Organic & Biomolecular Chemistry

Cite this: *Org. Biomol. Chem.*, 2011, **9**, 2784

www.rsc.org/obc

PAPER

Sequential two-step multienzyme synthesis of tumor-associated sialyl T-antigens and derivatives†

Kam Lau, Hai Yu, Vireak Thon, Zahra Khedri, Meghan E. Leon, Bao K. Tran and Xi Chen*

Received 30th December 2010, Accepted 1st February 2011
DOI: 10.1039/c0ob01269f

A series of α2–3-sialylated β1–3-linked galactosides, including sialyl T-antigens, 3′-sialyl *galacto*-*N*-biose, 3′-sialyl lacto-*N*-biose, and their derivatives containing natural and non-natural sialic acid forms have been synthesized from simple monosaccharides using an efficient sequential two-step multienzyme approach.

Introduction

α2–3-Sialylated β1–3-linked galactosides including 3′-sialyl type 1 (Siaα2–3Galβ1–3GlcNAcOR), 3′-sialyl type 3 or core 1 (Siaα2–3Galβ1–3GalNAcαOR), and 3′-sialyl type 4 (Siaα2–3Galβ1–3GalNAcβOR) glycan structures are important sialylated carbohydrate moieties on glycoproteins and glycolipids in nature. For example, α2–3-sialylated type 1 glycan (Siaα2–3Galβ1–3GlcNAcOR) is a part of important cancer-related antigens such as sialyl Lewis[a] antigen presented in the glycoproteins and glycolipids on the cell surface.[1] It is also a common structure presented in human milk oligosaccharides (HMOs).[2] α2–3-Linked sialyl type 3 glycans (Siaα2–3Galβ1–3GalNAcαOR) have been found in glycoproteins.[3] Sialyl core 1 structure attached to a serine or a threonine residue on peptides or proteins is also called sialyl T-antigen (sialyl Thomsen-Friedenreich antigen or sialyl TF-antigen, Siaα2–3Galβ1–3GalNAcα1-*O*-Ser/Thr). It has been found at an elevated level in *O*-linked glycoproteins such as mucins in certain cancers including breast cancer.[4] Desialylation of sialyl T-antigen in erythrocytes may cause hemolytic-uremic syndrome (HUS), a disease predominantly affecting children.[5] Sialyl T-antigens and derivatives have been used in diagnosis for early stages of HUS. They have also been considered as potential cancer vaccines.[6] 3′-Sialyl type 4 structures (Siaα2–3Galβ1–3GalNAcβOR) are presented on complex gangliosides, a family of sialic acid-containing glycosphingolipids, such as GM1b, GD1a, GT1a, GT1b, GQ1b, and c-series ganglioside GQ1c, GP1c, GH1c, *etc*.[7]

We recently reported an efficient one-pot two-enzyme approach for preparative-scale (50–100 mg scale) synthesis of β1–3-linked galactosides including T-antigens, *galacto*-*N*-biose (Galβ1–3GalNAc), lacto-*N*-biose (Galβ1–3GlcNAc), and their derivatives using a recombinant *Escherichia coli* K-12 galactokinase (GalK) and a novel D-galactosyl-β1–3-*N*-acetyl-D-hexosamine phosphorylase cloned from *Bifidobacterium infantis* (BiGalHexNAcP).[8] Here we show that the one-pot two-enzyme β1–3-galactosylation system can be easily scaled up to prepare a larger amount of β1–3-linked galactosyl disaccharides as demonstrated here for the synthesis of Galβ1–3GalNAcαProN_3. We also present data that the β1–3-linked galactodisaccharide products obtained from the one-pot two-enzyme reactions can be used as sialyltransferase acceptors for synthesizing α2–3-sialosides containing natural and non-natural sialic acid forms in a one-pot three-enzyme sialylation system[9] containing a sialic acid aldolase, a CMP-sialic acid synthetase, and a recombinant *Pasteurella multocida* multifunctional α2–3-sialyltransferase (PmST1).[10] Using this two-step multienzyme reaction process in sequence, α2–3-sialylated β1–3-linked galactoside trisaccharides containing the common sialic acid form *N*-acetylneuraminic acid (Neu5Ac), a non-human sialic acid form *N*-glycolylneuraminic acid (Neu5Gc), a less common sialic acid form 2-keto-3-deoxy-D-*glycero*-D-*galacto*-nonulosonic acid (Kdn), and their azido-derivatives have been successfully prepared. Overall, this sequential two-step multienzyme method allows facile access of diverse sialyltrisaccharides from simple monosaccharides, including *N*-acetyl-D-galactosamine (GalNAc), *N*-acetyl-D-glucosamine (GlcNAc), or their derivatives as simple acceptor substrates and D-galactose (Gal) as well as D-mannose, *N*-acetyl-D-mannosamine (ManNAc), or their derivatives as donor precursors of carbohydrate phosphorylases or glycosyltransferases.

Results and discussion

Sequential two-step multienzyme approach for the synthesis of sialyl galactosides

As shown in Fig. 1, α2–3-sialylated β1–3-linked galactoside trisaccharides containing different sialic acid forms and various underlying glycans can be readily obtained from simple monosaccharides using an efficient two-step multienzyme approach carried out in sequence.

Step I is a one-pot two-enzyme reaction involving a recombinant *E. coli* K-12 galactokinase (GalK) and a *B. infantis* phosphorylase

Department of Chemistry, University of California, One Shields Avenue, Davis, California, USA. E-mail: chen@chem.ucdavis.edu; Fax: +1-5307528995; Tel: +1-5307546037

† Electronic supplementary information (ESI) available: NMR spectra of sialylated galactosides. See DOI: 10.1039/c0ob01269f

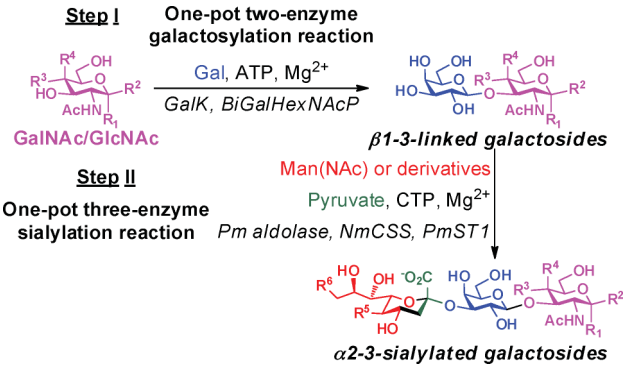


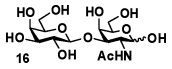
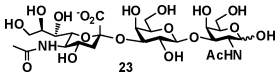
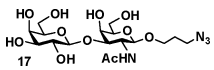
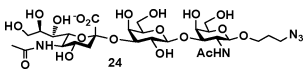
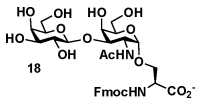
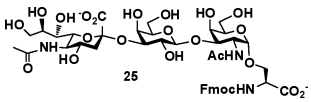
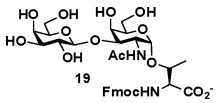
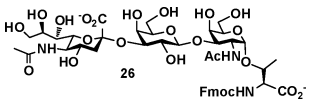
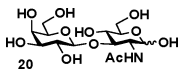
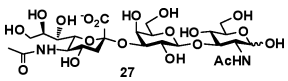
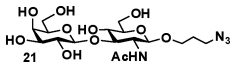
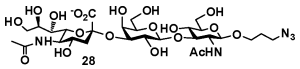
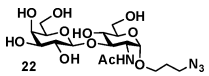
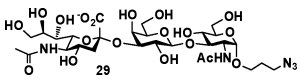
Fig. 1 Sequential two-step multienzyme synthesis of α2–3-sialylated β1–3-linked galactoside trisaccharides containing different sialic acid forms and diverse underlying glycans. Enzymes: GalK, *E. coli* K-12 galactokinase; BiGalHexNAcP, *B. infantis* D-galactosyl-β1–3-*N*-acetyl-D-hexosamine phosphorylase; Pm aldolase, *P. multocida* sialic acid aldolase; NmCSS, *N. meningitidis* CMP-sialic acid synthetase; PmST1, *P. multocida* multifunctional α2–3-sialyltransferase. Compounds: Gal, D-galactose; GalNAc, *N*-acetyl-D-galactosamine; GlcNAc, *N*-acetyl-D-glucosamine; ATP, adenosine 5′-triphosphate; Man, D-mannose; ManNAc, *N*-acetyl-D-mannosamine; CTP, cytidine 5′-triphosphate; Sia, sialic acid.

(BiGalHexNAcP).[8] GalK is responsible for synthesizing galactosyl-1-phosphate (Gal-1-P) from Gal and adenosine 5′-triphosphate (ATP). The Gal-1-P generated *in situ* can be directly used as a donor substrate by the BiGalHexNAcP which catalyzes the transfer of Gal to simple monosaccharides such as GalNAc, GlcNAc, and their derivatives for the synthesis of a library of β1–3-linked galactodisaccharides in excellent yields. Therefore, in this step, simple monosaccharides are readily linked together by a two-enzyme process to form disaccharides.

The disaccharides formed in Step I can be purified and used as sialyltransferase acceptors in Step II for the formation of α2–3-sialylated β1–3-linked galactoside trisaccharides. Step II is a one-pot three-enzyme reaction involving a sialic acid aldolase, a CMP-sialic acid synthetase, and an α2–3-sialyltransferase.[9] The sialic acid aldolase is responsible for synthesizing different sialic acid forms from mannose, ManNAc, and their derivatives in the presence of pyruvate. The sialic acids generated *in situ* can then be activated by the CMP-sialic acid synthetase to form CMP-sialic acids, the donor substrates of the α2–3-sialyltransferase. The α2–3-sialyltransferase is then responsible for the formation of α2–3-linked sialylgalactosides. A *Pasteurella multocida* sialic acid aldolase,[11] an *Neisseria meningitidis* CMP-sialic acid synthetase,[12] and a *Pasteurella multocida* multifunctional α2–3-sialyltransferase[10] were chosen for this reaction due to their superior expression levels in *E. coli* and promiscuous substrate specificities. In this step, different sialic acid forms are formed from mannose, ManNAc, or their derivatives, activated, and linked to various disaccharides formed in Step I for the synthesis of trisaccharides with different terminal sialic acids by a three-enzyme process.

The application of the sequential two-step multienzyme approach has been explored for the synthesis of two libraries of sialylgalactosides. One library contains α2–3-linked sialylgalactosides with different terminal sialic acid forms and the same underlying β1–3-linked galactoside (Galβ1–3GalNAcαProN$_3$) (Fig. 2). The other contains α2–3-linked sialylgalactosides with a common Neu5Ac form, and different underlying β1–3-linked galactosides (Table 1).

Table 1 Enzymatic synthesis of α2–3-linked sialosides with terminal Neu5Ac and different underlying β1–3-linked galactosides

Acceptors	Products	Yields (%)
		90
		92
		86
		89
		94
		91
		89

Preparation of sialyl T-antigens containing different sialic acid forms

For preparative-scale synthesis of sialyl T-antigens containing different sialic acid forms, galactoside disaccharide with a propyl azide aglycon, Galβ1–3GalNAcαProN$_3$ **3**, was synthesized at pH 6.5 from 3-azidopropyl α-*N*-acetyl-D-galactosamine (GalNAcαProN$_3$) **1**[13] and D-galactose **2** using the Step I one-pot two-enzyme galactosylation reaction (Fig. 1). As shown in Fig. 2, compound **3** was readily obtained in an excellent 91% yield after purification by silica gel and gel filtration chromatography. The obtained Galβ1–3GalNAcαProN$_3$ **3** was then used as a sialyltransferase acceptor for the synthesis of sialyl T-antigens containing different terminal sialic acid forms using the Step II one-pot three-enzyme sialylation reaction at pH 8.5.

As shown in Fig. 2, PmST1 exhibited promiscuous donor substrate specificity and catalyzed the transfer of sialic acids from *in situ* generated CMP-sialic acids to Galβ1–3GalNAcαProN$_3$ **3** synthesized by the Step I reaction to produce sialyltrisaccharides **10–15** containing different terminal sialic acid forms in good to excellent (85–97%) yields. The sialosides containing three naturally occurring sialic acid forms, Neu5Acα2–3Galβ1–3GalNAcαProN$_3$ **10**,[15] Neu5Gcα2–3Galβ1–3GalNAcαProN$_3$ **11**,[15] and Kdnα2–3Galβ1–3GalNAcαProN$_3$ **12** were obtained in 87%, 94%, and 89% yields, respectively, from ManNAc **4**, ManNGc **5**, and mannose **6** as six-carbon sialic acid precursors. Neu5Gc is a non-human sialic acid which is overexpressed in certain cancer

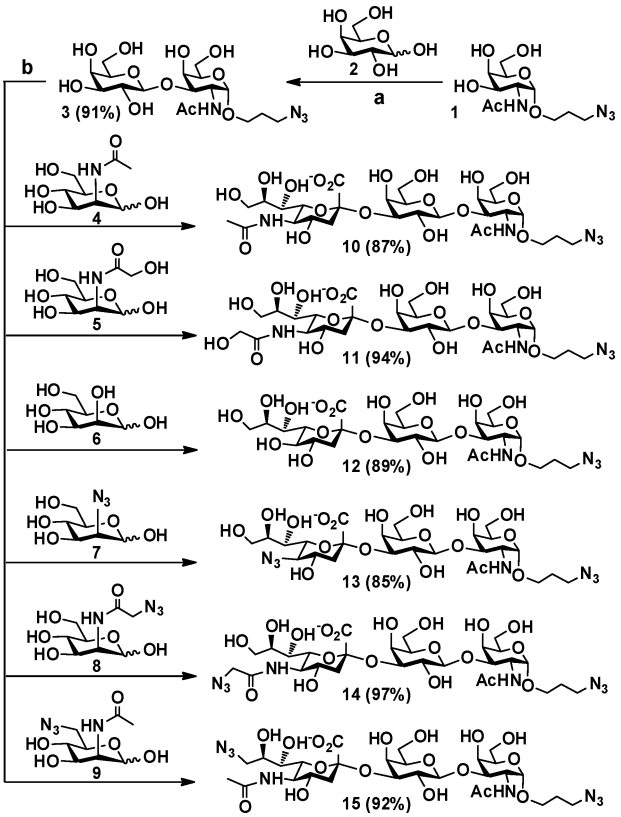


Fig. 2 Synthesis of sialyl T-antigens Siaα2–3Galβ1–3GalNAcαProN$_3$ containing different natural and non-natural sialic acid forms. Reagents and conditions: (a) ATP, Mg^{2+}, Tris-HCl buffer (pH 6.5), GalK, BiGalHexNAcP; (b) pyruvate, CTP, Mg^{2+}, Tris-HCl buffer (pH 8.5), Pm aldolase, NmCSS, and PmST1.

cells.[5,14] Sialyl T-antigen trisaccharides containing non-natural sialic acids including Kdn5N$_3$α2–3Galβ1–3GalNAcαProN$_3$ **13**, Neu5AcN$_3$α2–3Galβ1–3GalNAcαProN$_3$ **14**, and Neu5Ac9N$_3$α2–3Galβ1–3GalNAcαProN$_3$ **15**, which contain an azido group at the C-5 or C-9 position of the terminal sialic acid residue, were also obtained in high yields (85–97%) from C2- or C6-modified ManNAc or mannose derivatives 2-azido-mannose (Man2N$_3$) **7**, *N*-azidoacetyl mannosamine (ManNAcN$_3$) **8**, and *N*-acetyl-6-azido-mannosamine (ManNAc6N$_3$) **9**, respectively. The structures of all purified sialyl T-antigen trisaccharide products were confirmed by nuclear magnetic resonance (NMR) spectroscopy and high-resolution mass spectrometry (HRMS).

Preparation of α2–3-linked sialosides with terminal Neu5Ac and different underlying β1–3-linked galactosides

Sialylgalactosides with the most common sialic acid form Neu5Ac and different underlying β1–3-linked galactosides (Table 1) were also synthesized using the sequential two-step multienzyme approach. To do this, β1–3-linked galactodisaccharides **16–22** were synthesized from D-galactose and the corresponding GalNAc, GlcNAc, or derivatives such as the BiGalHexNAcP acceptors in the Step I one-pot two-enzyme galactosylation process (Fig. 1) as described previously.[8] These β1–3-linked galactodisaccharides **16–22** were then used as sialyltransferase acceptors in the Step II one-pot three-enzyme sialylation process for the synthesis of α2–3-linked sialosides containing Neu5Ac formed from ManNAc **4** as a sialic acid precursor. As shown in Table 1, preparative-scale sialylation of *galacto*-*N*-biose (GNB) Galβ1–3GalNAc **16** and Galβ1–3GalNAcβProN$_3$ **17** by Neu5Ac formed from ManNAc and pyruvate successfully produced the sialylated products 3′-sialyl *galacto*-*N*-biose (GNB) Neu5Acα2–3Galβ1–3GalNAc **23**[16] and Neu5cα2–3Galβ1–3GalNAcβProN$_3$ **24**[15] in 90% and 92% yields, respectively. The one-pot three-enzyme system was also applied in the sialylation of biologically important T-antigens Galβ1–3GalNAcα1-*O*-Ser and Galβ1–3GalNAcα1-*O*-Thr to produce the desired trisaccharide products **25** and **26** in excellent 86% and 89% yields, respectively. Similarly, α2–3-sialylation of lacto-*N*-biose (LNB) Galβ1–3GlcNAc **20** with Neu5Ac produced 3′-sialyl lacto-*N*-biose (LNB) Neu5Acα2–3Galβ1–3GlcNAc **27**[17] in 94% yield. In addition, α2–3-sialylation of type 1 disaccharides **21** and **22** with Neu5NAc using the one-pot three-enzyme system produced sialylated products Neu5Acα2–3Galβ1–3GlcNAcβProN$_3$ **28**[15] and Neu5Acα2–3Galβ1–3GlcNAcαProN$_3$ **29** in 91% and 89% yields, respectively.

Conclusions

In conclusion, we have developed an efficient sequential two-step multienzyme approach for the synthesis of a series of α2–3-sialylated β1–3-galactosyl disaccharides containing different underlying glycans and various natural and non-natural terminal sialic acid residues. Both LNB (Galβ1–3GlcNAc) and GNB (Galβ1–3GalNAc) type disaccharides, including T-antigens are excellent acceptor substrates for the *Pasteurella multodica* multifunctional α2–3-sialyltransferase (PmST1). The sequential two-step multienzyme synthetic approach thus allows facile access of complex trisaccharides from simple monosaccharides and derivatives.

Experimental

General methods

Chemicals were purchased and used without further purification. ^{1}H NMR (600 MHz) and ^{13}C NMR (150 MHz) spectra were recorded on a Varian VNMRS 600 MHz spectrometer. High resolution electrospray ionization (ESI) mass spectra were obtained at the Mass Spectrometry Facility in the University of California-Davis. Optical rotation was recorded on an Autopol IV Automatic polarimeter at 589 nm wavelength. Infrared spectra were recorded on a PerkinElmer Spectrum 100 ATR-FTIR. Silica gel 60 Å was used for flash column chromatography. Thin-layer chromatography (TLC) was performed on silica gel plates using *p*-anisaldehyde sugar stain or 5% sulfuric acid in ethanol stain for detection. Gel filtration chromatography was performed using a column (100 cm × 2.5 cm) packed with BioGel P-2 Fine resins.

General procedures for sequential two-step multienzyme synthesis of α2–3-sialylated β1–3-linked galactosides

Step I: One-pot two-enzyme synthesis of β1–3-linked galactoside Galβ1–3GalNAcαProN$_3$ 3. The process started with chemical synthesis of GalNAcαProN$_3$ **1** similar to that described previously.[13] To obtain Galβ1–3GalNAcαProN$_3$ **3**, GalNAcαProN$_3$ **1** (300 mg, 0.98 mmol), galactose (266 mg,

1.48 mmol), ATP (816 mg, 1.48 mmol) and $MgCl_2$ (132 mg, 22 mM) were dissolved in a Tris-HCl buffer (30 mL, 100 mM, pH 6.5). After the addition of GalK (10.8 mg) and BiGalHexNAcP (9.6 mg), the reaction was carried out by incubating the solution in an incubator shaker for 24 h at 37 °C. The reaction was then stopped by adding cold EtOH (30 mL) and the mixture was centrifuged to remove the precipitates. The filtrate was concentrated on a rotary evaporator and purified by a BioGel P-2 filtration column (eluted with water) and a silica gel column (EtOAc–MeOH–H_2O = 7 : 2 : 1, by volume) to produce Galβ1–3GalNAcαProN$_3$ **3** (418 mg, 91%). ^{1}H NMR (600 MHz, D_2O): δ 4.90 (d, 1H, J = 3.6 Hz), 4.48 (d, 1H, J = 7.8 Hz), 4.35 (dd, 1H, J = 10.8 and 3.6 Hz), 4.26 (d, 1H, J = 2.4 Hz), 4.04 (dd, 1H, J = 10.8 and 3.0 Hz), 4.00 (t, 1H, J = 6.0 Hz), 3.92 (d, 1H, J = 3.0 Hz), 3.83–3.74 (m, 5H), 3.68–3.63 (m, 2H), 3.58–3.45 (m, 4H), 2.04 (s, 3H), 1.92 (m, 2H); ^{13}C NMR (150 MHz, D_2O): δ 174.65, 104.89, 97.41, 77.40, 75.18, 72.71, 70.81, 68.92, 68.79, 65.13, 61.39, 61.19, 48.87, 48.41, 28.17, 22.22.

Compounds **16–22** were synthesized using the same procedures as reported before[8] which were similar to that described above for preparing compound **3**.

Step II: One-pot three-enzyme preparative-scale synthesis of α2–3-linked sialosides. β1–3-Linked galactosides (50–100 mg) produced from Step I, a sialic acid precursor (mannose, ManNAc, or their derivatives, 1.5 equiv.), sodium pyruvate (7.5 equiv.), and CTP (1.5 equiv.) were dissolved in Tris-HCl buffer (10 mL, 100 mM, pH 8.5) containing $MgCl_2$ (20 mM) and appropriate amounts of Pm aldolase (0.7–1.0 mg), NmCSS (0.5–0.8 mg), and PmST1 (0.2–0.4 mg). The reactions were carried out by incubating the reaction mixture in an incubator shaker at 37 °C for 1–2 h with agitation at 140 rpm. The product formation was monitored by TLC developed with EtOAc–MeOH–H_2O–HOAc = 4 : 2 : 1 : 0.1 (by volume) and stained with *p*-anisaldehyde sugar stain. When an optimal yield was achieved, the reaction was stopped by adding the same volume (10 mL) of cold EtOH and incubation at 4 °C for 30 min. The mixture was then centrifuged and the precipitates were removed. The supernatant was concentrated, passed through a BioGel P-2 gel filtration column, and eluted with water to obtain the partially purified product. A silica gel column was then used to obtain pure sialylated products using EtOAc–MeOH–H_2O = 6 : 2 : 1 (by volume) as the mobile phase.

Neu5Acα2–3Galβ1–3GalNAcαProN$_3$ (10). 54 mg, yield 87%. $[\alpha]_D^{24}$ = +4.16 (c = 1.0 H_2O). Wavenumber$_{max}$ (film)/cm^{-1} 3292 (s, OH), 2944 (s, C–H alkene), 2101 (s, N_3), 1621 (s, C═O, carboxylic acid), 1560 (m, C═O, amide), 1023 (s, C–N). ^{1}H NMR (600 MHz, D_2O): δ 4.87 (d, 1H, J = 3.6 Hz), 4.51 (d, 1H, J = 7.8 Hz), 4.29–4.27 (m, 1H), 4.21 (m, 1H), 4.05–3.31 (m, 21H), 2.72 (dd, 1H, J = 4.8 and 12.6 Hz), 1.99 (s, 6H), 1.87 (m, 2H), 1.75 (t, 1H, J = 12.3 Hz). ^{13}C NMR (150 MHz, D_2O): δ 175.10, 174.69, 174.03, 104.58, 99.83, 97.30, 77.52, 75.76, 74.90, 72.92, 71.95, 70.75, 69.21, 68.69, 68.51, 68.16, 67.51, 65.03, 62.62, 61.35, 61.10, 59.58, 51.78, 48.81, 48.30, 39.83, 28.09, 22.17. HRMS (ESI) m/z calcd for $C_{28}H_{47}N_5O_{19}Na$ (M+Na) 780.2763, found 780.2750.

Neu5Gcα2–3Galβ1–3GalNAcαProN$_3$ (11). 58 mg, yield 94%. $[\alpha]_D^{24}$ = +5.31 (c = 1.0 H_2O). Wavenumber$_{max}$ (film)/cm^{-1} 3287 (s, OH), 3012 (s, C–H alkene), 2101 (s, N_3), 1614 (s, C═O, carboxylic acid), 1562 (m, C═O, amide), 1026 (s, C–N). ^{1}H NMR (600 MHz, D_2O): δ 4.88 (d, 1H, J = 4.2 Hz), 4.52 (d, 1H, J = 7.8 Hz), 4.29 (dd, 1H, J = 3.6 and 11.4 Hz), 4.22 (d, 1H, J = 2.4 Hz), 4.09 (s, 2H), 4.07–3.67 (m, 15H), 3.63–3.40 (m, 6H), 2.75 (dd, 1H, J = 4.8 Hz and 12.6 Hz), 2.00 (s, 3H), 1.89 (m, 2H), 1.78 (t, 1H, J = 12.6 Hz). ^{13}C NMR (150 MHz, D_2O): δ 175.88, 174.71, 174.07, 104.60, 99.85, 97.32, 77.51, 75.77, 74.92, 72.65, 72.03, 70.77, 69.22, 68.71, 68.26, 68.10, 67.50, 65.05, 62.59, 61.36, 61.10, 51.49, 48.83, 48.31, 39.93, 28.10, 22.18. HRMS (ESI) m/z calcd for $C_{28}H_{48}N_5O_{20}$ (M+H) 774.2893, found 774.2898.

Kdnα2–3Galβ1–3GalNAcαProN$_3$ (12). 55 mg, yield 89%. $[\alpha]_D^{24}$ = +1.57 (c = 1.1 H_2O). Wavenumber$_{max}$ (film)/cm^{-1} 3291 (s, OH), 2932 (s, C–H alkene), 2102 (s, N_3), 1611 (s, C═O, carboxylic acid), 1560 (m, C═O, amide), 1029 (s, C–N). ^{1}H NMR (600 MHz, D_2O): δ 4.86 (d, 1H, J = 3.6 Hz), 4.51 (d, 1H, J = 7.8 Hz), 4.30 (dd, 1H, J = 3.2 and 11.2 Hz), 4.22 (m, 1H), 4.04–4.02 (m, 2H), 3.97 (dd, 1H, J = 4.8 and 7.2 Hz), 3.90–3.66 (m, 10H), 3.62–3.45 (m, 8H), 2.69 (dd, 1H, J = 12.6 Hz and 4.8 Hz), 2.00 (s, 3H), 1.89 (m, 2H), 1.72 (t, 1H, J = 12.6 Hz). ^{13}C NMR (150 MHz, D_2O): δ 174.46, 173.94, 104.37, 99.51, 97.07, 77.18, 75.49, 74.69, 73.70, 72.03, 70.52, 70.14, 69.59, 68.95, 68.48, 67.56, 67.16, 64.80, 62.43, 61.10, 60.87, 48.59, 39.32, 27.85, 26.29, 21.93. HRMS (ESI) m/z calcd for $C_{26}H_{45}N_4O_{19}$ (M+H) 717.2678, found 717.2688.

Kdn5N$_3$α2–3Galβ1–3GalNAcαProN$_3$ (13). 51 mg, yield 85%. $[\alpha]_D^{24}$ = +6.56 (c = 1.0 in H_2O). Wavenumber$_{max}$ (film)/cm^{-1} 3291 (s, OH), 2927 (s, C–H alkene), 2106 (s, N_3), 1606 (s, C═O, carboxylic acid), 1555 (m, C═O, amide), 1027 (s, C–N). ^{1}H NMR (600 MHz, D_2O): δ 4.92 (d, 1H, J = 3.0 Hz), 4.55 (d, 1H, J = 7.8 Hz), 4.34 (dd, 1H, J = 3.0 and 10.8 Hz), 4.26 (m, 1H)), 4.07–3.89 (m, 6H), 3.85–3.65 (m, 9H), 3.57–3.45 (m, 6H), 2.76 (dd, 1H, J = 4.8 and 12.6 Hz), 2.05 (s, 3H), 1.92 (m, 2H), 1.80 (t, 1H, J = 12.6 Hz). ^{13}C NMR (150 MHz, D_2O): δ 174.73, 173.87, 104.62, 99.87, 97.38, 77.47, 75.82, 74.96, 72.97, 72.21, 70.82, 69.65, 69.24, 68.78, 68.55, 67.46, 65.14, 62.76, 62.65, 61.40, 61.16, 48.89, 48.39, 39.82, 28.15, 22.25. HRMS (ESI) m/z calcd for $C_{26}H_{44}N_7O_{18}$ (M+H) 742.2743, found 742.2753.

Neu5AcN$_3$α2–3Galβ1–3GalNAcαProN$_3$ (14). 58 mg, yield 97%. $[\alpha]_D^{24}$ = +7.81 (c = 1.0 in H_2O). Wavenumber$_{max}$ (film)/cm^{-1} 3285 (s, OH), 2926 (s, C–H alkene), 2100 (s, N_3), 1611 (s, C═O, carboxylic acid), 1563 (m, C═O, amide), 1029 (s, C–N). ^{1}H NMR (600 MHz, D_2O): δ 4.88 (d, 1H, J = 3.0 Hz), 4.52 (d, 1H, J = 7.8 Hz), 4.29 (dd, 1H, J = 3.0 and 10.8 Hz), 4.22 (m, 1H)), 4.06–4.02 (m, 4H), 3.97–3.69 (m, 12H), 3.62–3.42 (m, 7H), 2.74 (dd, 1H, J = 4.8 Hz and 12.6 Hz), 2.00 (s, 3H), 1.88 (m, 2H), 1.77 (t, 1H, J = 12.6 Hz). ^{13}C NMR (150 MHz, D_2O): δ 174.71, 174.04, 171.28, 104.60, 99.87, 97.33, 77.52, 75.79, 74.93, 72.61, 72.06, 70.78, 69.23, 68.72, 68.36, 68.13, 67.52, 65.07, 62.63, 61.37, 61.13, 52.05, 51.89, 48.83, 48.33, 39.91, 28.12, 22.21. HRMS (ESI) m/z calcd for $C_{28}H_{47}N_8O_{19}$ (M+H) 799.2957, found 799.2969.

Neu5Ac9N$_3$α2–3Galβ1–3GalNAcαProN$_3$ (15). 56 mg, yield 92%. $[\alpha]_D^{24}$ = +5.31 (c = 1.0 H_2O). Wavenumber$_{max}$ (film)/cm^{-1} 3291 (s, OH), 2932 (s, C–H alkene), 2101 (s, N_3), 1608 (s, C═O, carboxylic acid), 1560 (m, C═O, amide), 1030 (s, C–N). ^{1}H NMR (600 MHz, D_2O): δ 4.93 (d, 1H, J = 3.6 Hz), 4.56 (d, 1H, J = 7.8 Hz), 4.34 (dd, 1H, J = 3.6 and 10.8 Hz), 4.26 (d, 1H, J = 2.4 Hz), 4.21–4.00 (m, 4H), 3.95 (d, 1H, J = 3.0 Hz), 3.88–3.61 (m, 11H), 3.58–3.46 (m, 5H), 2.77 (dd, 1H, J = 4.8 and 12.6 Hz), 2.05 (s, 6H), 1.93 (m, 2H), 1.80 (t, 1H, J = 12.6 Hz). ^{13}C NMR

(150 MHz, D_2O): δ 175.08, 174.68, 173.93, 104.55, 99.88, 97.32, 77.48, 75.82, 74.92, 72.74, 70.78, 70.67, 69.25, 68.86. 68.69, 68.51, 67.45, 65.09, 61.36, 61.12, 53.20, 51.79, 48.85, 48.35, 39.91, 28.11, 22.20. HRMS (ESI) *m/z* calcd for $C_{28}H_{47}N_8O_{18}$ (M+H) 783.3008, found 783.3022.

Neu5Acα2–3Galβ1–3GalNAc (23). 55 mg, yield 90%. $[\alpha]_D^{24}$ = +1.48 (*c* = 1.0 H_2O). $Wavenumber_{max}$ (film)/cm^{-1} 3283 (s, OH), 2937 (s, C–H alkene), 1614 (s, C═O, carboxylic acid), 1560 (m, C═O, amide), 1043 (s, C–N). ^1H NMR (600 MHz, D_2O): δ 5.21 (d, 0.6H, *J* = 3.6 Hz, H-1′α), 4.67 (d, 0.4H, *J* = 8.4 Hz, H-1′β), 4.55 (d, 0.6H, *J* = 7.8 Hz), 4.49 (d, 0.4H, *J* = 7.8 Hz), 4.26–3.51 (m, 19H), 2.74–2.73 (m, 1H), 2.00 (s, 6H), 1.77 (t, 1H, *J* = 12.3 Hz). ^{13}C NMR (150 MHz, D_2O): δ 175.11, 174.81, 174.08, 174.04, 173.95, 104.74, 104.57, 99.83, 95.31, 91.29, 80.38, 77.36, 75.81, 75.73, 74.97, 74.89, 72.93, 71.96, 70.33, 69.22, 69.15, 68.70, 68.51, 68.18, 68.06, 67.51, 62.63, 61.33, 61.11, 61.08, 52.46, 51.80, 49.12, 39.86, 22.47, 22.23, 22.19. HRMS (ESI) *m/z* calcd for $C_{25}H_{42}N_2O_{19}Na$ (M+Na) 697.2279, found 697.2270.

Neu5Acα2–3Galβ1–3GalNAcβProN$_3$ (24). 60 mg, yield 92%. $[\alpha]_D^{24}$ = –0.83 (*c* = 1.0 H_2O). $Wavenumber_{max}$ (film)/cm^{-1} 3280 (s, OH), 2936 (s, C–H alkene), 2103 (s, N_3), 1636 (s, C═O, carboxylic acid), 1563 (m, C═O, amide), 1027 (s, C–N). ^1H NMR (600 MHz, D_2O): δ 4.53 (d, 1H, *J* = 7.8 Hz), 4.52 (d, 1H, *J* = 8.4 Hz), 4.19 (d, 1H, *J* = 3.0 Hz), 4.08 (dd, 1H, *J* = 3.0 and 10.2 Hz), 4.04–3.98 (m, 2H), 3.90 (d, 1H, *J* = 2.4 Hz), 3.87–3.54 (m, 16H), 3.40 (t, 2H, *J* = 6.6 Hz), 2.77 (dd, 1H, *J* = 4.5 and 12.3 Hz), 2.05 (s, 3H), 2.04 (s, 3H), 1.86 (m, 2H), 1.80 (t, 1H, *J* = 12.3 Hz). ^{13}C NMR (150 MHz, D_2O): δ 175.24, 174.95, 174.14, 104.80, 101.64, 99.95, 80.24, 75.82, 74.98, 73.04, 72.80, 72.04, 69.26, 68.59, 68.32, 68.12, 67.64, 67.27, 62.89, 62.77, 61.19, 51.93, 51.43, 48.07, 39.99, 28.36, 22.56, 22.31. HRMS (ESI) *m/z* calcd for $C_{28}H_{47}N_5O_{19}Na$ (M+Na) 780.2763, found 780.2744.

Neu5Acα2–3Galβ1–3GalNAcα1-*O*-Ser (25). 49 mg, yield 86%. $[\alpha]_D^{24}$ = +3.63 (*c* = 0.9 H_2O). $Wavenumber_{max}$ (film)/cm^{-1} 3286 (s, OH), 2942 (s, C–H alkene), 1636 (s, C═O, carboxylic acid), 1560 (m, C═O, amide), 1401 (m, ArC═C), 1029 (s, C–N). ^1H NMR (600 MHz, D_2O): δ 7.85–7.37 (m, 8H), 4.79 (d, 1H, *J* = 3.6 Hz), 4.56–3.97 (m, 8H), 3.89–3.42 (m, 17H), 2.70 (dd, 1H, *J* = 4.8 and 12.6 Hz), 1.98 (s, 3H), 1.86 (s, 3H), 1.74 (t, 1H, *J* = 12.3 Hz). ^{13}C NMR (150 MHz, D_2O): δ 176.42, 175.05, 174.66, 174.07, 157.74, 143.84, 141.72, 141.07, 140.75, 128.16, 127.64, 125.13, 120.30, 104.47, 99.79, 98.02, 83.77, 77.33, 75.68, 74.61, 74.02, 72.88, 71.88, 70.63, 69.29, 69.10, 68.47, 68.04, 67.38, 66.39, 62.50, 60.91, 60.47, 51.74, 48.51, 46.92, 39.78, 22.17, 22.13. HRMS (ESI) *m/z* calcd for $C_{43}H_{57}N_3O_{23}Na$ (M+Na) 1006.3281, found 1006.3267.

Neu5Acα2–3Galβ1–3GalNAcα1-*O*-Thr (26). 52 mg, yield 89%. $[\alpha]_D^{24}$ = +3.55 (*c* = 1.0 H_2O). $Wavenumber_{max}$ (film)/cm^{-1} 3275 (s, OH), 2932 (s, C—H alkene), 1600 (s, C═O, carboxylic acid), 1562 (m, C═O, amide), 1446 (m, ArC═C), 1032 (s, C–N). ^1H NMR (600 MHz, D_2O): δ 7.77–7.28 (m, 8H), 4.77 (d, 1H, *J* = 3.6 Hz), 4.71 (dd, 1H, *J* = 3.0 and 12.0 Hz), 4.48 (m, 1H), 4.36 (d, 1H, *J* = 7.8 Hz), 4.14–3.72 (m, 13H), 3.67–3.42 (m, 14H), 2.70 (dd, 1H, *J* = 4.8 and 12.6 Hz), 1.98 (s, 3H), 1.88 (s, 3H), 1.75 (t, 1H, *J* = 12.0 Hz), 0.96 (d, 3H, *J* = 6.6 Hz). ^{13}C NMR (150 MHz, D_2O): δ 176.70, 175.08, 174.62, 174.13, 158.23, 144.12, 143.69, 141.76, 141.12, 128.10, 127.64, 127.58, 125.05, 124.93, 120.23, 104.46, 99.92, 98.98, 77.45, 77.28, 75.67, 74.73, 73.77, 72.92, 71.88, 70.76, 69.40, 69.17, 68.72, 68.55, 68.13, 67.54, 65.88, 61.19, 61.03, 60.72, 60.60, 51.80, 48.60, 47.32, 39.68, 22.41, 22.16, 18.28. HRMS (ESI) *m/z* calcd for $C_{44}H_{59}N_3O_{23}Na$ (M+Na) 1020.3437, found 1020.3421.

Neu5Acα2–3Galβ1–3GlcNAc (27). 69 mg, yield 94%. $[\alpha]_D^{24}$ = –0.34 (*c* = 1.0 in H_2O). $Wavenumber_{max}$ (film)/cm^{-1} 3284 (s, OH), 2891 (s, C–H alkene), 1610 (s, C═O, carboxylic acid), 1565 (m, C═O, amide), 1046 (s, C–N). ^1H NMR (600 MHz, D_2O): δ 5.00 (d, 0.6H, *J* = 3.6 Hz, H-1′α), 4.57 (d, 0.4H, *J* = 8.4 Hz, H-1′β), 4.35 (d, 0.6H, *J* = 7.8 Hz), 4.31 (d, 0.4H, *J* = 7.8 Hz), 3.91–3.34 (m, 19H), 2.58 (dd, 0.6H, *J* = 4.8 and 12.6 Hz), 2.54 (dd, 0.4H, *J* = 4.8 and 12.6 Hz), 1.85 (s, 6H), 1.60 (t, 1H, *J* = 12.3 Hz). ^{13}C NMR (150 MHz, D_2O): δ 175.10, 175.05, 174.92, 174.66, 173.98, 103.63, 103.53, 99.78, 94.84. 91.17, 82.87, 80.41, 75.82, 75.78, 75.60, 75.22, 75.18, 72.93, 71.98, 71.39, 69.29, 69.22, 68.89, 68.82, 68.52, 68.17, 67.39, 62.59, 61.13, 60.87, 60.71, 59.45, 55.62, 52.98, 51.79, 39.90, 22.48, 22.22, 22.19. HRMS (ESI) *m/z* calcd for $C_{25}H_{42}N_2O_{19}Na$ (M+Na) 697.2279, found 697.2252.

Neu5Acα2–3Galβ1–3GlcNAcβProN$_3$ (28). 103 mg, yield 91%. $[\alpha]_D^{24}$ = –1.72 (*c* = 1.0 H_2O). $Wavenumber_{max}$ (film)/cm^{-1} 3278 (s, OH), 2932 (s, C–H alkene), 2101 (s, N_3), 1612 (s, C═O, carboxylic acid), 1560 (m, C═O, amide), 1030 (s, C–N). ^1H NMR (600 MHz, D_2O): δ 4.52 (d, 1H, *J* = 8.4 Hz), 4.46 (d, 1H, *J* = 8.4 Hz), 4.05 (dd, 1H, *J* = 3.0 and 9.6 Hz), 3.90–3.44 (m, 20H), 3.36 (t, 2H, *J* = 8.4 Hz), 2.72 (dd, 1H, *J* = 4.8 and 12.6 Hz), 2.00 (s, 3H), 1.99 (s, 3H), 1.80 (m, 2H), 1.75 (t, 1H, *J* = 12.3 Hz). ^{13}C NMR (150 MHz, D_2O): δ 175.08, 174.69, 174.04, 103.56, 101.02, 99.76, 82.63, 75.72, 75.49, 75.21, 72.91, 71.96, 69.19, 68.84, 68.51, 68.14, 67.36, 67.26, 62.55, 61.14, 60.84, 59.43, 51.77, 47.90, 39.88, 28.22, 22.42, 22.16. HRMS (ESI) *m/z* calcd for $C_{28}H_{47}N_5O_{19}Na$ (M+Na) 780.2763, found 780.2746.

Neu5Acα2–3Galβ1–3GlcNAcαProN$_3$ (29). 74 mg, yield 89%. $[\alpha]_D^{24}$ = +1.46 (*c* = 1.2 H_2O). $Wavenumber_{max}$ (film)/cm^{-1} 3284 (s, OH), 2929 (s, C–H alkene), 2101 (s, N_3), 1612 (s, C═O, carboxylic acid), 1560 (m, C═O, amide), 1029 (s, C–N). ^1H NMR (600 MHz, D_2O): δ 4.88 (d, 1H, *J* = 3.6 Hz), 4.53 (d, 1H, *J* = 7.8 Hz), 4.11 (dd, 1H, *J* = 3.6 and 10.8 Hz) 3.96–3.83 (m, 10H), 3.72–3.47 (m, 12H), 2.78 (dd, 1H, *J* = 4.8 and 12.6 Hz), 2.05 (s, 6H), 1.93 (m, 2H), 1.80 (t, 1H, *J* = 12.3 Hz). ^{13}C NMR (150 MHz, D_2O): δ 175.16, 174.60, 174.05, 103.51, 99.86, 97.26, 80.57, 75.85, 75.22, 72.98, 72.01, 71.80, 69.34, 68.87, 68.54, 68.22, 67.46, 65.15, 62.66, 61.44, 61.17, 60.72, 52.65, 51.85, 48.36, 39.94, 28.14, 22.24. HRMS (ESI) *m/z* calcd for $C_{28}H_{47}N_5O_{19}Na$ (M+Na) 780.2763, found 780.2756.

Acknowledgements

This work was supported by NIH grants R01GM076360 and U01CA128442. X. Chen is a Camille Dreyfus Teacher-Scholar and a UC-Davis Chancellor's Fellow. We thank Ryan Davis at the Department of Chemistry in the University of California-Davis for his help with the FTIR measurements.

Notes and references

1 (*a*) B. Anderson, L. E. Davis and M. Venegas, *Adv. Exp. Med. Biol.*, 1988, **228**, 601; (*b*) J. L. Magnani, *Arch. Biochem. Biophys.*, 2004, **426**, 122; (*c*) M. Ugorski and A. Laskowska, *Acta Biochim. Pol.*, 2002, **49**, 303; (*d*) D. S. Sanders and M. A. Kerr, *Mol. Pathol.*, 1999, **52**, 174;

(*e*) R. L. Schnaar, A. Suzuki and P. Stanley, Glycosphingolipids. In *"Essentials of Glycobiology,"* ed. A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, G. W. Hart and M. E. Etzler, Cold Spring Harbor Laboratory Press, New York, 2008, 2nd edn, pp. 129–141, .
2 C. Kunz, S. Rudloff, W. Baier, N. Klein and S. Strobel, *Annu. Rev. Nutr.*, 2000, **20**, 699.
3 E. Maes, D. Florea, A. Coppin and G. Strecker, *Eur. J. Biochem.*, 1999, **264**, 301.
4 (*a*) S. J. Storr, L. Royle, C. J. Chapman, U. M. Hamid, J. F. Robertson, A. Murray, R. A. Dwek and P. M. Rudd, *Glycobiology*, 2008, **18**, 456; (*b*) I. Brockhausen, *EMBO Rep.*, 2006, **7**, 599.
5 A. Varki, *Trends Mol. Med.*, 2008, **14**, 351.
6 J. Cohen, *Science*, 1993, **262**, 841.
7 (*a*) J. N. Scarsdale, J. H. Prestegard and R. K. Yu, *Biochemistry*, 1990, **29**, 9843; (*b*) S. Sabesan, J. O. Duus, T. Fukunaga, K. Bock and S. Ludvigsen, *J. Am. Chem. Soc.*, 1991, **113**, 3236; (*c*) J. M. Wiseman and J. B. Li, *Anal. Chem.*, 2010, **82**, 8866; (*d*) R. K. Yu and S. Ando, *Adv. Exp. Med. Biol.*, 1980, **125**, 33; (*e*) M. Saito, H. Kitamura and K. Sugiyama, *J. Neurochem.*, 2001, **78**, 64; (*f*) M. Saito, H. Kitamura and K. Sugiyama, *Biochim. Biophys. Acta, Gen. Subj.*, 2002, **1571**, 18.
8 H. Yu, V. Thon, K. Lau, L. Cai, Y. Chen, S. Mu, Y. Li, P. G. Wang and X. Chen, *Chem. Commun.*, 2010, **46**, 7507.
9 H. Yu, H. A. Chokhawala, S. Huang and X. Chen, *Nat. Protoc.*, 2006, **1**, 2485.
10 H. Yu, H. Chokhawala, R. Karpel, H. Yu, B. Wu, J. Zhang, Y. Zhang, Q. Jia and X. Chen, *J. Am. Chem. Soc.*, 2005, **127**, 17618.
11 Y. Li, H. Yu, H. Cao, K. Lau, S. Muthana, V. K. Tiwari, B. Son and X. Chen, *Appl. Microbiol. Biotechnol.*, 2008, **79**, 963.
12 H. Yu, H. Yu, R. Karpel and X. Chen, *Bioorg. Med. Chem.*, 2004, **12**, 6427.
13 H. Yu, H. A. Chokhawala, A. Varki and X. Chen, *Org. Biomol. Chem.*, 2007, **5**, 2458.
14 V. Padler-Karavani, H. Yu, H. Cao, H. A. Chokhawala, F. Karp, N. Varki, X. Chen and A. Varki, *Glycobiology*, 2008, **18**, 818.
15 S. Huang, H. Yu and X. Chen, *Sci. China Chem.*, 2011, **54**, 117.
16 (*a*) M. Izumi, G.-J. Shen, S. Wacowich-Sgarbi, T. Nakatani, O. Plettenburg and C.-H. Wong, *J. Am. Chem. Soc.*, 2001, **123**, 10909; (*b*) V. Kren and J. Thiem, *Angew. Chem., Int. Ed. Engl.*, 1995, **34**, 893; (*c*) H. Paulsen and U. Von Deessen, *Carbohydr. Res.*, 1988, **175**, 283.
17 (*a*) S.-G. Lee and B.-G. Kim, *Enzyme Microb. Technol.*, 2001, **28**, 161; (*b*) A. Vetere, M. Miletich, M. Bosco and S. Paoletti, *Eur. J. Biochem.*, 2000, **267**, 942; (*c*) A. Lubineau, C. Auge and P. Francois, *Carbohydr. Res.*, 1992, **228**, 137; (*d*) J. Xia, J. L. Alderfer, C. F. Piskorz, R. D. Locke and K. L. Matta, *Carbohydr. Res.*, 2000, **328**, 47.