Efficient chemoenzymatic synthesis of novel galacto-N-biose derivatives and their

sialylated forms

Supporting Information

Lei Li,*^{a,b} Yonghui Liu,^b Tiehai Li,^a Wenjun Wang,^b Zaikuan Yu,^a Cheng Ma,^a Jingyao Qu,^a Wei Zhao,^b Xi Chen,^c and Peng G Wang*^a

^aDepartment of Chemistry and Center for Diagnostics & Therapeutics, Georgia State University, Atlanta, GA 30303, USA

^bState Key Laboratory of Medicinal Chemical Biology and College of Pharmacy, Nankai University, Tianjin 300071, PR China

^cDepartment of Chemistry, University of California, Davis, One Shields Avenue, Davis, CA 95616, USA

*Corresponding author. E-mail: <u>lli22@gsu.edu</u> or <u>pwang11@gsu.edu</u>

Table of Contents

1.	Materials and general methods	S2
2.	Substrate Specificity assay of BiGalK	S2
3.	Enzymatic synthesis of Gal-1-P derivatives	S3
4.	One-pot two-enzyme synthesis of GNB, lacto- <i>N</i> -biose and derivatives	S3
5.	One-pot two-enzyme synthesis of sialyl GNB and derivatives	S4
6.	One-pot two-enzyme synthesis of di-sialyl GNB derivatives	S4
7.	HRMS and NMR data of Gal-1-P derivatives	S5
8.	HRMS and NMR data of GNB, lacto- <i>N</i> -biose, and derivatives	S6
9.	HRMS and NMR data of sialyl GNB and derivatives	S9
10.	HRMS and NMR data of di-sialyl GNB derivatives	S 11
11.	NMR spectra of synthesized compounds	.S12
12.	References	. S 31

1. Materials and general methods

Bifidobacterium infantis ATCC15697 D-galactosyl- β 1–3-*N*-acetyl-D-hexosamine phosphorylase (BiGalHexNAcP),¹ *B. infantis* ATCC15697 galactokinase (BiGalK),² *Neisseria meningitidis* CMP-Sia synthetase (NmCSS),³ *Pasteurella multocida* α 2,3-sialyltransferase 1 mutant with reduced sialidase activity (PmST1 E271F/313Y),⁴ *Photobacterium damselae* α 2–6-sialyltransferase (Pd2,6ST)⁵ were expressed and purified as previously reported. Galactose (Gal), 2-deoxy-galactose (Gal2deoxy), galactosamine (GalNH₂), *N*-acetyl-glucosamine (GlcNAc), *N*-acetyl-galactosamine (GalNAc), 6-deoxy-galactose (Gal6deoxy, also known as D-Fucose), 6-azido-6-deoxy-galactose (Gal6N₃), galacturonic acid (GalA), and D-talose (Tal) were purchased from Sigma. Cytidine 5'-triphosphate (CTP), *N*-acetylneuraminic acid (Neu5Ac), 2-azido-2-deoxy-galactose (GalN₃) and 2-fluoro-2-deoxy-galactose (GalF) were purchased from Carbosynth.

2. Substrate Specificity Assay of BiGalK

BiGalK reactions were carried out at 37 °C in a total volume of 100 μ L in Tris-HCl (100 mM, pH 6.5) containing Gal or derivatives (20 mM), ATP (22 mM), MgCl₂ (10 mM), and BiGalK (5 μ g or 50 μ g). These conditions were similar to those used for later preparative-scale synthesis. After 2 hr and 12 hr, an aliquot of 30 μ L was withdrawn from each reaction mixture and added into 30 μ L ethanol to quench reactions. The samples were stored at -20 °C until being analyzed by Thin-Layer Chromatography (TLC) and **2,4-dinitrosalicyclic acid** (**DNS**) method. For TLC analysis, 0.25 μ L of each sample was directly spotted on silica gel plates, developed by isopropanol/NH₄OH/H₂O = 7:3:2 (v/v/v), and stained with *p*-anisaldehyde sugar stain. DNS analysis was carried out as previously described using 40 μ L of each sample.⁶ Specific activity was also determined as previously reported in a pH 8.0 buffer.²

Substrate		BiGalK Specific Activity (µmol min ⁻¹ mg ⁻¹	Specific Activity	Duoduot	Conversion (%)		
			(µmol min ⁻¹ mg ⁻¹)		rroullet	2 h	12 h
1	Gal	5 µg	158 ²	31	Gal-1-P	95.2	NA
2	Gal2deoxy	5 µg	99.7 ²	32	Gal2deoxy-1-P	96.0	NA
3	GalNH ₂	5 µg	12.5 ²	33	Gal <mark>NH</mark> 2-1-P	65.3	93.1
4	GalN ₃	5 µg	3.26	34	Gal <mark>N</mark> 3-1-P	20.3	63.2
5	GalF	5 µg	76.3	35	Gal <mark>F</mark> -1-P	90.0	NA
6	Tal	5 µg	4.67	36	Tal-1-P	30.6	82.5
7	Gal <mark>6deoxy</mark>	50 µg	0.0781 ²	37	Gal <mark>6deoxy</mark> -1-P	13.4	60.4
8	GalA	50 µg	0.0906 ²	38	GalA-1-P	5.30	53.1
9	Gal <mark>6N</mark> 3	50 µg	0.201	39	Gal <mark>6N₃-1-P</mark>	34.8	75.9

Table S1. Yields of the BiGalK catalyzed reactions monitored by DNS assays. NA, not assayed

Figure S1. TLC analysis of BiGalK catalyzed reactions. Lanes with numbers: reactions with corresponding monosaccharide at 0 hr; Lanes after numbered lanes: reactions at 2 hr and 12 hr. Black vertical arrows point to monosaccharides and derivatives, red horizontal arrows point to corresponding sugar 1-phosphate product.



3. Enzymatic synthesis of Gal-1-P derivatives

To double confirm the activity of BiGalK towards various Gal derivatives, we performed mg scale synthesis of GalN₃-1-P (**34**), GalF-1-P (**35**), Tal-1-P (**36**) and Gal6N₃-1-P (**39**). Briefly, a 5 mL reaction system contains Tris-HCl (100 mM, pH 6.5), Gal derivatives (20 mM), ATP (22 mM), MgCl₂ (10 mM), and BiGalK (0.5 mg for **35**, 1 mg for **34** and **35**, 2.5 mg for **39**). Reactions were carried out at 37 °C and monitored by TLC until 80% of Gal derivatives were converted. Target proteins were purified as previously reported² and subjected to MS and NMR analysis to confirm structures. The synthesis of Gal-1-P (**31**), 2-deoxy-Gal-1-P (**32**), GalNH₂-1-P (**33**), 6-deoxy-Gal-1-P (**37**) and GalA-1-P (**38**) using BiGalK had be performed previously.²

4. One-pot two-enzyme synthesis of GNB, Lacto-N-biose and derivatives

For the synthesis of each β 1–3-galactoside derivative, an 8 mL reaction system contains: 100 mM Tris-HCl (pH 6.5), one of the Gal derivatives (20 mM, 26–33 mg), ATP (22 mM), GlcNAc or GalNAc (22 mM), MgCl₂ (15 mM), BiGalHexNAcP (1.5 mg) and excess amounts of BiGalK (0.1 mg for **1**, **2**, **5**; 0.5 mg for **3**, **4**, **6**; 1.5 mg for **9**; and 2.5 mg for **7**, **8**). Reactions were allowed to incubate at 37 °C for 24 hours with brief agitation. The product formation was monitored by TLC developed with isopropanol/NH₄OH/H₂O = 7:3:2 (v/v/v) and stained with *p*-anisaldehyde sugar stain. The reaction was quenched by adding the same volume of ice-cold EtOH and incubating at 4 °C for 30 min. The mixture was then centrifuged to remove precipitates. The supernatant was concentrated and purified by Bio-gel P-2 gel chromatography.

Donors	Acceptors	Products	Yields (%)
но он но Но Но 1 Gal		HO OH OH HO HO COH HO ACHN 41 Gal β 1–3GlcNAc	93
но он но он 2 Gal2deoxy		HO OH HO OH ACNH 42 Gal2deoxyβ1-3GlcNAc	76
HO OH HO HO HO H2N 3 GalNH2		HO OH OH HO HO HO HO H ₂ N AcHN 43 GalNH ₂ β1–3GlcNAc	ND
но_он ноо,,,,он 4 GalN ₃		HO OH OH HO HO HO HO N ₃ AcHN 44 GalN ₃ β 1–3GlcNAc	ND
но он но Гелион 5 GalF	но Сон но Асни 40 GicNAc	HO OH OH HO F ACHN 45 GalFβ1-3GlcNAc	26
но ^{ОН} но СОД 6 Tal		но он но но н	ND
носн _а но но но 7 Gal6deoxy		HO _{CH3} HO _{HO} HO _{AcHN} OH 47 Gal6deoxyβ1–3GlcNAc	83
HOCO2H HO HO B GalA	HOCOH HO HO 8 GalA	но <u>со</u> дн он но но но то но асни 48 GalAβ1–3GlcNAc	85
но _{N3} но но 9 Gal <mark>6N₃</mark>		$\begin{array}{c} HO & N_{3} \\ HO & HO & HO \\ HO & ACHN \end{array} OH \\ \begin{array}{c} HO & ACHN \\ \end{array} OH \\ \end{array} OH \\ \begin{array}{c} HO & ACHN \\ \end{array} OH \\ \end{array} OH \\ \begin{array}{c} HO & ACHN \\ \end{array} OH \\ \end{array} OH \\ \begin{array}{c} HO & ACHN \\ \end{array} OH \\ \end{array} OH \\ \begin{array}{c} HO & ACHN \\ \end{array} OH \\ OH \\$	68

 Table S2. Synthesis of Lacto-N-biose and derivatives via one-pot two-enzyme system shown in Scheme 1.

 GlcNAc was used as an acceptor instead of GalNAc. ND. not detected.

5. One-pot two-enzyme synthesis of sialyl GNB and derivatives

Reaction systems (2 mL) contains: 100 mM Tris-HCl (pH 8.0), one of the GNB or derivatives (20 mM, compounds **12**, **13**, **16**, **18**, **19**, **20**), CTP (22 mM), Neu5Ac (22 mM), MgCl₂ (15 mM), NmCSS (0.1 mg/mL), and PmST1 E271F/313Y (50 μ g/mL). Reactions were allowed to incubate at 37 °C for 4 hours with brief agitation. The product formation was monitored by TLC developed with isopropanol/NH₄OH/H₂O = 7:3:2 (v/v/v) and stained with *p*-anisaldehyde sugar stain. The reaction was quenched by adding the same volume of ice-cold EtOH and incubating at 4 °C for 30 min. The mixture was then centrifuged to remove precipitates. The supernatant was concentrated and purified by Bio-gel P-2 gel chromatography.

6. One-pot two-enzyme synthesis of di-sialyl GNB derivatives

Reaction systems (0.2–1 mL) contains: 100 mM Tris-HCl (pH 8.0), one of the sialyl GNB derivatives (20 mM, compounds **25**, **26**, **27**), CTP (22 mM), Neu5Ac (22 mM), MgCl₂ (15 mM), NmCSS (0.1 mg/mL), and Pd2,6ST (0.1 mg/mL). Reactions were allowed to incubate at 37 °C for overnight with brief agitation. The product formation was monitored by TLC developed with isopropanol/NH₄OH/H₂O = 7:3:2 (v/v/v) and stained with *p*-anisaldehyde sugar stain. The reaction was quenched by adding the same volume of ice-cold EtOH and incubating at 4 °C for 30 min. The mixture was then centrifuged to remove precipitates. The supernatant was concentrated and purified by Bio-gel P-2 gel chromatography.

7. HRMS and NMR data of Gal-1-P derivatives

NMR spectra were recorded in D_2O with a Bruker AV 400 spectrometer ¹H NMR at 400 MHz or 600MHz, ¹³C NMR at 100 MHz, ³¹P NMR at 162 MHz, ¹⁹F NMR at 376 MHz). Chemical shifts were reported in δ (ppm) from an internal standard of TMS (δ 0.00). Coupling constants are reported in hertz. High-resolution electrospray-ionization mass spectra (HRESIMS) were obtained on a Thermo Orbitrap elite.

GalN₃-1-P (30 mg), 42

¹H NMR (400 MHz, D₂O) δ 5.57 (dd, J = 3.6, 7.6 Hz, 1H), 4.15 (t, J = 6.4 Hz, 1H), 4.08 (dd, J = 3.2, 10.8 Hz, 1H), 4.02 (d, J = 2.8 Hz, 1H), 3.74–3.70 (m, 2H), 3.52 (ddd, J = 2.8, 5.2, 10.8 Hz, 1H); ¹³C NMR (100 MHz, D₂O) δ 92.69 (d, J_{C-P} = 5.5 Hz), 70.46, 67.87, 66.28, 60.15, 58.84 (d, J_{C-P} = 7.7 Hz); ¹³P NMR (162 MHz, D₂O) δ -0.13;

HRESIMS calcd for $C_6H_{11}N_3O_8P(M-H^+)^-m/z$ 284.0289, found m/z 284.0286.

GalF-1-P (31 mg), 43

¹H NMR (400 MHz, D₂O) δ 5.65 (dd, J = 3.6, 8.0 Hz, 1H), 4.70 (dddd, J = 1.6, 3.6, 11.6, 55.2 Hz, 1H), 4.19-4.15 (m, 2H), 4.05 (t, J = 3.6 Hz, 1H), 3.77–3.68 (m, 2H); ¹³C NMR (100 MHz, D₂O) δ 90.53 (dd, J = 4.9, 22.0 Hz), 87.46 (dd, J = 6.9, 183.1 Hz), 70.19, 68.83 (d, J = 8.2 Hz), 66.67 (d, J = 17.3 Hz), 60.05; ³¹P NMR (162 MHz, D₂O) δ 1.06; ¹⁹F NMR (376 MHz, D₂O) δ -207.63;

HRESIMS calcd for $C_6H_{11}FO_8P(M-H^+)$ m/z 261.0181, found m/z 261.0176.

Tal-1-P (28 mg), 44

¹H NMR (400 MHz, D₂O) δ 5.43 (dd, *J* = 1.6, 8.4 Hz, 1H), 4.10 (dd, *J* = 4.4, 8.0 Hz, 1H), 3.97 (t, *J* = 3.2 Hz, 1H), 3.91–3.86 (m, 2H); 3.80 (dd, *J* = 8.0, 12.0 Hz, 1H), 3.72 (dd, *J* = 4.0, 12.0 Hz, 1H); ¹³C NMR (100 MHz, D₂O) δ 94.71 (d, *J*_{*C-P*} = 4.8 Hz), 70.98, 69.83 (d, *J*_{*C-P*} = 7.3 Hz), 68.59, 63.77, 60.60; ¹³P NMR (162 MHz, D₂O) δ 0.64;

HRESIMS calcd for $C_6H_{12}O_9P(M-H^+)^-$ m/z 259.0224, found m/z 259.0219.

Gal6N₃-1-P (34 mg), 47

¹H NMR (400 MHz, D₂O) δ 5.47 (dd, J = 3.6, 7.2 Hz, 1H), 4.19 (t, J = 6.8 Hz, 1H), 3.97 (d, J = 2.8 Hz, 1H), 3.88 (dd, J = 3.2, 10.4 Hz, 1H), 3.77–3.71 (m, 1H), 3.56–3.44 (m, 2H); ¹³C NMR (100 MHz, D₂O) δ 93.44 (d, J_{C-P} = 5.8 Hz), 68.54, 68.33, 68.16, 67.46 (d, J_{C-P} = 7.1 Hz), 49.54; ¹³P NMR (162 MHz, D₂O) δ 0.45; HRESIMS calcd for C₆H₁₁N₃O₈P (M-H⁺)⁻ m/z 284.0289, found m/z 284.0284.

8. HRMS and NMR data of GNB, Lacto-N-biose and derivatives

NMR spectra were recorded in D_2O with a Bruker AV 400 spectrometer (¹H NMR at 400 MHz, ¹³C NMR at 100 MHz, ¹⁹F NMR at 376 MHz). Chemical shifts were reported in δ (ppm) from an internal standard of TMS (δ 0.00). Coupling constants are reported in hertz. High-resolution electrospray-ionization mass spectra (HRESIMS) were obtained on a Varian QFT-ESI mass spectrometer.

Galβ1-3GalNAc (61.7 mg), 12

¹H NMR (400 MHz, D₂O) δ 5.08 (d, *J* = 3.6 Hz, 1H), 4.54 (d, *J* = 8.5 Hz, 1H), 4.35 (d, *J* = 7.8 Hz, 1H), 4.29 (d, *J* = 7.8 Hz, 1H), 4.21–4.12 (m, 3H), 4.04 (d, *J* = 2.8 Hz, 2H), 3.99 (d, *J* = 6.9 Hz, 1H), 3.91–3.85 (m, 4H), 3.66 – 3.60 (m, 9H), 3.55–3.49 (m, 3H), 3.42–3.36 (m, 2H), 1.88 (s, 6H). ¹³C NMR (100 MHz, D₂O) δ 175.13, 174.60, 104.30, 104.08, 95.15, 91.15, 80.32, 77.27, 75.31, 74.93, 72.59, 70.23, 68.21, 67.54, 61.03, 52.41, 48.99, 23.22, 22.01.

HRESIMS calcd for $C_{14}H_{25}NO_{11}$ (M+ Na)⁺ m/z 406.1325, found m/z 406.1326.

$Gal2deoxy\beta1-3GalNAc (49.9 mg), 13$

¹H NMR (400 MHz, D₂O) δ 5.08 (d, J = 3.7 Hz, 1H), 4.60 (dd, J = 9.8, 2.1 Hz, 1H), 4.57–4.52 (m, 2H), 4.15–4.08 (m, 2H), 4.04–3.96 (m, 2H), 3.89 (dd, J = 11.2, 3.0 Hz, 1H), 3.82 (dd, J = 10.8, 8.4 Hz, 1H), 3.75–3.68 (m, 3H), 3.65–3.56 (m, 9H), 3.44–3.38 (m, 2H), 1.91 (s, 6H), 1.81–1.73 (m, 2H), 1.55 (td, J = 12.2, 9.9 Hz, 2H). ¹³C NMR (100 MHz, D₂O) δ 174.57, 174.36, 101.81, 95.00, 91.07, 79.91, 76.90, 75.21, 74.84, 70.19, 68.49, 67.81, 67.71, 66.64, 61.28, 61.11, 60.89, 52.49, 49.00, 33.58, 22.05, 21.84. HRESIMS calcd for C₁₄H₂₅NO₁₀ (M+ Na)⁺ m/z 390.1376, found m/z 390.1380.

GalF_β1–3GalNAc (18.3 mg), 16

¹H NMR (400 MHz, D₂O): δ 5.21 (d, 0.6 H, J = 3.6 Hz, H'-1α), 4.72 (d, 0.4 H, J = 8.4 Hz, H'-1β), 4.70 (t, 0.6 H, J = 8.4 Hz), 4.40 (t, 0.4 H, J = 8.4 Hz), 4.33 (dd, 0.4 H, J = 3.6, 11.2 Hz), 4.27 (t, 0.6 H, J = 8.0 Hz), 4.23 (d, 0.4 H, J = 2.8 Hz), 4.17 (d, 0.6 H, J = 3.2 Hz), 4.14–3.73 (m, 10.0 H), 2.02 (s, 3 H). ¹³C NMR (100 MHz, D₂O): δ 175.07, 174.78, 102.22 (J = 23.2 Hz), 102.11 (J = 23.0 Hz), 94.99, 91.33, 91.24 (J = 179.9 Hz), 80.76, 77.71, 75.19, 75.16, 74.88, 71.27 (J = 16.9 Hz), 71.26 (J = 17.2 Hz), 70.33, 69.22, 69.13, 68.66, 67.88, 61.47, 61.22, 61.00, 60.77, 52.31, 48.84, 21.18, 21.99. ¹⁹F NMR (376 MHz, D₂O): δ -207.25, -207.35. HRESIMS calcd for C₁₄H₂₄FNO₁₀ (M+ Na)⁺ m/z 408.1282, found m/z 408.1270.

Gal6deoxyβ1–3GalNAc (54.3 mg), 18

¹H NMR (400 MHz, D₂O) δ 5.10 (d, *J* = 3.7 Hz, 1H), 4.57 (d, *J* = 8.5 Hz, 1H), 4.35 (d, *J* = 7.9 Hz, 1H), 4.30 (d, *J* = 7.8 Hz, 1H), 4.17 (dd, *J* = 11.1, 3.7 Hz, 1H), 4.11 (d, *J* = 2.7 Hz, 1H), 4.04 (d, *J* = 3.1 Hz, 1H), 4.01 (d, *J* = 6.2 Hz, 1H), 3.89 (dd, *J* = 8.3, 2.8 Hz, 1H), 3.88–3.83 (m, 1H), 3.72 (dd, *J* = 10.9, 3.1 Hz, 1H), 3.69–3.57 (m, 9H), 3.53–3.47 (m, 2H), 3.37 (d, *J* = 7.8 Hz, 1H), 3.34 (d, *J* = 7.9 Hz, 1H), 1.91 (s, 6H), 1.14–1.09 (m, 6H). ¹³C NMR (100 MHz, D₂O) δ 174.93, 174.63, 104.54, 104.37, 95.14, 91.17, 79.86, 76.83, 74.85, 72.72, 71.22, 70.62, 70.42, 70.36, 70.21, 68.68, 68.01, 61.22, 61.01, 52.45, 48.98, 22.25, 22.02, 15.47. HRESIMS calcd for C₁₄H₂₅NO₁₀ (M+ Na)⁺ m/z 390.1376, found m/z 390.1375.

GalAβ1–3GalNAc (56.5 mg), 19

¹H NMR (400 MHz, D₂O) δ 5.04 (d, *J* = 3.5 Hz, 1H), 4.61 (d, *J* = 8.0 Hz, 1H), 4.33 (d, *J* = 7.8 Hz, 1H), 4.29 (d, *J* = 7.7 Hz, 1H), 3.93 (dd, *J* = 10.6, 3.5 Hz, 1H), 3.81–3.73 (m, 5H), 3.72–3.56 (m, 8H), 3.54–3.34 (m, 6H), 1.90

(s, 6H). ¹³C NMR (100 MHz, D₂O) δ 175.43, 175.18, 104.20, 104.07, 95.35, 91.66, 83.21, 80.76, 76.08, 75.91, 75.86, 73.17, 71.84, 71.34, 69.32, 69.17, 61.63, 61.35, 61.18, 56.24, 53.52, 22.86, 22.61. HRESIMS calcd for C₁₄H₂₃NO₁₂ (M+ Na)⁺ m/z 420.1118, found m/z 420.1117.

Gal6N₃β1-3GalNAc (52.4 mg), 20

¹H NMR (400 MHz, D₂O): δ 5.21 (d, 0.6 H, J = 4.0 Hz, H'-1α), 4.70 (d, 0.4 H, J = 8.4 Hz, H'-1β), 4.52 (d, 0.6 H, J = 7.6 Hz), 4.47 (d, 0.4 H, J = 7.6 Hz), 4.32 (dd, 0.4 H, J = 4.0 and 11.2 Hz), 4.23 (d, 0.6 H, J = 2.8 Hz), 4.17 (d, 0.4 H, J = 2.8 Hz), 4.12 (t, 0.6 H, J = 6.0 Hz), 4.03 (dd, 0.6 H, J = 2.8 and 10.8 Hz), 4.00 (t, 0.4 H, J = 10.8 Hz), 3.88–3.50 (m, 8.0 H), 3.29 (dd, 1.0 H, J = 3.2 Hz and 13.2 Hz), 2.03 (s, 3 H). ¹³C NMR (100 MHz, D₂O): δ 174.95, 174.67, 104.82, 104.70, 95.13, 91.25, 79.82, 76.89, 74.87, 73.84, 72.29, 70.46, 70.41, 70.29, 69.05, 68.92, 68.21, 61.18, 60.96, 52.45, 51.02, 48.90, 22.26, 22.03. HRESIMS calcd for C₁₄H₂₄N₄O₁₀ (M+ Na)⁺ m/z 431.1390, found m/z 431.1378.

Galβ1–3GlcNAc (60.4 mg), 41

¹H NMR (400 MHz, D₂O) δ 5.05 (d, *J* = 3.5 Hz, 1H), 4.62 (d, *J* = 8.3 Hz, 1H), 4.32 (d, *J* = 7.8 Hz, 1H), 4.28 (d, *J* = 7.8 Hz, 1H), 4.11–4.06 (m, 2H), 3.97–3.93 (m, 2H), 3.92 (d, *J* = 3.5 Hz, 1H), 3.81–3.73 (m, 3H), 3.73–3.65 (m, 3H), 3.60 (d, *J* = 7.5 Hz, 9H), 3.57–3.53 (m, 1H), 3.51–3.38 (m, 3H), 1.90 (s, 6H). ¹³C NMR (100 MHz, D₂O) δ 174.77, 174.48, 103.14, 103.01, 94.64, 90.96, 83.15, 80.61, 75.53, 75.29, 72.65, 71.27, 70.25, 70.08, 68.82, 61.20, 60.58, 52.83, 23.23, 22.20, 21.95.

HRESIMS calcd for $C_{14}H_{25}NO_{11}$ (M+ Na)⁺ m/z 406.1325, found m/z 406.1327.

Gal2deoxyβ1–3GlcNAc (47.4 mg), 42

¹H NMR (400 MHz, D₂O) δ 5.04 (d, J = 3.4 Hz, 1H), 4.63 (d, J = 8.2 Hz, 1H), 4.59–4.52 (m, 2H), 3.88 (dd, J = 10.6, 3.5 Hz, 1H), 3.82–3.58 (m, 15H), 3.48–3.43 (m, 2H), 3.43 – 3.35 (m, 2H), 1.93 (s, 6H), 1.76 (dd, J = 11.1, 3.6 Hz, 2H), 1.62–1.49 (m, 2H). ¹³C NMR (100 MHz, D₂O) δ 174.39, 174.20, 100.80, 94.37, 90.94, 80.46, 75.51, 71.26, 68.58, 67.69, 66.63, 61.36, 60.65, 60.49, 55.73, 52.89, 33.66, 22.01, 21.80. HRESIMS calcd for C₁₄H₂₅NO₁₀ (M+ Na)⁺ m/z 390.1376, found m/z 390.1377.

GalF_β1–3GlcNAc (17.0 mg), 45

¹H NMR (400 MHz, D₂O): δ 5.13 (d, 0.6 H, J = 3.6 Hz, H'-1α), 4.72 (d, 0.4 H, J = 8.4 Hz, H'-1β), 4.69 (dd, 0.6 H, J = 8.0, 3.6 Hz), 4.67 (dd, 0.4 H, J = 8.0, 4.0 Hz), 4.36 (t, 0.6 H, J = 8.0 Hz), 4.23 (t, 0.4 H, J = 8.0 Hz), 4.06-3.44 (m, 11.0 H), 1.98 (s, 3 H). ¹³C NMR (100 MHz, D₂O): δ 174.73, 174.52, 101.30 (d, J = 23.4 Hz), 101.12 (d, J = 23.5 Hz), 94.41, 91.33 (d, J = 179.7 Hz), 91.11, 83.31, 80.80, 75.49, 75.42, 75.36, 71.35, 71.26 (d, J = 17.8 Hz), 71.23 (d, J = 17.6 Hz), 69.11, 69.02, 68.52, 68.48, 61.31, 60.76, 60.65, 60.49, 55.46, 52.72, 22.01, 21.77. ¹⁹F NMR (376 MHz, D₂O): δ -207.43, -207.44.

HRESIMS calcd for $C_{14}H_{24}FNO_{10}$ (M+ Na)⁺ m/z 408.1282, found m/z 408.1278.

Gal6deoxyβ1-3GlcNAc (51.8 mg), 47

¹H NMR (400 MHz, D₂O) δ 5.10 (d, *J* = 3.7 Hz, 1H), 4.57 (d, *J* = 8.5 Hz, 1H), 4.35 (d, *J* = 7.9 Hz, 1H), 4.30 (d, *J* = 7.8 Hz, 1H), 4.17 (dd, *J* = 11.1, 3.7 Hz, 1H), 4.11 (d, *J* = 2.7 Hz, 1H), 4.04 (d, *J* = 3.1 Hz, 1H), 4.01 (d, *J* = 6.2 Hz, 1H), 3.89 (dd, *J* = 8.3, 2.8 Hz, 1H), 3.87–3.83 (m, 1H), 3.72 (dd, *J* = 10.9, 3.1 Hz, 1H), 3.67–3.59 (m, 7H), 3.53–3.47 (m, 2H), 3.37 (d, *J* = 7.8 Hz, 1H), 3.34 (d, *J* = 7.9 Hz, 1H), 1.91 (s, 6H), 1.12 (d, *J* = 2.2 Hz, 3H), 1.11 (d, *J* = 2.3 Hz, 3H). ¹³C NMR (100 MHz, D₂O) δ 174.93, 174.63, 104.54, 104.37, 95.14, 91.17, 79.86,

76.83, 74.85, 72.72, 71.22, 70.62, 70.42, 70.36, 70.21, 68.68, 68.01, 61.22, 61.01, 52.45, 48.98, 22.25, 22.02, 15.47.

HRESIMS calcd for $C_{14}H_{25}NO_{10}$ (M+ Na)⁺ m/z 390.1376, found m/z 390.1377.

GalA_β1–3GlcNAc (57.1 mg), 48

¹H NMR (400 MHz, D₂O) δ 5.05 (d, *J* = 3.5 Hz, 0.5H), 4.62 (d, *J* = 8.3 Hz, 0.5H), 4.34–4.26 (m, 1H), 4.08 (d, *J* = 3.3 Hz, 1H), 3.98–3.90 (m, 1.5H), 3.82–3.60 (m, 4.5H), 3.58–3.53 (m, 0.5H), 3.51–3.34 (m, 2.5H), 1.90 (s, 3H); ¹³C NMR (100 MHz, D₂O) δ 174.48, 174.42, 103.14, 103.01, 94.64, 90.96, 83.15, 80.61, 75.53, 75.29, 72.65, 72.62, 71.27, 70.25, 70.08, 70.04, 68.82, 68.77, 61.20, 60.58, 55.51, 52.83, 23.23, 21.95. HRESIMS calcd for C₁₄H₂₃NO₁₂ (M+ Na)⁺ m/z 420.1118, found m/z 420.1115.

Gal6N₃β1-3GlcNAc (46.8 mg), 49

¹H NMR (400 MHz, D₂O): δ 5.16 (d, 0.6 H, J = 3.6 Hz, H'-1α), 4.74 (d, 0.4 H, J = 8.0 Hz, H'-1β), 4.47 (d, 0.6 H, J = 8.0 Hz), 4.43 (d, 0.4 H, J = 8.0 Hz), 4.07–3.48 (m, 12H), 2.02 (s, 3 H). ¹³C NMR (100 MHz, D₂O): δ 103.61, 103.49, 94.67, 91.01, 83.01, 80.71, 75.36, 73.21, 73.16, 72.28, 72.35, 71.13, 70.52, 68.84, 68.75, 61.21, 60.74, 60.58, 55.57, 52.80, 50.90, 22.29, 22.04.

HRESIMS calcd for $C_{14}H_{24}N_4O_{10}$ (M+ Na)⁺ m/z 431.1390, found m/z 431.1385.

9. HRMS and NMR data of sialyl GNB and derivatives

NMR spectra were recorded in D_2O with a Bruker AV 400 spectrometer ¹H NMR at 400 MHz or 600MHz, ¹³C NMR at 150 HMz, ³¹P NMR at 162 MHz, ¹⁹F NMR at 376 MHz). Chemical shifts were reported in δ (ppm) from an internal standard of TMS (δ 0.00). Coupling constants are reported in hertz. High-resolution electrospray-ionization mass spectra (HRESIMS) were obtained on a Thermo Orbitrap elite.

Neu5Aca2-3Galβ1-3GalNAc (25.3 mg), 22

¹H NMR (600 MHz, D₂O) δ 5.24 (d, *J* = 3.6 Hz, 0.5H), 4.70 (d, *J* = 8.4 Hz, 0.5H), 4.58 (d, *J* = 7.8 Hz, 0.5H), 4.52 (d, *J* = 7.8 Hz, 0.5H), 4.29 (d, *J* = 3.6, 11.4 Hz, 0.5H), 4.26 (d, *J* = 2.4 Hz, 0.5H), 4.18 (d, *J* = 3.0 Hz, 0.5H), 4.14 (d, *J* = 6.0 Hz, 0.5H), 4.10–3.99 (m, 2H), 3.97–3.94 (m, 1H), 3.90–3.70 (m, 9H), 3.67–3.59 (m, 5H), 3.57–3.53 (m, 1H), 2.78–2.75 (m, 1H), 2.04 (s, 6H), 1.79 (t, *J* = 12.0 Hz, 1H); ¹³C NMR (150 MHz, D₂O) δ 174.99, 174.68, 173.90, 104.60, 104.43, 99.71, 95.20, 91.16, 80.24, 77.21, 75.69, 75.61, 74.83, 74.76, 72.80, 71.83, 70.20, 69.10, 69.02, 68.56, 68.37, 68.06, 67.92, 67.38, 62.51, 61.19, 60.97, 60.94, 52.33, 51.67, 48.98, 39.73, 22.33, 22.08, 22.04;

HRESIMS calcd for $C_{25}H_{41}N_2O_{19}$ (M-H⁺)⁻ m/z 673.2309, found m/z 673.2291.

Neu5Aca2-3Gal2deoxyB1-3GalNAc (21.6 mg), 23

¹H NMR (600 MHz, D₂O) δ 5.22 (d, J = 3.6 Hz, 0.5H), 4.75–4.68 (m, 2.1H), 4.32 (d, J = 3.6 Hz, 0.4H), 4.26–4.23 (m, 1H), 4.17–4.12 (m, 2H), 4.02 (dd, J = 3.0, 11.4 Hz, 0.5H), 3.95 (dd, J = 9.0, 11.4 Hz, 0.5H), 3.89–3.86 (m, 1H), 3.85–3.71 (m, 8H), 3.69–3.59 (m, 4H), 3.55–3.52 (m, 1H), 2.75 (dd, J = 4.2, 12.0 Hz, 1H), 2.07 (s, 3H), 2.04 (s, 3H), 2.00–1.96 (m, 1H), 1.83–1.76 (m, 1H), 1.71 (t, J = 12.0 Hz, 1H); ¹³C NMR (150 MHz, D₂O) δ 175.04, 174.64, 174.39, 173.79, 101.69, 101.66, 100.66, 95.13, 91.13, 79.93, 76.87, 75.20, 75.15, 74.88, 72.69, 72.11, 72.06, 71.10, 70.23, 68.54, 68.31, 68.01, 67.84, 65.89, 62.58, 61.34, 61.31, 61.20, 60.97, 51.73, 48.98, 40.38, 33.84, 22.23, 22.03, 22.01;

HRESIMS calcd for $C_{25}H_{41}N_2O_{18}$ (M-H⁺)⁻ m/z 657.2360, found m/z 657.2342.

Neu5Aca2-3GalFβ1-3GalNAc (23.5 mg), 24

¹H NMR (600 MHz, D₂O) δ 5.21 (d, *J* = 3.6 Hz, 0.5H), 4.70 (d, *J* = 8.4 Hz, 1H), 4.33-4.27 (m, 3H), 4.22 (d, *J* = 3.0 Hz, 0.5H), 4.20–4.13 (m, 2H), 4.10–4.01 (m, 2H), 3.96 (t, *J* = 3.0 Hz, 1H), 3.88–3.81 (m, 3H), 3.79–3.67 (m, 6H), 3.64 (dd, *J* = 6.0, 12.0 Hz, 1H), 3.61–3.56 (m, 2H), 2.70 (dd, *J* = 3.2, 12.0 Hz, 1H), 2.03 (s, 3H), 2.02 (s, 3H), 1.76 (t, *J* = 12.0 Hz, 1H); ¹³C NMR (150 MHz, D₂O) δ 175.08, 174.93, 174.68, 173.62, 102.38 (t, *J* = 23.1 Hz), 99.91, 96.16, 94.98, 91.32, 89.43, 82.97, 82.91, 80.87, 77.76, 74.92, 74.87, 74.24, 73.78 (d, *J* = 16.1 Hz), 72.75, 71.89, 70.31, 69.32, 68.67, 68.34, 68.18 (d, *J* = 8.6 Hz), 68.05, 67.88, 63.78, 63.75, 62.45, 62.45, 62.43, 61.16, 60.94, 60.73, 60.71, 52.25, 51.61, 48.77, 39.68, 22.19, 22.03, 22.00;

¹⁹F NMR (376 MHz, D_2O) δ -205.46, -205.56;

HRESIMS calcd for $C_{25}H_{40}FN_2O_{18}$ (M-H⁺)⁻ m/z 675.2266, found m/z 675.2247.

Neu5Aca2-3Gal6deoxyβ1-3GalNAc (22.3 mg), 25

¹H NMR (600 MHz, D₂O) δ 5.23 (d, *J* = 3.6 Hz, 0.5H), 4.69 (d, *J* = 8.4 Hz, 0.5H), 4.54 (d, *J* = 7.8 Hz, 0.5H), 4.48 (d, *J* = 7.8 Hz, 0.5H), 4.28 (dd, *J* = 3.6, 10.8 Hz, 0.5H), 4.23 (d, *J* = 2.4 Hz, 0.5H), 4.16–4.12 (m, 1H), 4.09–3.97 (m, 2H), 3.89-3.59 (m, 13H), 3.52–3.48 (m, 1H), 2.78-2.76 (m, 1H), 2.03 (s, 6H), 1.78 (t, *J* = 12.0 Hz, 1H), 1.22–1.21 (m, 3H); ¹³C NMR (150 MHz, D₂O) δ 174.99, 174.65, 173.89, 173.85, 104.31, 104.15, 99.60, 95.21, 91.15, 79.99, 76.98, 75.96, 75.90, 74.87, 72.77, 71.84, 70.47, 70.43, 70.23, 69.95, 68.81, 68.72, 68.54,

68.38, 68.04, 67.90, 62.49, 61.26, 61.04, 52.36, 51.68, 49.01, 39.86, 22.34, 22.09, 22.04, 15.50; HRESIMS calcd for $C_{25}H_{41}N_2O_{18}$ (M-H⁺)⁻ m/z 657.2360, found m/z 657.2341.

Neu5Aca2-3GalAβ1-3GalNAc (4.6 mg), 26

¹H NMR (600 MHz, D₂O) δ 5.25 (d, *J* = 3.6 Hz, 0.4H), 4.69 (d, *J* = 9.0 Hz, 1H), 4.58 (d, *J* = 7.8 Hz, 0.6H), 4.51 (d, *J* = 7.8 Hz, 0.4H), 4.40 (t, *J* = 5.4 Hz, 1H), 4.34–4.23 (m, 4H), 4.14-4.10 (m, 1.6H), 4.05–4.00 (m, 2H), 3.92–3.83 (m, 3H), 3.80-3.59 (m, 5H), 3.58-3.54 (m, 1H), 2.77–2.74 (m, 1H), 2.04 (s, 3H), 2.03 (s, 3H), 1.82 (t, *J* = 12.0 Hz, 1H); ¹³C NMR (150 MHz, D₂O) δ 174.95, 174.67, 173.92, 173.88, 104.04, 103.82, 99.98, 96.51, 91.15, 89.35, 82.67, 82.61, 80.51, 77.42, 75.80, 75.74, 75.07, 74.98, 74.26, 72.82, 71.78, 70.36, 69.23, 68.98, 68.77, 68.44, 68.11, 68.06, 67.42, 64.16, 64.12, 62.53, 61.29, 61.07, 52.33, 51.64, 49.04, 39.55, 22.32, 22.07, 22.04;

HRESIMS calcd for $C_{25}H_{39}N_2O_{20}$ (M-H⁺)⁻ m/z 687.2102, found m/z 687.2080.

Neu5Aca2-3Gal6N₃β1-3GalNAc (22.1 mg), 27

¹H NMR (600 MHz, D₂O) δ 5.23 (d, *J* = 3.6 Hz, 0.5H), 4.75-4.70 (m, 1.5H), 4.61 (d, *J* = 7.8 Hz, 0.6H), 4.55 (d, *J* = 7.8 Hz, 0.4H), 4.32 (dd, *J* = 3.6, 11.4 Hz, 0.5H), 4.25 (d, *J* = 3.0 Hz, 0.5H), 4.18 (d, *J* = 3.0 Hz, 0.4H), 4.13 (d, *J* = 6.0 Hz, 0.6H), 4.10–3.98 (m, 2H), 3.91-3.51 (m, 14H), 3.32–3.29 (m, 1H), 2.76 (dd, *J* = 4.8, 12.6 Hz, 1H), 2.04 (s, 3H), 2.03 (s, 3H), 1.79 (t, *J* = 12.0 Hz, 1H); ¹³C NMR (150 MHz, D₂O) δ 174.98, 174.71, 173.87, 104.66, 104.52, 99.69, 99.68, 95.16, 91.23, 80.07, 77.12, 75.40, 75.34, 74.88, 73.60, 72.79, 71.84, 70.30, 68.89, 68.81, 68.56, 68.37, 68.11, 68.05, 67.87, 63.28, 62.47, 61.20, 60.98, 52.34, 52.27, 51.66, 51.01, 48.91, 39.69, 22.33, 22.09, 22.04;

HRESIMS calcd for $C_{25}H_{40}N_4O_{18}$ (M-H⁺)⁻ m/z 698.2374, found m/z 698.2347.

10. HRMS and NMR data of di-sialyl GNB derivatives

NMR spectra were recorded in D_2O with a Bruker AV 400 spectrometer ¹H NMR at 400 MHz or 600MHz, ¹³C NMR at 150 HMz, ³¹P NMR at 162 MHz, ¹⁹F NMR at 376 MHz). Chemical shifts were reported in δ (ppm) from an internal standard of TMS (δ 0.00). Coupling constants are reported in hertz. High-resolution electrospray-ionization mass spectra (HRESIMS) were obtained on a Thermo Orbitrap elite.

Neu5Aca2-3Gal6deoxyβ1-3(Neu5Aca2-6)GalNAc (15.4 mg), 28

¹H NMR (600 MHz, D₂O) δ 5.22 (d, *J* = 3.6 Hz, 0.5H), 4.67 (d, *J* = 8.4 Hz, 0.5H), 4.53 (d, *J* = 7.8 Hz, 0.5H), 4.47 (d, *J* = 7.8 Hz, 0.5H), 4.37 (t, *J* = 3.6 Hz, 1H), 4.33 (t, *J* = 3.2 Hz, 1H), 4.28–4.25 (m, 3H), 4.23-4.19 (m, 2H), 4.08–3.81 (m, 8H), 3.75–3.57 (m, 10H), 3.51–3.47 (m, 1H), 2.78–2.71 (m, 2H), 2.04 (s, 6H), 2.03 (s, 3H), 1.78 (t, *J* = 12.0 Hz, 1H), 1.72-1.67 (m, 1H), 1.22–1.21 (m, 3H); ¹³C NMR (150 MHz, D₂O) δ 175.02, 174.98, 174.63, 173.89, 173.86, 173.50, 173.46, 104.34, 104.20, 100.44, 99.61, 96.44, 89.29, 82.77, 82.71, 76.84, 75.96, 75.90, 74.22, 72.77, 72.61, 71.83, 71.67, 70.48, 70.44, 69.98, 69.20, 68.89, 68.81, 68.71, 68.58, 68.39, 68.22, 68.05, 64.44, 64.40, 62.64, 62.47, 51.86, 51.68, 40.15, 39.86, 22.35, 22.11, 22.04, 15.51; HRESIMS calcd for C₃₆H₃₈N₃O₂₆ (M-H⁺)⁻ m/z 948.3314, found m/z 948.3328.

Neu5Aca2-3GalAβ1-3(Neu5Aca2-6)GalNAc (2.2 mg), 29

¹H NMR (600 MHz, D₂O) δ 5.23 (d, *J* = 3.6 Hz, 0.5H), 4.74–4.67 (m, 2H), 4.57 (d, *J* = 7.8 Hz, 0.5H), 4.51 (d, *J* = 7.8 Hz, 0.5H), 4.33 (d, *J* = 3.0 Hz, 0.5H), 4.29–4.27 (m, 1H), 4.30–4.22 (m, 1H), 4.14–4.10 (m, 1H), 4.07–3.80 (m, 10H), 3.73-3.51 (m, 11H), 2.78–2.72 (m, 2H), 2.04 (s, 6H), 2.03 (s, 3H), 1.82 (t, *J* = 12.0 Hz, 1H), 1.73–1.68 (m, 2H); ¹³C NMR was not able to obtain because of small amount of the compound. HRESIMS calcd for C₃₆H₅₆N₃O₂₈ (M-H⁺)⁻ m/z 978.3056, found m/z 978.3063.

Neu5Aca2-3Gal6N₃β1-3(Neu5Aca2-6)GalNAc (14.8 mg), 30

¹H NMR (600 MHz, D₂O) δ 5.19 (d, *J* = 3.6 Hz, 0.5H), 4.69–4.66 (m, 1H), 4.57 (d, *J* = 7.8 Hz, 0.5H), 4.51 (d, *J* = 7.8 Hz, 0.5H), 4.27 (dd, *J* = 3.6, 10.8 Hz, 0.5H), 4.22–4.14 (m, 2H), 4.07–3.96 (m, 2H), 3.90–3.77 (m, 10H), 3.69–3.50 (m, 12H), 3.28 (dd, *J* = 3.0, 7.2 Hz, 1H), 2.74–2.67 (m, 2H), 2.01 (s, 9H), 1.76 (t, *J* = 12.0 Hz, 1H), 1.68–1.63 (m, 1H); ¹³C NMR (150 MHz, D₂O) δ 175.01, 174.96, 174.65, 173.87, 173.85, 173.38, 104.66, 104.54, 100.39, 99.69, 95.10, 91.21, 77.10, 75.38, 75.32, 73.47, 73.28, 72.26, 72.56, 71.81, 71.61, 68.88, 68.85, 68.81, 68.78, 68.35, 68.24, 68.17, 68.02, 67.84, 63.92, 62.60, 62.45, 51.83, 51.64, 50.94, 40.13, 39.64, 22.31, 22.07, 22.01;

HRESIMS calcd for $C_{36}H_{57}N_6O_{26}$ (M-H⁺)⁻ m/z 989.3328, found m/z 989.3335.

11. NMR Spectra of synthesized compounds ¹H NMR, ¹³C NMR and ³¹P NMR spectra of Gal2N₃-1-P, **34**



¹H NMR, ¹³C NMR, ³¹P NMR and ¹⁹F NMR of GalF-1-P, **35**





¹H NMR, ¹³C NMR, and ³¹P NMR of Tal-1-P, **36**





¹H NMR, ¹³C NMR and ³¹P NMR of Gal6N₃-1-P, **39**











 ^1H NMR and ^{13}C NMR of Gal6deoxy $\beta1\text{--}3\text{GalNAc},$ 18





¹H NMR and ¹³C NMR of Gal6N₃ β 1–3GalNAc, **28**





 ^1H NMR and ^{13}C NMR of Gal2deoxy $\beta1\text{--}3\text{GlcNAc},$ **42**





¹H NMR, ¹³C NMR and ¹⁹F NMR of GalF β 1–3GlcNAc, **45**

¹H NMR and ¹³C NMR of Gal6deoxy β 1–3GlcNAc, **47**





¹H NMR and ¹³C NMR of Neu5Ac α 2–3Gal β 1–3GlcNAc, **22**





 ^1H NMR, ^{13}C NMR and ^{19}F NMR of Neu5Aca2–3GalF\beta1–3GlcNAc, **24**





 ^1H NMR and ^{13}C NMR of Neu5Aca2–3Gal6deoxy β 1–3GlcNAc, **25**









¹H NMR and ¹³C NMR of Neu5Ac α 2–3Gal6deoxy β 1–3(Neu5Ac α 2–6)GalNAc, **28**



¹H NMR of Neu5Ac α 2–3GalA β 1–3(Neu5Ac α 2–6)GalNAc, **29**



¹H NMR and ¹³C NMR of Neu5Ac α 2–3Gal6N₃ β 1–3(Neu5Ac α 2–6)GalNAc, **30**





12. References

- 1 H. Yu, V. Thon, K. Lau, L. Cai, Y. Chen, S. Mu, Y. Li, P. G. Wang and X. Chen, *Chem Commun (Camb)*, 2010, **46**, 7507.
- L. Li, Y. Liu, W. Wang, J. Cheng, W. Zhao and P. Wang, *Carbohydr Res*, 2012, **355**, 35.
- 3 H. Yu, H. Yu, R. Karpel and X. Chen, *Bioorg Med Chem*, 2004, **12**, 6427.
- 4 G. Sugiarto, K. Lau, Y. Li, Z. Khedri, H. Yu, D. T. Le and X. Chen, *Mol Biosyst*, 2011, **7**, 3021.
- 5 H. Yu, S. Huang, H. Chokhawala, M. Sun, H. Zheng and X. Chen, *Angew Chem Int Ed Engl*, 2006, **45**, 3938.
- 6 L. Li, Y. Liu, Y. Wan, Y. Li, X. Chen, W. Zhao and P. G. Wang, *Org Lett*, 2013, **15**, 5528.