Graphical Abstract

Donor substrate promiscuity of bacterial β 1–3-*N*-acetylglucosaminyltransferases and acceptor substrate flexibility of β 1–4-galactosyltransferases

Yanhong Li^{a,†}, Mengyang Xue^{a,b,c,†}, Xue Sheng^{a,d}, Hai Yu^a, Jie Zeng^{a,e}, Vireak Thon^{a,f,g}, Yi Chen^a, Musleh M. Muthana^{a,h}, Peng G. Wang^{b,f}, Xi Chen^{a,*}

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

Leave this area blank for abstract info.

Donor substrate promiscuity of bacterial β 1–3-*N*-acetylglucosaminyltransferases and acceptor substrate flexibility of β 1–4-galactosyltransferases

Yanhong Li^{a,†}, Mengyang Xue^{a,b,c,†}, Xue Sheng^{a,d}, Hai Yu^a, Jie Zeng^{a,e}, Vireak Thon^{a,f,g}, Yi Chen^a, Musleh M. Muthana^{a,h}, Peng G. Wang^{b,f}, Xi Chen^{a,*}

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

^bNational Glycoengineering Research Center and Shandong Province Key Laboratory of Carbohydrate Chemistry and Glycobiology, Shandong University, Jinan, Shandong 250100, China

^cCurrent address: Shandong Province Key Laboratory of Emergency and Critical Care Medicine, Shandong University, Jinan, Shandong 250100, China

^dCurrent address: Department of Chemistry and Biochemistry, University of California, San Diego, CA 92093, USA

^eSchool of Food Science, Henan Institute of Science and Technology, Xinxiang, Henan 453003, China

^fDepartment of Chemistry, Georgia State University, Atlanta, GA 30303, USA

^gCurrent address: Laboratory of Bacterial Polysaccharides, Food and Drug Administration, Bethesda, MD 20892, USA

^hCurrent address: Children's National Medical Center, Washington DC, 20010, USA

ARTICLE INFO

Article history:

Received

Received in revised form

Accepted

Available online

Keywords:

Carbohydrate

Glycosyltransferase

Donor substrate promiscuity

Acceptor substrate specificity

Enzymatic synthesis

One-pot multienzyme

ABSTRACT

β 1–3-*N*-Acetylglucosaminyltransferases (β 3GlcNAcTs) and β 1–4-galactosyltransferases (β 4GalTs) have been broadly used in enzymatic synthesis of *N*-acetylglucosamine (GlcNAc)-containing oligosaccharides and glycoconjugates including poly-LacNAc, and lacto-*N*-neotetraose (LNnT) found in the milk of human and other mammals. In order to explore oligosaccharides and derivatives that can be synthesized by the combination of β 3GlcNAcTs and β 4GalTs, donor substrate specificity studies of two bacterial β 3GlcNAcTs from *Helicobacter pylori* (Hp β 3GlcNAcT) and *Neisseria meningitidis* (NmLgtA), respectively, using a library of 39 sugar nucleotides were carried out. The two β 3GlcNAcTs have complementary donor substrate promiscuity and 13 different trisaccharides were produced. They were used to investigate the acceptor substrate specificities of three β 4GalTs from *Neisseria meningitidis* (NmLgtB), *Helicobacter pylori* (Hp β 4GalT), and bovine (B β 4GalT), respectively. Ten of the 13 trisaccharides were shown to be tolerable acceptors for at least one of these β 4GalTs. The application of NmLgtA in one-pot multienzyme (OPME) synthesis of two trisaccharides including GalNAc β 1–3Gal β 1–4Glc β MU and Gal β 1–3Gal β 1–4Glc was demonstrated. The study provides important information for using these glycosyltransferases as powerful catalysts in enzymatic and chemoenzymatic syntheses of oligosaccharides and derivatives which can be useful probes and reagents.

2015 Elsevier Ltd. All rights reserved.

1. Introduction

β 1–3-*N*-Acetylglucosaminyltransferases (β 3GlcNAcTs) and β 1–4-galactosyltransferases (β 4GalTs) have been broadly used in enzymatic synthesis of Gal β 1–4GlcNAc β 1–3Gal-containing oligosaccharides and glycoconjugates such as poly-*N*-acetylglucosamine [poly-LacNAc, (-3Gal β 1–4GlcNAc β 1-)_n]-containing structures^{1–3} with or without additional fucosylation and/or sialylation^{4–10} as well as lacto-*N*-neotetraose (LNnT, Gal β 1–4GlcNAc β 1–3Gal β 1–4Glc)^{11–13} and longer LacNAc-containing oligosaccharides¹³ found in the milk of human and other mammals. Both mammalian and bacterial glycosyltransferases have been used in the synthesis of these structures.^{6, 8, 13–15}

For β 3GlcNAcTs, mammalian enzymes were found to be quite specific on the acceptor substrates.^{1, 15, 16} In comparison, recombinant β 3GlcNAcT from *Neisseria meningitidis* (NmLgtA) was shown to be more promiscuous and was able to use both α - and β -galactosides as acceptors.^{15, 17} It was found to have low¹⁵ or no activity towards type 1 (Gal β 1–3GlcNAc) or type 3 (Gal β 1–3GalNAc α) galactoside acceptors.^{1, 6} Branched glycans or longer LacNAc repeats were less efficient acceptors for NmLgtA. More recently, a β 3GlcNAcT from *Helicobacter pylori* (Hp β 3GlcNAcT) was cloned and characterized.^{1, 18} Hp β 3GlcNAcT was shown to be specific for the β -linked galactoside acceptors¹⁴ and prefer type 2 (Gal β 1–4GlcNAc) acceptors but can also use type 1 (Gal β 1–3GlcNAc), type 3 (Gal β 1–3GalNAc α), and type 4 (Gal β 1–3GalNAc β) acceptors on both linear and branched glycan structures.¹ Regarding donor substrate specificity, only a limited number of UDP-sugars have been tested for NmLgtA which was shown to tolerate uridine 5'-

*Corresponding author. Tel: 530-754-6037; fax: 530-752-8995.

E-mail: xiichen@ucdavis.edu

† These authors contributed equally.

diphosphate-*N*-acetylglucosamine (UDP-GlcNAc), uridine 5'-diphosphate-*N*-acetylgalactosamine (UDP-GalNAc), UDP-4-deoxy-GalNAc, UDP-6-deoxy-GalNAc/GlcNAc, and UDP-*N*-propanoylglucosamine.^{15, 19} The donor substrate specificity of Hpβ3GlcNAcT has not been explored in detail.

Bovine β4GalT 1 (Bβ4GalT) is the most well studied mammalian glycosyltransferase that has been used for biosynthesis.^{20, 21} In addition, β4GalTs have been cloned from *Neisseria meningitidis*, *Neisseria gonorrhoeae*, and *Helicobacter pylori*.²¹⁻²⁵ *Neisseria meningitidis* β4GalT (NmLgtB) and *Helicobacter pylori* β4GalT (Hpβ4GalT), and Bβ4GalT have been used broadly for enzymatic and chemoenzymatic synthesis of β1–4-linked galactosides.^{13, 21, 26-30} NmLgtB has been fused with a *Streptococcus thermophilus* UDP-galactose 4-epimerase (NmLgtB-StGalE) to allow the use of a less expensive sugar nucleotide uridine 5'-diphosphate glucose (UDP-Glc) as the starting material which can be converted to uridine 5'-diphosphate galactose (UDP-Gal) by StGalE for NmLgtB-catalyzed synthesis of β1–4-linked galactosides.³¹ Although the tolerance of the modifications on the acceptor substrates of β-4GalTs has been tested using monosaccharide derivatives,^{21, 32} their tolerance of modifications on longer oligosaccharides such as trisaccharides is less clear.

In order to explore the scope of oligosaccharide derivatives that can be synthesized by the combination of β3GlcNAcTs and β4GalTs, herein we report the donor substrate specificities of NmLgtA and Hpβ3GlcNAcT using 4-methylumbelliferyl (MU) β-lactoside Galβ1–4GlcβMU (or LacβMU) as an acceptor and a library of 39 sugar nucleotides including 37 UDP-sugars, cytidine 5'-monophosphate *N*-acetylneuraminic acid (CMP-Neu5Ac), and guanosine 5'-diphosphate L-fucose (GDP-Fuc) as potential donors. Thirteen trisaccharides have been successfully obtained by NmLgtA or Hpβ3GlcNAcT-catalyzed reactions and have been used to investigate the acceptor substrate specificities of NmLgtB, Hpβ4GalT, and Bβ4GalT (Scheme 1). We have also demonstrated the application of NmLgtA in one-pot multienzyme (OPME) preparative-scale synthesis of two trisaccharides including GalNAcβ1–3Galβ1–4GlcβMU and Galβ1–3Galβ1–4Glc from simple monosaccharides and suitable acceptors. The study provides important information for choosing suitable catalysts in enzymatic and chemoenzymatic synthesis of desired glycans and glycoconjugates that can be useful probes or reagents.

Scheme 1. Donor substrate specificity and acceptor substrate specificity studies of β3GlcNAcTs and β4GalTs, respectively. LacβMU was used as an acceptor for β3GlcNAcTs (NmLgtA or Hpβ3GlcNAcT) and the resulting trisaccharide products were used as acceptor substrates for β4GalTs (NmLgtB, Hpβ4GalT, or Bβ4GalT).

2. Results and discussion

2.1. Cloning and expression of β3GlcNAcTs and β4GalTs

Both NmLgtA¹⁵ and Hpβ3GlcNAcT¹ have been cloned and expressed in *Escherichia coli* previously. In our attempts, initial cloning into pET15b vectors led to a low expression of soluble and active enzymes in *Escherichia coli* BL21(DE3) cells. In

order to improve the protein expression, NmLgtA and Hpβ3GlcNAcT were cloned into pMAL-c2X and pMAL-c4X vectors, respectively, both as N-terminal maltose-binding protein (MBP)-fused and C-terminal hexa-histidine (His₆)-tagged recombinant proteins. Soluble and active enzymes could be obtained by inducing *Escherichia coli* BL21(DE3) cells with 0.1 mM of isopropyl β-D-1-thiogalactopyranoside (IPTG) followed by incubation at 20 °C for 20 hours. Purification was achieved by one-step nickel-nitrilotriacetic acid (Ni²⁺-NTA) affinity chromatography. About 10 mg of NmLgtA and only about 1 mg of Hpβ3GlcNAcT could be obtained from 1 liter of *Escherichia coli* cell culture. The differences of their expression levels could be contributed by their protein sequence differences as the two share only 16% protein sequence similarity (<http://imed.med.ucm.es/Tools/sias.html>) (Fig. 1A). Sodium dodecylsulfate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis indicated that the apparent molecular weights of purified Hpβ3GlcNAcT (Fig. 1B) and NmLgtA (Fig. 1C) were about 88 kDa and 76 kDa respectively, which were consistent with their calculated molecular weights of 88.9 kDa and 76.6 kDa. Recombinant Bβ4GalT was usually expressed in *Escherichia coli* as inclusion bodies and active soluble enzyme was obtained after proper refolding procedures.^{33, 34} We found that when cloned in vector plasmid pET15b as an N-His₆-tagged fusion protein with an N-terminal 128 amino acid truncation, soluble and active Bβ4GalT was readily obtained without refolding although its expression level (< 1 mg L⁻¹ culture) was low (Fig. S1). In comparison, NmLgtB and Hpβ4GalT (previously named as Hp1–4GalT) were obtained at expression levels of 12.5 mg and 5 mg per liter of culture, respectively, as reported previously.²¹

A.

NmLgtA	-----MDEHVSVLICRYNVEVFAQSIAA	24
Hpβ3GlcNAcT	MTSASSHSFKEQDFHPIAFADFNKYLIPAGACYSLESTIARFNKKIRPTLHATUVV	57
NmLgtA	VVCTWRN-----LDILIVDDGSTDGTATACRFCECGRIFILACFRNSGLI	72
Hpβ3GlcNAcT	GLNEEDTKTKNQITEPFKEFAVEVKDIEFEDTIENPFDEETFSKSNVLEVKYEL	114
NmLgtA	PSNIGLDELAFSGGGGEYIARTEDADIAAPWIBK-IUGEMKRSITAMCAWLEV	128
Hpβ3GlcNAcT	ATTFPKYSKRMVRSQVETVFCNEESALFLNKEINENYEVIEVEKHHMMECFLECN	171
NmLgtA	ISEEKDGNRIARSHHGHG-----IMKKPTRHEDIADEFFGNHNNNTM	172
Hpβ3GlcNAcT	LDYQRRKNTLIMHMLLNGNEARGELDFPKWCNPNMKALGIEYCVFYTYTKDFSN	228
NmLgtA	IMRRSVIDGGLRYNTERNAEFCQHYVEVKEGLRAY-YEPFVHYTHANQVSSGV	228
Hpβ3GlcNAcT	AVINENYKKTILEARSPPTIIPGHWGAVPWYEPFGLKALWLNHAKTEFFKGL	285
NmLgtA	SIRCEPIAG-YCTDNRNDFLCMG-----PKTRFDSLEYRCIKAVAYELLEKHLP	278
Hpβ3GlcNAcT	TKKMTNESFYTTMRECHYESSVKSSKEIVKAPYLFKFSYLEVVEKERRIHSRVF	342
NmLgtA	SEDFDARRRLYCQFRRPTETPAGAWLEFAADG-SMRRTFLRCYQFGLHRLNNR	333
Hpβ3GlcNAcT	ELTGNVYKRFNNLLIEVFLEKALVLRVRSKILVLCFPHGIVKRIILRLRLRRRG	398

B.

C.

Figure 1. Protein sequence alignment (A) and SDS-PAGE analysis (8% Tris-Glycine gel) of Hpβ3GlcNAcT (B) and NmLgtA (C). Lanes: BI, whole cell extract before induction; AI, whole cell extract after induction; L, lysate after induction; P, Ni²⁺-NTA column purified protein; M, protein markers (Bio-Rad precision Plus Protein Standards, 10–250 kDa).

2.2. pH Profile and metal effect of β3GlcNAcTs

When 4-methylumbelliferyl β-lactoside (LacβMU) was used as the acceptor, the observed optimal pH for both NmLgtA and Hpβ3GlcNAcT was 8.0 (Fig. 2). The slight different optimal pH observed for NmLgtA compared to the previously reported pH

Table 1. Donor substrate specificities of Hpβ3GlcNAcT and NmLgtA analyzed by HRMS and CE or HPLC. LacβMU was used as the acceptor. Reactions were carried out at 37 °C for 24 hours.

#	Donors	NmLgtA		Hpβ3GlcNAcT	
		HRMS ^a	CE ^{b,c} or HPLC ^d	HRMS ^a	CE ^b or HPLC ^d
1	UDP-GlcNAc	90–95%	81.2±0.8% ^b	55–60%	66.2±1.3% ^b
2	UDP-GlcNTFA	100%	100±0% ^b	90–95%	92.9±0.3% ^b
3	UDP-GlcNAcN ₃	25–35%	29.1±0.1% ^b	20–30%	26.3±0.4% ^b
4	UDP-GlcNGc	30–40%	36.1±0.8% ^b	55–60%	77.5±0.1% ^b
5	UDP-GlcNAcPh	20–25%	12.9±0.5% ^b	0	0
6	UDP-GlcNAcPhPh	<5%	N.D.	0	0
8	UDP-GlcNH ₂	20–25%	peaks together ^d	0	0
10	UDP-GlcNAc6NH ₂	20–25%	peaks together ^d	70–75%	peaks together ^d
11	UDP-GlcNAc6N ₃	0	0	30–35%	27.8±1.1% ^b
16	UDP-GlcNAc6S	0	0	65–70%	peaks together ^d
21	UDP-Glc	20–25%	23.2±2.6% ^b	0	0
23	UDP-GalNAc	80–90%	peaks together ^d	0	0
25	UDP-Gal	40–45%	39.6±0.6% ^c	0	0

^a Yields estimated by HRMS.

^b Quantitative yields obtained by CE using sodium tetraborate (25 mM, pH 9.4) as the running buffer.

^c Quantitative yields obtained by CE using sodium tetraborate (50 mM, pH 11.0) as the running buffer.

^d Results obtained by HPLC. The acceptor and the product peaks had the same retention time and could not be separated.

N.D.: not detected by CE or HPLC.

7.5¹⁵ may be caused by the differences of the buffers (Tris-HCl buffer for this study versus sodium cacodylate buffer used before) used. Both enzymes retain activities in a pH range of 6.0–9.0 although at pH 10.0, only NmLgtA was active.

Figure 2. pH profiles of NmLgtA (filled diamond) and Hpβ3GlcNAcT (open square). Buffers used were: MES, pH 5.0–6.5; Tris-HCl, pH 7.0–9.0; CAPS, pH 10.0–11.0.

The effects of several divalent metal ions including Mn²⁺, Mg²⁺, and Ca²⁺, as well as ethylenediaminetetraacetic acid (EDTA) on the activity of both β3GlcNAcTs were investigated. As shown in Fig. S2, both Hpβ3GlcNAcT (Fig. S2A) NmLgtA (Fig. S2B) required a divalent metal ion for activity as adding 10 mM of EDTA completely abolished their activities. Hpβ3GlcNAcT exhibited optimal activity in the presence of Mg²⁺. Increasing the concentration of Mg²⁺ from 1 mM to 10 mM did not affect the activity of Hpβ3GlcNAcT significantly, while its activity dropped slightly when the concentration of Mg²⁺ increased from 10 mM to 20 mM. Mn²⁺ was shown to be a less efficient cofactor for Hpβ3GlcNAcT and increasing the concentration of Mn²⁺ from 1 mM to 20 mM caused gradual decrease of Hpβ3GlcNAcT activity. Ca²⁺ was shown to be a less

efficient cofactor for Hpβ3GlcNAcT compared to Mg²⁺ and its concentration ranging from 1–20 mM did not affect the Hpβ3GlcNAcT activity significantly. Unlike Hpβ3GlcNAcT, NmLgtA exhibited the best activity in the presence of Mn²⁺ with an optimal activity observed in the presence of 10 mM of Mn²⁺. Increasing or decreasing the concentrations of Mn²⁺ led to the decrease of NmLgtA activity. Although a less effective cofactor compared to Mn²⁺ at concentrations at or lower than 10 mM, increasing the concentration of Mg²⁺ from 1 mM to 20 mM led to a gradual increase of NmLgtA activity. Ca²⁺ was not a suitable cofactor for NmLgtA. The results of the metal effect of NmLgtA were consistent to those reported previously.¹⁵

2.3. Donor Substrate Specificities of β3GlcNAcTs

Once the optimal conditions were identified for Hpβ3GlcNAcT and NmLgtA, donor substrate specificity studies were carried out using a library of 39 sugar nucleotides (Fig. S3) including 37 UDP-sugars as well as CMP-Neu5Ac and GDP-Fuc. UDP-sugars used in the assays included UDP-GlcNAc (1) and its 2'-, 3'-, and/or 6'-modified derivatives (2–20), UDP-glucose (UDP-Glc) (21), UDP-glucuronic acid (UDP-GlcA) (22), UDP-N-acetylgalactosamine (UDP-GalNAc) (23) and its 6'-modified derivative (24), UDP-galactose (UDP-Gal) (25) and its 6'-modified derivative (26), UDP-galacturonic acid (UDP-GalA) (27), UDP-N-acetylmannosamine (UDP-ManNAc) (28) and its 2'-modified derivatives (29–35), UDP-mannose (36) and its 2'-modified derivative (37), as well as CMP-Neu5Ac (38), and GDP-Fuc (39) (Fig. S3). The reactions were analyzed by capillary electrophoresis (CE) equipped with a photodiode array (PDA) detector and were monitored at 315 nm to determine the ratio of the UV active fluorescent trisaccharide product formed and the fluorescent LacβMU acceptor. High performance liquid chromatography (HPLC) was also used to assist the analysis for some of the substrates. The presence of each trisaccharide product was further confirmed by high resolution mass spectrometry (HRMS).

Table 2. Apparent kinetic data for Hpβ3GlcNAcT and NmLgtA.

Substrate	NmLgtA			Hpβ3GlcNAcT		
	k_{cat} (s ⁻¹)	K_M (mM)	k_{cat}/K_M (s ⁻¹ mM ⁻¹)	k_{cat} (s ⁻¹)	K_M (mM)	k_{cat}/K_M (s ⁻¹ mM ⁻¹)
UDP-GlcNAc (1)	0.55 ± 0.13	0.47 ± 0.11	1.16	0.46 ± 0.08	1.66 ± 0.28	0.68
UDP-GlcNTFA (2)	0.19 ± 0.06	0.59 ± 0.20	0.32	0.32 ± 0.08	1.23 ± 0.31	0.26
UDP-GlcNAcN ₃ (3)	0.15 ± 0.03	1.34 ± 0.24	0.11	0.29 ± 0.07	4.74 ± 1.09	0.062

As shown in Table 1, both β3GlcNAcTs exhibited promiscuous and somewhat complementary substrate specificities. NmLgtA was able to use 11 UDP-sugars and Hpβ3GlcNAcT could take 7 UDP-sugars tested as donor substrates. The estimated yields based on high-resolution mass spectrometry (HRMS) matched well with the capillary electrophoresis and/or high performance liquid chromatography (HPLC) results. For the reaction using UDP-Gal (25), the product and the acceptor peaks had the same retention time and could not be separated by CE using sodium tetraborate (25 mM, pH 9.4) as a running buffer but could be resolved by CE using sodium tetraborate (50 mM, pH 11.0) as a running buffer. For reactions using compounds 8, 10, 16, and 23, neither CE nor HPLC could separate the acceptor and the product peaks. However, the product formation was verified by HRMS analysis. In addition to UDP-GlcNAc (1), both β3GlcNAcTs could use 2'-modified UDP-GlcNAc derivatives including UDP-GlcNTFA (2), UDP-GlcNAcN₃ (3), and UDP-GlcNGc (4) as well as a 6'-modified UDP-GlcNAc derivative UDP-GlcNAc6NH₂ (10) as donor substrates. UDP-GlcNTFA (2) containing an *N*-trifluoroacetyl group was an excellent donor substrate for both β3GlcNAcTs. In contrary to a previous report indicating UDP-GlcNAcN₃ (3) containing an *N*-azidoacetyl group as an intolerable donor for NmLgtA,¹⁹ the compound was confirmed by both HRMS and CE assays to be an acceptable donor substrate for both β3GlcNAcTs although the yields were not high.

Quite interestingly, two β3GlcNAcTs showed some complementary donor substrate specificities. For example, NmLgtA could tolerate UDP-GlcNAc derivatives with a small or a large substituent at C-2 of the GlcNAc including UDP-GlcNAcPh (5) with an *N*-phenylacetyl group, UDP-GlcNAcPhPh (6) with an *N*-diphenylacetyl group, and UDP-GlcNH₂ (8) with an amino group. However, these UDP-sugars were not acceptable donor substrates for Hpβ3GlcNAcT. On the other hand, 6'-modifications of UDP-GlcNAc were more tolerable by Hpβ3GlcNAcT compared to NmLgtA. For example, UDP-GlcNAc6NH₂ (10) containing a 6'-amino group was a better donor substrate for Hpβ3GlcNAcT and the reaction was achieved with a higher yield than that with NmLgtA. In addition, UDP-GlcNAc6N₃ (11) containing an 6'-azido group and UDP-GlcNAc6S (16) containing an 6'-*O*-sulfate group were suitable donor substrates for Hpβ3GlcNAcT, but not for NmLgtA. It was quite exciting to notice the well tolerance of UDP-GlcNAc6S (16) which contains a relatively large and negatively charged *O*-sulfate group at C-6 of GlcNAc as the donor substrate by Hpβ3GlcNAcT. In consistent with previous reports,^{15, 19} UDP-GalNAc (23) was an excellent donor substrate for NmLgtA. The compound, however, was not a donor for Hpβ3GlcNAcT. To our surprise, unlike a previous report indicating both UDP-Glc (21) and UDP-Gal (25) were not donor substrate for NmLgtA,¹⁵ our results showed that both of these UDP-sugars were suitable donor substrates for NmLgtA but not for Hpβ3GlcNAcT.

Neither β3GlcNAcTs could use UDP-GlcA (22), UDP-GalNAc (24), UDP-Gal (26), UDP-GalA (27), UDP-ManNAc (28) or its derivatives (29–35), UDP-mannose (36), UDP-mannose (37), CMP-Neu5Ac (38), and GDP-Fuc (39) (Fig. S3) as donor substrates under the reaction conditions used.

2.4. Kinetic studies of β3GlcNAcTs

Due to their potential in synthetic applications and chemical biological studies, three UDP-sugars including UDP-GlcNAc (1), UDP-GlcNTFA (2), and UDP-GlcNAcN₃ (3) that could be used by both β3GlcNAcTs were chosen for kinetic studies. As shown in Table 2, in general the catalytic efficiency (k_{cat}/K_M) of NmLgtA was higher than that of Hpβ3GlcNAcT no matter which substrate among the three was used. Compared to UDP-GlcNTFA (2), the catalytic efficiencies of both β3GlcNAcTs were higher (~3-fold) when UDP-GlcNAc (1) was used as the donor substrate mainly due to the higher k_{cat} values observed for UDP-GlcNAc (1). Compared to UDP-GlcNTFA (2), the catalytic efficiencies of both β3GlcNAcTs were lower (3–4-fold) when UDP-GlcNAcN₃ (3) was used as the donor substrate mainly due to the higher K_M values observed for UDP-GlcNAcN₃ (3).

2.5. Acceptor substrate specificities of β4GalTs

Thirteen trisaccharides could be obtained by reactions catalyzed by NmLgtA and/or Hpβ3GlcNAcT. They were used to test the acceptor substrate specificities of three β4GalTs including NmLgtB, Hpβ4GalT, and Bβ4GalT with UDP-Gal as the donor substrate. Protein sequence alignment (Fig. S4) showed that these enzymes share very low sequence identity. For example, 7.3% sequence identity was found for Hpβ4GalT and NmLgtB, 6.2% for Bβ4GalT and Hpβ4GalT, and 2.9% for NmLgtB and Bβ4GalT. The overall sequence similarity among the three was less than 16%.

Despite their low sequence homology, NmLgtB, Hpβ4GalT, and Bβ4GalT are effective β4GalTs and can use several trisaccharides and derivatives produced by β3GlcNAcT-catalyzed reactions as acceptors. As shown in Table 3, ten of the 13 trisaccharides could be used as acceptors by either NmLgtB or Hpβ4GalT or both. Bβ4GalT could use only four as acceptors (HRMS data are shown in Fig. S5). Besides GlcNAc-containing trisaccharide GlcNAcβ1–3LacβMU (1A), all three β4GalTs could use 2'-modified GlcNAc-containing trisaccharides GlcNTFAβ1–3LacβMU (2A), GlcNAcN₃β1–3LacβMU (3A), and GlcNAcPhβ1–3LacβMU (5A) as acceptors. While GlcNTFA-containing trisaccharide GlcNTFAβ1–3LacβMU (2A) was well tolerated by Bβ4GalT, other 2'-modifications on GlcNAc decrease its activity more significantly. 6'-Modification completely abolished its activity. While not acceptable by Bβ4GalT, GlcNGcβ1–3LacβMU (4A) was an excellent acceptor for both NmLgtB and Hpβ4GalT.

In consistent with those observed previously using monosaccharide derivatives as acceptors,²¹ NmLgtB was more promiscuous towards acceptors containing a terminal 2'-modified GlcNAc, while Hpβ4GalT was more tolerable towards acceptors containing a terminal 6'-modified GlcNAc. For example, trisaccharide GlcNAcPhPhβ1–3Galβ1–4GlcβMU (6A) containing a terminal *N*-diphenylacetylglucosamine (GlcNAcPhPh) was a suitable acceptor for NmLgtB but not for Hpβ4GalT. On the other hand, GlcNAc6NH₂β1–3LacβMU (10A) and GlcNAc6Sβ1–3LacβMU (16A) could be used as acceptors by Hpβ4GalT but not NmLgtB. In addition, a high yield (100%) was achieved when GlcNAc6N₃β1–3LacβMU (11A) was used as an acceptor for Hpβ4GalT compared to a low yield (5–10%) for NmLgtB. Quite excitingly, the Glcβ1–3LacβMU (21A) obtained by NmLgtA-catalyzed reaction can be

Table 3. Acceptor substrate specificities of β 4GalTs using a sequential two-step process.^a

#	Acceptor	β 3GlcNAcT*	NmLgtB	Hp β 4GalT	B β 4GalT
1A	GlcNAc β 1-3Lac β MU	90–95%	NmLgtA	100%	90–95%
2A	GlcNTFA β 1-3Lac β MU	100%	NmLgtA	95–100%	60–65%
3A	GlcNAcN ₃ β 1-3Lac β MU	25–35%	NmLgtA	100%	35–40%
4A	GlcNGc β 1-3Lac β MU	55–60%	Hp β 3GlcNAcT	90–95%	0
5A	GlcNAcPh β 1-3Lac β MU	20–25%	NmLgtA	100%	20–25%
6A	GlcNAcPhPh β 1-3Lac β MU	<5%	NmLgtA	0	0
8A	GlcNH ₂ β 1-3Lac β MU	20–25%	NmLgtA	0	0
10A	GlcNAc6NH ₂ β 1-3Lac β MU	70–75%	Hp β 3GlcNAcT	0	0
11A	GlcNAc6N ₃ β 1-3Lac β MU	30–35%	Hp β 3GlcNAcT	5–10%	0
16A	GlcNAc6S β 1-3Lac β MU	65–70%	Hp β 3GlcNAcT	0	0
21A	Glc β 1-3Lac β MU	20–25%	NmLgtA	95–100%	0
23A	GalNAc β 1-3Lac β MU	80–90%	NmLgtA	0	0
25A	Gal β 1-3Lac β MU	40–45%	NmLgtA	0	0

^a Yields estimated by HRMS. R = Lac β MU.* Lac β MU was used as the acceptor for β 3GlcNAcT-catalyzed reactions. The trisaccharide producing reactions with a higher yield (catalyzed by either Hp β 3GlcNAcT or NmLgtA) were used for the acceptor substrate specificity studies of β 1-4GalTs.

extended further by NmLgtB for the formation of presumably Lac β 1-3Lac β MU tetrasaccharide. It is interesting to notice that the trisaccharide derivatives formed by NmLgtA-catalyzed reaction were more tolerable acceptors by NmLgtB. Similarly, Hp β 3GlcNAcT and Hp β 4GalT seemed to also work as a good pair in tolerating modified glycan substrates.

2.6. Preparative-scale synthesis of GalNAc β 1-3Gal β 1-4Glc β ProN₃ and Gal β 1-3Gal β 1-4Glc

Taking advantage of the interesting tolerance of UDP-GalNAc and UDP-Gal as the donor substrates by NmLgtA, preparative-scale synthesis of trisaccharides GalNAc β 1-3Gal β 1-4Glc β ProN₃ (**1**) (Scheme 2A) and Gal β 1-3Gal β 1-4Glc (**2**) (Scheme 2B) were carried out.

GalNAc β 1-3Gal β 1-4Glc β OR component has been found in glycosphingolipids identified in rat granuloma, rat abdominal macrophage,³⁵ starfish *Asterias amurensis versicolor* Sladen.³⁶ As shown in Scheme 2A, GalNAc β 1-3Gal β 1-4Glc β ProN₃ (**1**) was synthesized from Gal β 1-4Glc β ProN₃ and simple free monosaccharide GalNAc using a one-pot four-enzyme (OP4E) system containing *Bifidobacterium longum* ATCC55813 *N*-acetylhexosamine 1-kinase (NahK),³⁷ *Pasteurella multocida* *N*-acetylglucosamine-1-phosphate uridylyltransferase (PmGlmU),³⁸ *Pasteurella multocida* inorganic pyrophosphatase (PmPpA),²¹ and NmLgtA. In this system, NahK was responsible for the formation of GalNAc-1-phosphate from GalNAc and adenosine 5'-triphosphate (ATP), PmGlmU catalyzed the formation of UDP-GalNAc which was used by NmLgtA to catalyze the transfer of GalNAc to the acceptor for the formation of the trisaccharide product. PmPpA was used to break down the pyrophosphate formed in the UDP-GalNAc production step to drive the reaction towards the formation of the sugar nucleotide. This OP4E system was quite efficient in synthesizing GalNAc β 1-3Gal β 1-4Glc β ProN₃ (**1**) and the yield was quantitative. Compared to a previous report,¹⁵ the OP4E process here avoided the use of relatively expensive UDP-GalNAc and allowed the production of the desired glycan in a high-yield more economic process. As shown in Table S1, comparing the ¹³C nuclear magnetic resonance (NMR) chemical shifts of the

disaccharide Lac β ProN₃ and the trisaccharide product indicated that significant changes occurred at C-3 of internal Gal residue (a downfield shift of 9.26 ppm from 72.37 ppm in the disaccharide starting material to 81.63 ppm in the trisaccharide product). This confirmed that the glycosylation took place at C-3 of the Gal residue of Lac β ProN₃. The anomeric proton of terminal GalNAc in the trisaccharide product had a chemical shift at 4.58 ppm as a doublet with a coupling constant $J_{1,2}$ = 8.8 Hz indicates the formation of a β -linkage. These results confirmed the structure of the trisaccharide product as GalNAc β 1-3Gal β 1-4Glc β ProN₃ (**1**).

Scheme 2 One-pot four-enzyme synthesis of GalNAc β 1-3Gal β 1-4Glc β ProN₃ (**23A**) (A) and Gal β 1-3Gal β 1-4Glc (**25A**) (B). Enzyme abbreviations: NahK, *Bifidobacterium longum* ATCC55813 *N*-acetylhexosamine 1-kinase;³⁷ PmGlmU, *Pasteurella multocida* *N*-acetylglucosamine-1-phosphate uridylyltransferase;³⁸ PmPpA, *Pasteurella multocida* inorganic pyrophosphatase;²¹ SpGalK, *Streptococcus pneumoniae*

TIGR4 galactokinase;³⁹ BLUSP, *Bifidobacterium longum* UDP-sugar pyrophosphorylase.⁴⁰

Gal β 1–3Gal β 1–4Glc (3'-galactosyllactose) has been found in human milk⁴¹ and eastern quoll (*Dasyurus viverrinus*) milk⁴² as well as in the colostrums of bovine,⁴³ ovine,⁴⁴ primate,⁴⁵ camel,⁴⁶ and horse.⁴⁷ It is also an abundant oligosaccharide in the milk of elephant⁴⁸ and marsupials such as brushtail possum (*Trichosurus vulpecula*),⁴⁹ tammar wallaby (*Macropus eugenii*) and grey kangaroo,⁵⁰ and koala (*Phascolarctos cinereus*).⁵¹ As shown in Scheme 2B, Gal β 1–3Gal β 1–4Glc (**2**) was synthesized from lactose and galactose (Gal) using a one-pot four-enzyme (OP4E) system containing a *Streptococcus pneumoniae* TIGR4 galactokinase (SpGalK),³⁹ a *Bifidobacterium longum* ATCC55813 UDP-sugar pyrophosphorylase (BLUSP),⁴⁰ PmPpA,²¹ and NmLgtA. SpGalK, BLUSP, and PmPpA were used previously for high efficient (85% yield) one-pot multienzyme (OPME) synthesis of UDP-Gal.⁴⁰ As UDP-Gal was a relatively weak substrate for NmLgtA, the synthetic yield (10%) of Gal β 1–3Gal β 1–4Glc (**25A**) was lower compared to the yield (40–45%) shown in Table 1, which may be caused by lactose as a less effective acceptor than the fluorophore-tagged beta-linked lactoside (Lac β MU) similar to that observed previously.¹⁵ Nevertheless, this is a tailor-made process of synthesizing 3'-galactosyllactose which is not readily available by reactions catalyzed by other glycosyltransferases. The chemical structure of the trisaccharide product was confirmed by NMR. As shown in Table S2, comparing the ¹³C NMR chemical shifts of the disaccharide lactose and the trisaccharide product indicated that significant changes occurred at C-3 of internal Gal residue (a downfield shift of 9.37 ppm from 72.38 ppm in the disaccharide starting material to 81.75 ppm in the trisaccharide product). This confirmed that the glycosylation took place at C-3 of the Gal residue of lactose. The anomeric proton of terminal Gal in the trisaccharide product had a chemical shift at 4.59 ppm as a doublet with a coupling constant $J_{1,2}$ = 8.0 Hz indicates the formation of a β -linkage. These results confirmed the structure of the trisaccharide product as Gal β 1–3Gal β 1–4Glc (**2**).

3. Conclusions

The donor substrate specificity of two β 1–3-*N*-acetylglucosaminyltransferases (Hp β 3GlcNAcT and NmLgtA) towards 39 sugar nucleotides and the acceptor substrate specificity of three β 1–4-galactosyltransferases (Hp β 4GalT, NmLgtB, and B β 4GalT) towards 13 trisaccharide acceptors formed by Hp β 3GlcNAcT and/or NmLgtA-catalyzed reactions were investigated for the synthesis of diverse LacNAc-containing glycans or derivatives. Compared to Hp β 3GlcNAcT, NmLgtA could tolerate more 2'-modified UDP-GlcNAc derivatives including those with large groups such as phenyl group and even diphenyl group or small groups such as NH₂. Besides UDP-GlcNAc and its derivatives, UDP-GalNAc, UDP-Glc, and UDP-Gal were also tolerable donor substrates for NmLgtA but not for Hp β 3GlcNAcT. However, for 6'-modified UDP-GlcNAc derivatives, Hp β 3GlcNAcT was a better catalyst than NmLgtA. Interestingly, enzymes from the same species can tolerate a similar set of modified substrates. For example, similarly to the tolerance of 6'-modified UDP-GlcNAc derivatives by Hp β 3GlcNAcT, β 4GalT from *H. pylori* (Hp β 4GalT) was the best catalyst in using 6'-modified GlcNAc derivatives as acceptors. Also, GlcNAcPhPh- and Glc-containing trisaccharides were only acceptable substrates for NmLgtB but not for Hp β 4GalT. In addition, preparative-scale enzymatic synthesis of trisaccharides GalNAc β 1–3Gal β 1–4Glc β ProN₃ and Gal β 1–3Gal β 1–4Glc from simple monosaccharides GalNAc and Gal, respectively, were successfully achieved using NmLgtA-containing one-pot four-

enzyme (OP4E) systems. The substrate promiscuity of β 3GlcNAcTs and β 4GalTs will allow efficient chemoenzymatic synthesis of diverse derivatives of LacNAc-containing structures as probes for protein binding study and for developing potential therapeutics.

4. Experimental section

4.1. Materials

Vector plasmids pMAL-c4X and pMAL-c2X were purchased from Amersham Biosciences (Piscataway, NJ, USA). Restriction enzymes *Eco*RI and *Sal*I were purchased from New England Biolabs, Inc. (Ipswich, MA, USA). Herculase II DNA polymerase was purchased from Stratagene (La Jolla, CA, USA). T4 DNA ligase and 1 kb DNA ladder were bought from Promega (Madison, WI, USA). QIAprep spin miniprep kit, QIAEX II gel extraction kit, and Nickel-nitrilotriacetic acid agarose (Ni²⁺-NTA agarose) were from Qiagen (Valencia, CA, USA). *Escherichia coli* electrocompetent DH5 α and chemically competent BL21 (DE3) cells were purchased from Invitrogen (Carlsbad, CA, USA). **Sugar nucleotides and derivatives were enzymatically or chemoenzymatically synthesized previously.**^{38, 40, 52–55}

4.2. Cloning of MBP-Hp β 3GlcNAcT-His₆ and MBP-NmLgtA-His₆

MBP-Hp β 3GlcNAcT-His₆ (GenBank accession number: NP_223749) was cloned from *Helicobacter pylori* J99 genomic DNA into pMAL-c4X vector as a fusion protein with N-terminal maltose-binding protein (MBP) and C-terminal His₆-tag. Primers used were: forward: 5'-CCGGAATTCATGACTTCAGCTTCAAGCCATTCTTTTAAAG-3' (*Eco*RI restriction site is underlined) and reverse: 5'-ACGCGTCGACTCAGTGGTGGTGGTGGTGGTGGCCCTTTTCAAGAGCTTGAGTT-3' (*Sal*I restriction site is underlined). MBP-NmLgtA-His₆ (GenBank accession number: U25839) was cloned from *Neisseria meningitidis* MC58 genomic DNA into pMAL-c2X vector as a fusion protein with N-terminal maltose-binding protein (MBP) and C-terminal His₆-tag. Primers used for polymerase chain reaction (PCR) are: Forward: 5'-GACCGAATTCATGCAGCCTTTAGTCAGCGTATTG-3' (*Eco*RI restriction site is underlined) and reverse: 5'-CAGCAAGCTTTTAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTTCAGCAATCGGTGCAAAATG-3' (*Hind*III restriction site is underlined). Polymer chain reactions (PCRs) were carried out in a reaction mixture (50 μ L) containing 10 $\times$ Herculase buffer (5 μ L), the genomic DNA (1 μ g), forward and reverse primers (1 μ M each), dNTP mixture (1 mM), and 5 U (1 μ L) of Herculase II DNA polymerase. The reaction mixture was subjected to 30 cycles of normal PCR amplification as followed: denaturation at 95 $^{\circ}$ C for 30 s, annealing at 50 $^{\circ}$ C for 45 s, elongation at 72 $^{\circ}$ C for 2 min. Followed by a final elongation at 72 $^{\circ}$ C for 10 min. The PCR product was subsequently purified, digested, and ligated with pre-digested pMAL-c4X vector (for MBP-Hp β 3GlcNAcT-His₆) or pMAL-c2X vector (for MBP-NmLgtA-His₆). The ligation products were transformed into electrocompetent *E. coli* DH5 α cells. The positive clones were confirmed by restriction digestion analysis and DNA sequencing (performed by UC-Davis Sequencing Facility). The correct plasmids with the target genes were subsequently transformed into expression host, *E. coli* BL21 (DE3) chemically competent cells.

4.3. Expression and purification

E. coli BL21 (DE3) strains harboring the recombinant plasmids with target genes were cultured in 50 mL Luria-Bertani (LB) rich medium (10 g L⁻¹ tryptone, 5 g L⁻¹ yeast extract, and 10

g L⁻¹ NaCl) containing 0.1 mg mL⁻¹ ampicillin with rapid shaking (220 rpm) at 37 °C overnight. Then 15 mL of the overnight cell culture was transferred into 1 L of LB rich medium with 0.1 mg mL⁻¹ ampicillin and incubated at 37 °C. When the OD_{600 nm} of the cell culture reached 0.8–1.0, isopropyl-1-thio-β-D-galactopyranoside (IPTG) was added to a final concentration of 0.1 mM to induce the over-expression of Hpβ3GlcNAcT, which was followed by incubation at 20 °C with slow shaking (190 rpm) for 20 h. Over-expression of NmLgtA was achieved by the same method.

All the cells were collected by centrifugation at 4 °C, 4000 rpm for 2 h. Harvested cells were resuspended with lysis buffer (100 mM, pH 8.0, Tris-HCl buffer containing 0.1 % Triton X-100). Lysis buffer (20 mL) was used to resuspend collected cells from one liter cell culture. The cells were broken by sonication and the method was as followed: amplitude at 68% for big tip, 3 s pulse on and 2 s pulse off for 60 cycles. Then the cell lysate was centrifuged at 4 °C, 12,000 rpm for 15 min, the supernatant was collected and loaded onto a Ni²⁺-NTA affinity column that was pre-equilibrated with 6 column volumes of binding buffer (50 mM, pH 7.5, Tris-HCl buffer, 10 mM imidazole, 0.5 M NaCl). The column was washed with 10 column volumes of binding buffer and 10 column volumes of washing buffer (Tris-HCl buffer, 50 mM, pH 7.5, 40 mM imidazole for NmLgtA and 60 mM imidazole for Hpβ3GlcNAcT, 0.5 M NaCl) sequentially to wash away the nonspecific binding protein. The target protein was eluted using Tris-HCl buffer (50 mM, pH 7.5) containing 200 mM of imidazole and 0.5 M of NaCl. Fractions containing the purified protein were combined and dialyzed against Tris-HCl buffer (20 mM, pH 7.5). The enzyme solution was stored at 4 °C with 10% glycerol. All procedures were performed at 4 °C.

4.4. pH profile by capillary electrophoresis (CE)

Standard enzymatic assays were carried out in reactions (10 μL) each containing a buffer (100 mM) with a pH in the range of 5.0–11.0, UDP-GlcNAc (1 mM), LacβMU (1 mM), MnCl₂ (10 mM), and Hpβ3GlcNAcT (3 μg) or NmLgtA (2 μg). Buffers used here were: MES, pH 5.0–6.5; Tris-HCl, pH 7.0–9.0; and CAPS, pH 10.0–11.0. Reactions were incubated at 30 °C for 10 min and quenched by heating at 98 °C for 2 min. Samples were centrifuged at 13,000 rpm for 5 min, the supernatants were analyzed by a P/ACETM Capillary Electrophoresis (CE) system equipped with a Photodiode Array (PDA) detector (Beckman Coulter, Inc., Fullerton, CA, USA). CE conditions were as follows: 75 μm i.d. capillary, 25 KV/80 μA, 6 s vacuum injections, monitored at 315 nm, the running buffer used was sodium tetraborate (25 mM, pH 9.4).

4.5. Effect of metal ions and EDTA

In order to investigate the effect of metal ions of Hpβ3GlcNAcT and NmLgtA, the reactions were carried out with the addition of different concentrations (1, 5, 10, and 20 mM) of different metal ions (MgCl₂, MnCl₂, CaCl₂) or EDTA (10 mM) in Tris-HCl buffer (100 mM, pH 8.0) containing UDP-GlcNAc (1 mM), LacβMU (1 mM), and Hpβ3GlcNAcT (3 μg) or NmLgtA (2 μg). Reactions were performed at 30 °C for 10 min. The reaction without metal ion nor EDTA was used as a control.

4.6. Donor substrate specificity of Hpβ3GlcNAcT and NmLgtA

All reactions were performed in duplicates. The enzymatic assay was carried out in 40 μL reaction system for both β3GlcNAcTs. For Hpβ3GlcNAcT, the reactions were carried out in Tris-HCl (100 mM, pH 8.0) containing LacβMU (2 mM), UDP-GlcNAc or its derivative or other sugar nucleotide (2 mM), a metal ion (10 mM OF MgCl₂ was used for Hpβ3GlcNAcT and

10 mM of MnCl₂ was used for NmLgtA), and Hpβ3GlcNAcT or NmLgtA (10 μg). The reaction mixtures were incubated at 30 °C for 24 h. The supernatants obtained after centrifugation were analyzed by CE as described for the pH profile. For reactions using UDP-Gal as the donor, the supernatants of the reaction mixtures after centrifugation were analyzed by CE using sodium tetraborate (50 mM, pH 11.0) as the running buffer to allow the separation of LacβMU and the trisaccharide product.

For HPLC analysis, aliquots of 1.0 μL from above supernatants were withdrawn and added to ice-cold 18% acetonitrile (99 μL) to make 100-fold dilution. The samples were kept on ice until aliquots of 10 μL were injected and analyzed by a Shimadzu LC-2010A system equipped with a membrane on-line degasser, a temperature control unit, and a fluorescence detector. A reverse-phase Premier C18 column (250 × 4.6 mm i.d., 5 μm particle size, Shimadzu) protected with a C18 guard column cartridge was used. The mobile phase was 18% (v/v) acetonitrile. The fluorescent 4-methylumbelliferyl compounds were detected with excitation at 325 nm and emission at 372 nm.

4.7. Acceptor substrate specificity of Hpβ4GalT, NmLgtB and Bβ4GalT

These assays were performed in duplicate in two steps using a sequential two-step process. An aliquot of 10 μL of the supernatant of the reaction mixture for the donor substrate specificity of β3GlcNAcT was drawn and added to a solution to give a final mixture of 20 μL containing Tris-HCl buffer (pH 7.5, 200 mM), MgCl₂ (10 mM), UDP-Gal (1 mM), and Hpβ4GalT (~2.0 μg), NmLgtB (~2.0 μg), or Bβ4GalT (~1.0 μg). The reaction mixtures were allowed to proceed for 24 h at 37 °C.

4.8. Kinetic study of Hpβ3GlcNAcT and NmLgtA by capillary electrophoresis (CE)

Kinetic assays were carried out in duplicates in a 10 μL system containing Tris-HCl buffer (100 mM, pH 8.0), MgCl₂ or MnCl₂ (10 mM), LacβMU (1 mM), UDP-GlcNAc or its derivatives at different concentration ranging from 0.1–10 mM (0.1 mM, 0.2 mM, 0.4 mM, 1 mM, 2 mM, 4 mM, 5 mM, 10 mM), and Hpβ3GlcNAcT (0.7 μg μL⁻¹) or NmLgtA (1.0 μg μL⁻¹). Reactions were incubated at 30 °C for 10 min, followed by heating at 98 °C for 2 min. The reaction mixtures were subsequently centrifuged at 13,000 rpm for 5 min to remove the protein precipitate. The supernatants were analyzed by CE as described for the pH profile.

4.9. One-pot four-enzyme preparative-scale synthesis of trisaccharide GalNAcβ1-3Galβ1-4GlcProN₃

To prepare the trisaccharide, a reaction mixture in Tris-HCl buffer (100 mM, pH 7.5) in a total volume of 5 mL containing LacβProN₃ (40 mg, 0.094 mmol), *N*-acetylgalactosamine (GalNAc, 27 mg, 0.122 mmol), adenosine 5'-triphosphate (ATP) (67 mg, 0.122 mmol), uridine 5'-triphosphate (UTP, 65 g, 0.122 mmol), MgCl₂ (20 mM). NahK (1.5 mg), PmGlmU (1.5 mg), NmLgtA (2.0 mg), and PmPpA (1.5 mg) were incubated in a shaker at 37 °C for 24 hrs. The product formation was monitored by TLC (EtOAc:MeOH:H₂O:HOAc = 4:2:1:0.2, by volume and detected by *p*-anisaldehyde sugar stain) and mass spectrometry. To stop the reaction, the reaction mixture was added with same volume (10 mL) of ethanol and incubated at 4 °C for 30 min. After centrifugation, the supernatant was concentrated and passed through a Bio Gel P-2 gel filtration column (water was used as an eluent). The fractions containing the product were collected, concentrated, further purified by silica gel column (EtOAc:MeOH:H₂O = 5:2:1, by volume). Further purification by preparative HPLC using C18 column provided the trisaccharide

GalNAc β 1–3Gal β 1–4Glc β ProN₃ as a white powder (58 mg, 98%). ¹H NMR (800 MHz, D₂O) δ 4.59 (d, *J* = 8.8 Hz, 1H), 4.46 (d, *J* = 8.0 Hz, 1H), 4.40 (d, *J* = 8.0 Hz, 1H), 4.13 (d, *J* = 3.6 Hz, 1H), 3.99–3.27 (m, 21H), 2.01 (s, 3H), 1.88 (m, 2H). ¹³C NMR (200 MHz, D₂O) δ 175.03, 103.20, 102.81, 101.97, 81.63, 78.20, 74.84, 74.74, 74.63, 74.22, 72.65, 70.59, 69.93 (2C), 68.34, 67.61, 67.23, 60.85, 60.82, 59.92, 47.73, 28.10, 22.09. HRMS (ESI) *m/z* calculated for C₂₃H₄₀N₄NaO₁₆ (M+Na) 651.2337, found 651.2335.

4.10. One-pot four-enzyme preparative-scale synthesis of trisaccharide Gal β 1–3Gal β 1–4Glc

To prepare the trisaccharide, a reaction mixture in Tris-HCl buffer (100 mM, pH 8.0) in a total volume of 10 mL containing lactose (Gal β 1–4Glc, 102 mg, 0.3 mmol), galactose (Gal, 65 mg, 0.36 mmol), adenosine 5'-triphosphate (ATP) (198 mg, 0.36 mmol), uridine 5'-triphosphate (UTP, 190 g, 0.36 mmol), MgCl₂ (10 mM). SpGalK (2.0 mg), BLUSP (2.0 mg), NmLgtA (2.0 mg), and PmPpA (1.5 mg) were incubated in a shaker at 37 °C for 24 hrs. The product formation was monitored by LC-MS. To stop the reaction, the reaction mixture was added with same volume (10 mL) of ethanol and incubated at 4 °C for 30 min. After centrifugation, the supernatant was concentrated and passed through a Bio Gel P-2 gel filtration column (water was used as an eluent). The fractions containing the product were collected, concentrated, further purified by C18 column (H₂O, 0–7 min; 0–50% ACN in H₂O, 7–35 min; 50% ACN in H₂O, 35–40min). The pure trisaccharide Gal β 1–3Gal β 1–4Glc was obtained by H₂O wash. The synthesis yield of this trisaccharide Gal β 1–3Gal β 1–4Glc was 10% (15 mg). ¹H NMR (800 MHz, D₂O) δ 5.20 (d, *J* = 3.2 Hz, 0.4H), 4.65 (d, *J* = 8.0 Hz, 0.6H), 4.46 (d, *J* = 7.2 Hz, 1H), 4.59 (d, *J* = 8.0 Hz, 1H), 4.49 (d, *J* = 8.0 Hz, 1H), 4.17 (d, *J* = 3.2 Hz, 1H), 3.94–3.26 (m, 17H); ¹³C NMR (200 MHz, D₂O) α -isomer: δ 104.21, 102.39, 91.70, 81.75, 78.08, 74.85, 74.65, 72.36, 71.27, 71.03, 70.07, 69.95, 68.43, 68.30, 68.28, 60.87, 59.94, 59.80. HRMS (ESI) *m/z* calculated for C₁₈H₃₂NaO₁₆ (M+Na) 527.1588, found 527.1584.

Acknowledgements

This work was supported by National Science Foundation grant CHE-1300449 as well as National Institutes of Health grants R01HD065122 and R01GM094523.

References and notes

- Peng, W.; Pranskevich, J.; Nycholat, C.; Gilbert, M.; Wakarchuk, W.; Paulson, J. C.; Razi, N. *Glycobiology* **2012**, *22*, 1453.
- Di Virgilio, S.; Glushka, J.; Moremen, K.; Pierce, M. *Glycobiology* **1999**, *9*, 353.
- Naruchi, K.; Hamamoto, T.; Kuroguchi, M.; Hinou, H.; Shimizu, H.; Matsushita, T.; Fujitani, N.; Kondo, H.; Nishimura, S. *J. Org. Chem.* **2006**, *71*, 9609.
- Nycholat, C. M.; Peng, W.; McBride, R.; Antonopoulos, A.; de Vries, R. P.; Polonskaya, Z.; Finn, M. G.; Dell, A.; Haslam, S. M.; Paulson, J. C. *J. Am. Chem. Soc.* **2013**, *135*, 18280.
- Renkonen, O. *Cell. Mol. Life Sci.* **2000**, *57*, 1423.
- Vasiliu, D.; Razi, N.; Zhang, Y.; Jacobsen, N.; Allin, K.; Liu, X.; Hoffmann, J.; Bohorov, O.; Blixt, O. *Carbohydr. Res.* **2006**, *341*, 1447.
- Sauerzapfe, B.; Krennek, K.; Schmiedel, J.; Wakarchuk, W. W.; Palantova, H.; Kren, V.; Elling, L. *Glycoconj. J.* **2009**, *26*, 141.
- Blixt, O.; Razi, N. *Methods Enzymol.* **2006**, *415*, 137.
- Nycholat, C. M.; McBride, R.; Ekiert, D. C.; Xu, R.; Rangarajan, J.; Peng, W.; Razi, N.; Gilbert, M.; Wakarchuk, W.; Wilson, I. A.; Paulson, J. C. *Angew. Chem. Int. Ed.* **2012**, *51*, 4860.
- Chien, W. T.; Liang, C. F.; Yu, C. C.; Lin, C. H.; Li, S. P.; Primadona, I.; Chen, Y. J.; Mong, K. K.; Lin, C. C. *Chem. Commun.* **2014**, *50*, 5786.
- Johnson, K. F. *Glycoconj. J.* **1999**, *16*, 141.
- Yu, H.; Lau, K.; Thon, V.; Autran, C. A.; Jantscher-Krenn, E.; Xue, M.; Li, Y.; Sugiarto, G.; Qu, J.; Mu, S.; Ding, L.; Bode, L.; Chen, X. *Angew. Chem. Int. Ed.* **2014**, *53*, 6687.
- Priem, B.; Gilbert, M.; Wakarchuk, W. W.; Heyraud, A.; Samain, E. *Glycobiology* **2002**, *12*, 235.
- Gebus, C.; Cottin, C.; Randrianisoa, M.; Drouillard, S.; Samain, E. *Carbohydr. Res.* **2012**, *361*, 83.
- Blixt, O.; van Die, I.; Norberg, T.; van den Eijnden, D. H. *Glycobiology* **1999**, *9*, 1061.
- Togayachi, A.; Kozono, Y.; Ishida, H.; Abe, S.; Suzuki, N.; Tsunoda, Y.; Hagiwara, K.; Kuno, A.; Ohkura, T.; Sato, N.; Sato, T.; Hirabayashi, J.; Ikehara, Y.; Tachibana, K.; Narimatsu, H. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 15829.
- Wakarchuk, W.; Martin, A.; Jennings, M. P.; Moxon, E. R.; Richards, J. C. *J. Biol. Chem.* **1996**, *271*, 19166.
- Logan, S. M.; Altman, E.; Mykytczuk, O.; Brisson, J. R.; Chandan, V.; Schur, M. J.; St Michael, F.; Masson, A.; Leclerc, S.; Hiratsuka, K.; Smirnova, N.; Li, J.; Wu, Y.; Wakarchuk, W. W. *Glycobiology* **2005**, *15*, 721.
- Guan, W.; Ban, L.; Cai, L.; Li, L.; Chen, W.; Liu, X.; Mrksich, M.; Wang, P. G. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 5025.
- Palcic, M. M. *Curr. Opin. Chem. Biol.* **2011**, *15*, 226.
- Lau, K.; Thon, V.; Yu, H.; Ding, L.; Chen, Y.; Muthana, M. M.; Wong, D.; Huang, R.; Chen, X. *Chem. Commun.* **2010**, *46*, 6066.
- Wakarchuk, W. W.; Cunningham, A.; Watson, D. C.; Young, N. M. *Protein Eng.* **1998**, *11*, 295.
- Park, J. E.; Lee, K. Y.; Do, S. I.; Lee, S. S. *J. Biochem. Mol. Biol.* **2002**, *35*, 330.
- Logan, S. M.; Conlan, J. W.; Monteiro, M. A.; Wakarchuk, W. W.; Altman, E. *Mol. Microbiol.* **2000**, *35*, 1156.
- Endo, T.; Koizumi, S.; Tabata, K.; Ozaki, A. *Glycobiology* **2000**, *10*, 809.
- Namdjou, D. J.; Chen, H. M.; Vinogradov, E.; Brochu, D.; Withers, S. G.; Wakarchuk, W. W. *ChemBiochem* **2008**, *9*, 1632.
- Dumon, C.; Bosso, C.; Utille, J. P.; Heyraud, A.; Samain, E. *ChemBiochem* **2006**, *7*, 359.
- Wong, C. H.; Wang, R.; Ichikawa, Y. *J. Org. Chem.* **1992**, *57*, 4343.
- Palcic, M. M.; Srivastava, O. P.; Hindsgaul, O. *Carbohydr. Res.* **1987**, *159*, 315.
- Berliner, L. J.; Davis, M. E.; Ebner, K. E.; Beyer, T. A.; Bell, J. E. *Mol. Cell. Biochem.* **1984**, *62*, 37.
- Blixt, O.; Brown, J.; Schur, M. J.; Wakarchuk, W.; Paulson, J. C. *J. Org. Chem.* **2001**, *66*, 2442.
- Brockhausen, I.; Benn, M.; Bhat, S.; Marone, S.; Riley, J. G.; Montoya-Peleaz, P.; Vlahakis, J. Z.; Paulsen, H.; Schutzbach, J. S.; Szarek, W. A. *Glycoconj. J.* **2006**, *23*, 525.
- Boeggeman, E. E.; Balaji, P. V.; Sethi, N.; Masibay, A. S.; Qasba, P. K. *Protein Eng.* **1993**, *6*, 779.
- Ramakrishnan, B.; Shah, P. S.; Qasba, P. K. *J. Biol. Chem.* **2001**, *276*, 37665.
- Hanada, E.; Handa, S.; Konno, K.; Yamakawa, T. *J. Biochem.* **1978**, *83*, 85.
- Higuchi, R.; Inukai, K.; Zhou, J. X.; Honda, M.; Komori, T.; Tsuji, S.; Nagai, Y. *Liebigs Ann. Chem.* **1993**, 359.
- Li, Y.; Yu, H.; Chen, Y.; Lau, K.; Cai, L.; Cao, H.; Tiwari, V. K.; Qu, J.; Thon, V.; Wang, P. G.; Chen, X. *Molecules* **2011**, *16*, 6396.
- Chen, Y.; Thon, V.; Li, Y.; Yu, H.; Ding, L.; Lau, K.; Qu, J.; Hie, L.; Chen, X. *Chem. Commun.* **2011**, *47*, 10815.
- Chen, M.; Chen, L. L.; Zou, Y.; Xue, M.; Liang, M.; Jin, L.; Guan, W. Y.; Shen, J.; Wang, W.; Wang, L.; Liu, J.; Wang, P. G. *Carbohydr. Res.* **2011**, *346*, 2421.
- Muthana, M. M.; Qu, J.; Li, Y.; Zhang, L.; Yu, H.; Ding, L.; Malekan, H.; Chen, X. *Chem. Commun.* **2012**, *48*, 2728.
- Donald, A. S.; Feeney, J. *Carbohydr. Res.* **1988**, *178*, 79.
- Urashima, T.; Sun, Y.; Fukuda, K.; Hirayama, K.; Taufik, E.; Nakamura, T.; Saito, T.; Merchant, J.; Green, B.; Messer, M. *Glycoconj. J.* **2015**, *32*, 361.
- Saito, T.; Itoh, T.; Adachi, S. *Carbohydr. Res.* **1987**, *165*, 43.
- Urashima, T.; Saito, T.; Nishimura, J.; Ariga, H. *Bioch. Biophys. Acta* **1989**, *992*, 375.
- Urashima, T.; Kawai, Y.; Nakamura, T.; Arai, I.; Saito, T.; Namiki, M.; Yamaoka, K.; Kawahara, K.; Messer, M. *Comp. Biochem. Physiol. C Pharmacol. Toxicol. Endocrinol.* **1999**, *124*, 295.
- Fukuda, K.; Yamamoto, A.; Ganzorig, K.; Khuukhenbaatar, J.; Senda, A.; Saito, T.; Urashima, T. *J. Dairy Sci.* **2010**, *93*, 5572.
- Difilippo, E.; Willems, H. A.; Vendrig, J. C.; Fink-Gremmels, J.; Gruppen, H.; Schols, H. A. *J. Agric. Food Chem.* **2015**, *63*, 4805.
- Kunz, C.; Rudloff, S.; Schad, W.; Braun, D. *Br. J. Nutr.* **1999**, *82*, 391.
- Urashima, T.; Fujita, S.; Fukuda, K.; Nakamura, T.; Saito, T.; Cowan, P.; Messer, M. *Glycoconj. J.* **2014**, *31*, 387.
- Messer, M.; Trifonoff, E.; Stern, W.; Collins, J. G.; Bradbury, J. H. *Carbohydr. Res.* **1980**, *83*, 327.

51. Urashima, T.; Taufik, E.; Fukuda, R.; Nakamura, T.; Fukuda, K.; Saito, T.; Messer, M. *Glycoconj. J.* **2013**, *30*, 801.
52. Zhang, L.; Muthana, M. M.; Yu, H.; McArthur, J. B.; Qu, J.; Chen, X. *Carbohydr. Res.* **2016**, *419*, 18.
53. Muthana, M. M.; Qu, J.; Xue, M.; Klyuchnik, T.; Siu, A.; Li, Y.; Zhang, L.; Yu, H.; Li, L.; Wang, P. G.; Chen, X. *Chem. Commun.* **2015**, *51*, 4595.
54. Yu, H.; Yu, H.; Karpel, R.; Chen, X. *Bioorg. Med. Chem.* **2004**, *12*, 6427.
55. Zhang, L.; Lau, K.; Cheng, J.; Yu, H.; Li, Y.; Sugiarto, G.; Huang, S.; Ding, L.; Thon, V.; Wang, P. G.; Chen, X. *Glycobiology* **2010**, *20*, 1077.