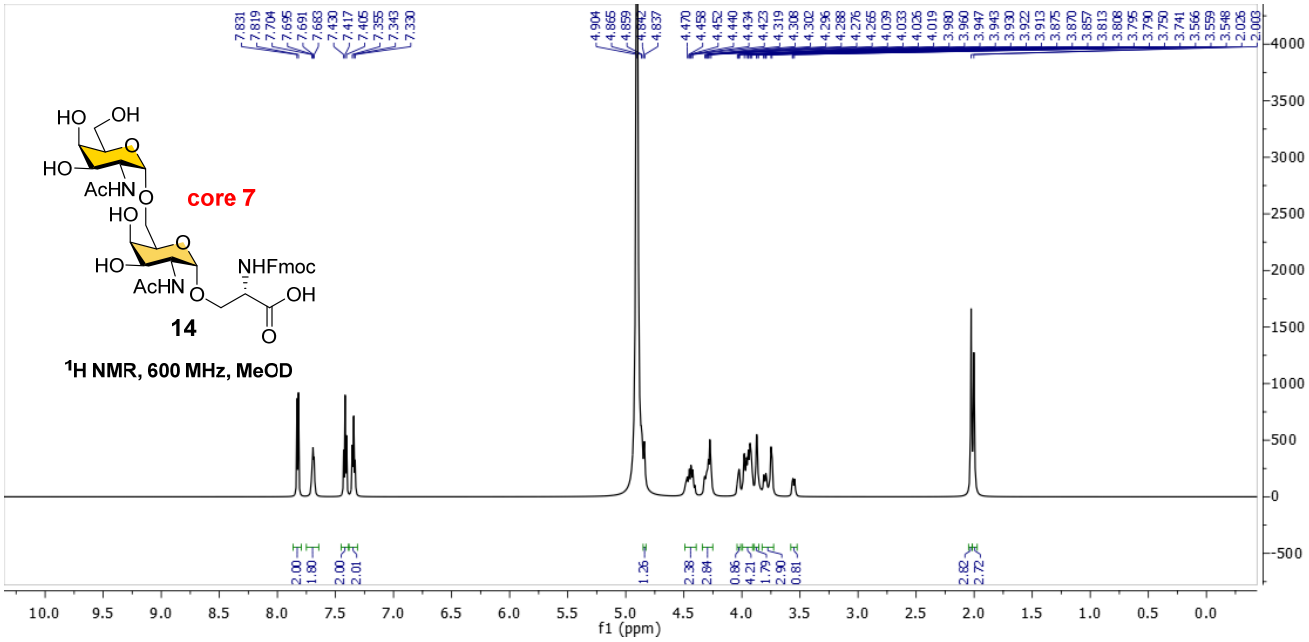

HO
OH
O
BnO
NHFmoc
N3
O
OtBu
O
12
13C NMR, 100 MHz, CDCl3
f1 (ppm)

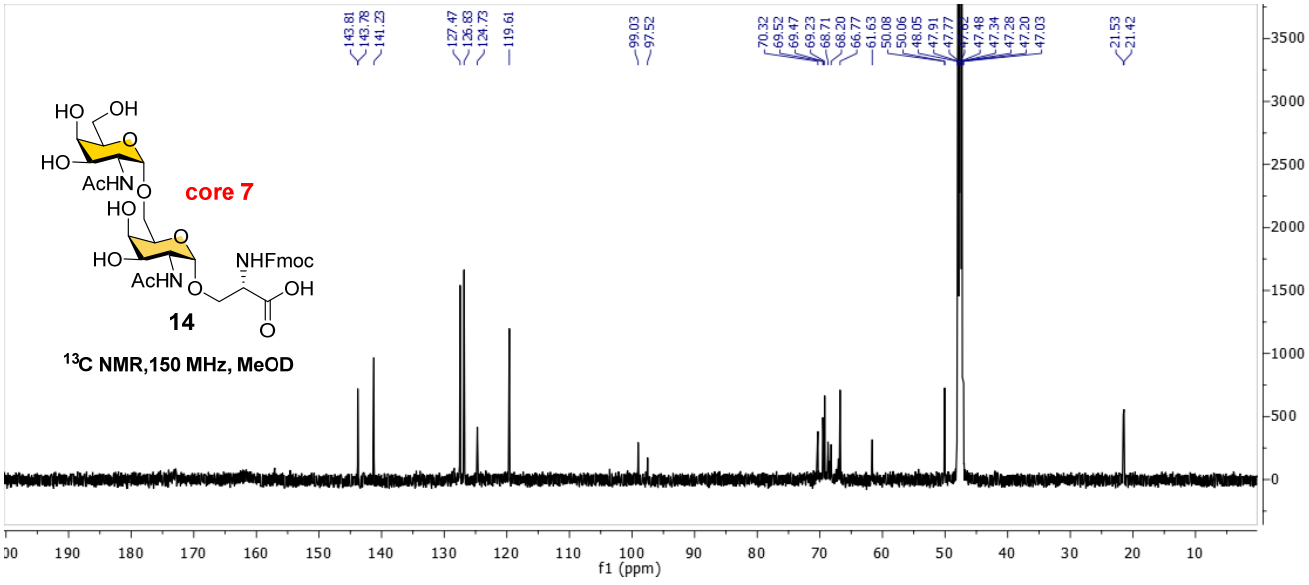

Ph
BnO
N3
HO
BnO
NHFmoc
OtBu
3
1H NMR, 600 MHz, CDCl3
f1 (ppm)

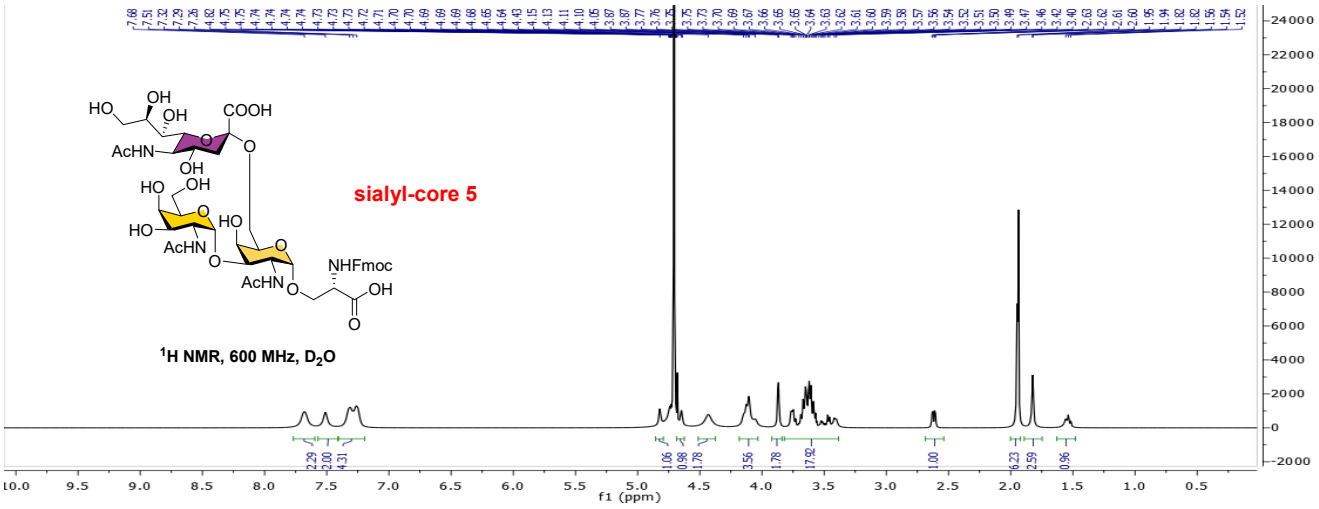

Ph
BnO
N3
HO
BnO
NHFmoc
OtBu
3
13C NMR, 150 MHz, CDCl3
f1 (ppm)

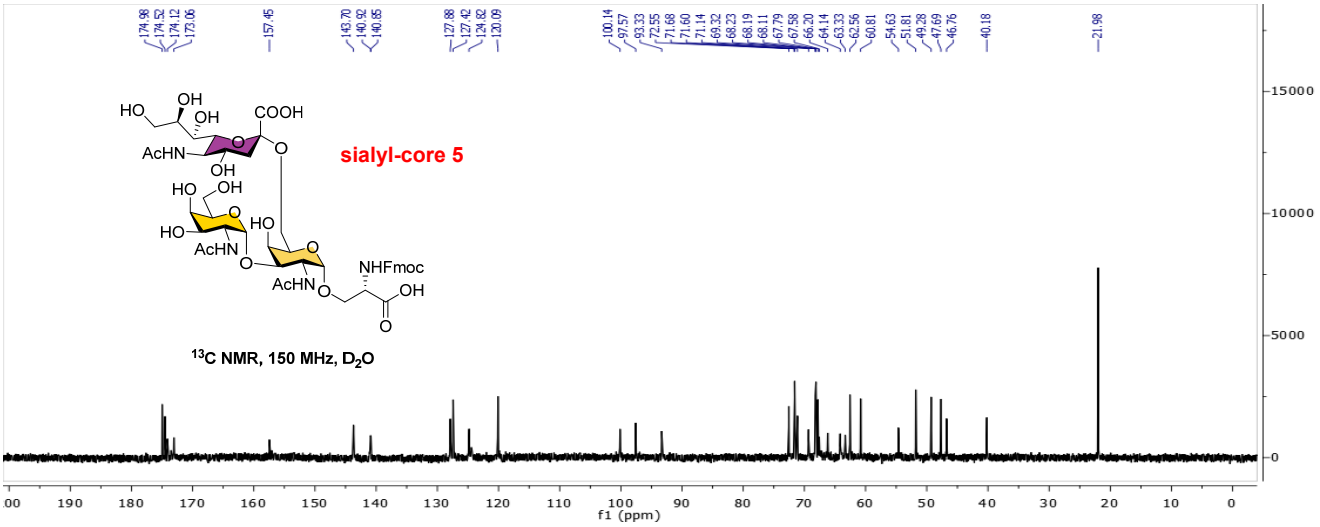

OH
OH
HO
AcHN
HO
HO
AcHN
NHFmoc
OtBu
13
^{1}H NMR, 600 MHz, MeOD
f1 (ppm)

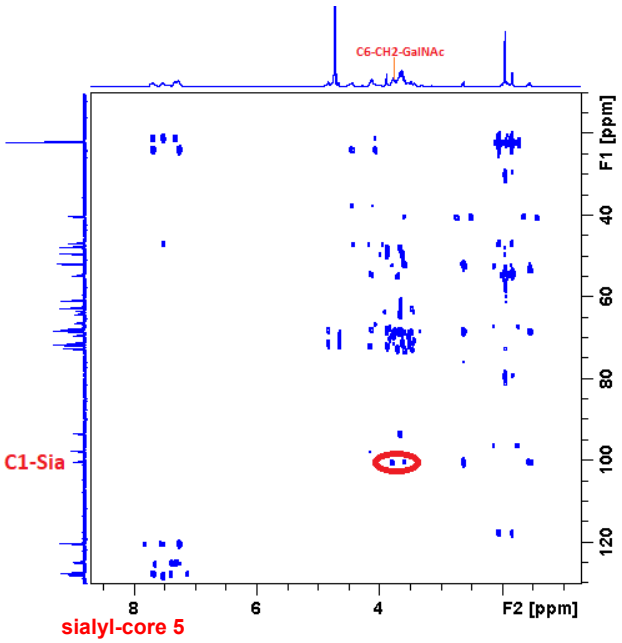

OH
OH
HO
AcHN
HO
HO
AcHN
NHFmoc
OtBu
13
^{13}C NMR, 150 MHz, MeOD
f1 (ppm)

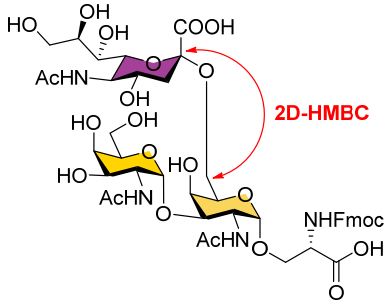

core 7
14
^{1}H NMR, 600 MHz, MeOD
f1 (ppm)

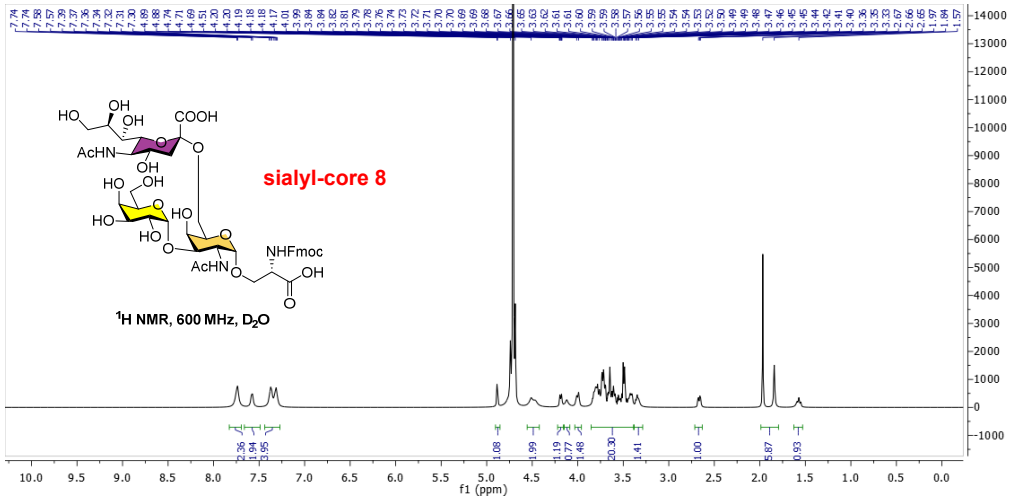

core 7
14
^{13}C NMR,150 MHz, MeOD
f1 (ppm)

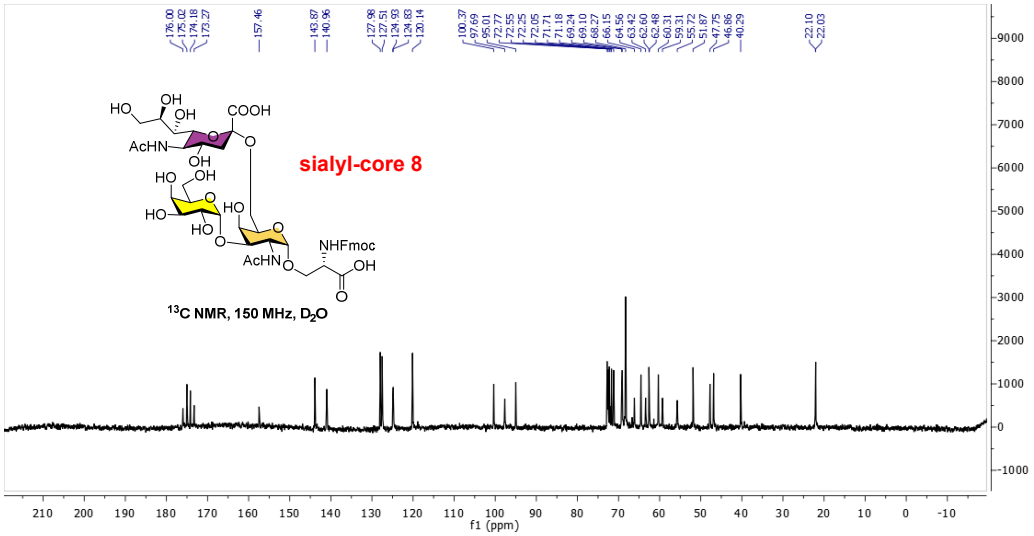

sialyl-core 5
^{1}H NMR, 600 MHz, D_2O
f1 (ppm)

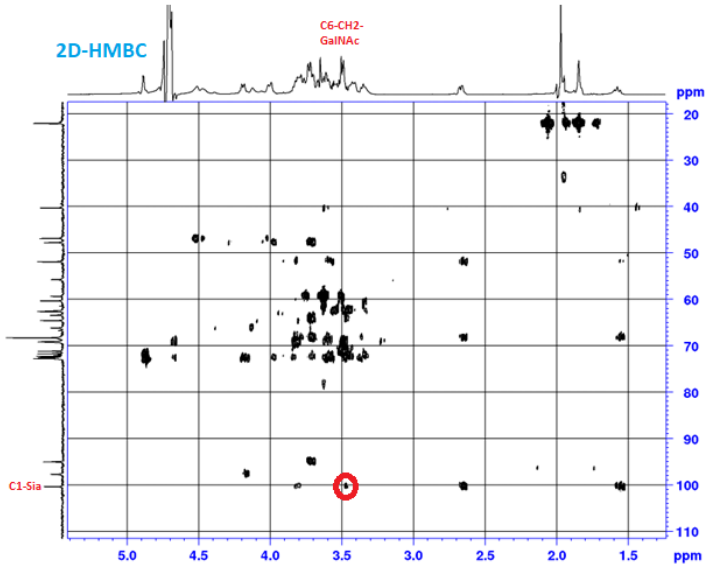

sialyl-core 5
^{13}C NMR, 150 MHz, D_2O
f1 (ppm)

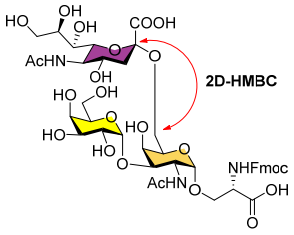

C6-CH2-GalNAc
C1-Sia
F1 [ppm]
F2 [ppm]
sialyl-core 5


2D-HMBC









sialyl-core 8

IX. References

1. H. Yu, H. Yu, R. Karpel and X. Chen, *Bioorg Med Chem*, 2004, **12**, 6427-6435.
2. H. Yu, S. Huang, H. Chokhawala, M. Sun, H. Zheng and X. Chen, *Angew Chem Int Ed Engl*, 2006, **45**, 3938-3944.
3. H. Yu, H. A. Chokhawala, S. Huang and X. Chen, *Nat Protoc*, 2006, **1**, 2485-2492.
4. S. Wang, C. Chen, M. R. Gadi, V. Saikam, D. Liu, H. Zhu, R. Bollag, K. Liu, X. Chen, F. Wang, P. G. Wang, P. Ling, W. Guan and L. Li, *Nature communications*, 2021, **12**, 3573.
5. S. Einarsson, B. Josefsson and S. Lagerkvist, *J Chromatogr A*, 1983, **282**, 609-618.
6. Z. Wu, Y. Liu, L. Li, X. F. Wan, H. Zhu, Y. Guo, M. Wei, W. Guan and P. G. Wang, *Org Biomol Chem*, 2017, **15**, 8946-8951.
7. X. Wang, Z. Zhong, L. Balmer and W. Wang, *Adv Exp Med Biol*, 2021, **1325**, 321-339.