



Published in final edited form as:

Bioorg Med Chem. 2016 April 15; 24(8): 1696–1705. doi:10.1016/j.bmc.2016.02.043.

## Donor substrate promiscuity of bacterial $\beta$ 1–3-*N*-acetylglucosaminyltransferases and acceptor substrate flexibility of $\beta$ 1–4-galactosyltransferases

Yanhong Li<sup>a,†</sup>, Mengyang Xue<sup>a,b,c,†</sup>, Xue Sheng<sup>a,d</sup>, Hai Yu<sup>a</sup>, Jie Zeng<sup>a,e</sup>, Vireak Thon<sup>a,f,g</sup>, Yi Chena<sup>a</sup>, Musleh M. Muthanaa<sup>a,h</sup>, Peng G. Wang<sup>b,f</sup>, and Xi Chena<sup>a,\*</sup>

a

b

e

f

### Abstract

$\beta$ 1–3-*N*-Acetylglucosaminyltransferases ( $\beta$ 3GlcNAcTs) and  $\beta$ 1–4-galactosyltransferases ( $\beta$ 4GalTs) have been broadly used in enzymatic synthesis of *N*-acetylglucosamine (LacNAc)-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  $\beta$ 3GlcNAcTs and  $\beta$ 4GalTs, donor substrate specificity studies of two bacterial  $\beta$ 3GlcNAcTs from *Helicobacter pylori* (Hp $\beta$ 3GlcNAcT) and *Neisseria meningitidis* (NmLgtA), respectively, using a library of 39 sugar nucleotides were carried out. The two  $\beta$ 3GlcNAcTs have complementary donor substrate promiscuity and 13 different trisaccharides were produced. They were used to investigate the acceptor substrate specificities of three  $\beta$ 4GalTs from *Neisseria meningitidis* (NmLgtB), *Helicobacter pylori* (Hp $\beta$ 4GalT), and bovine (B $\beta$ 4GalT), respectively. Ten of the 13 trisaccharides were shown to be tolerable acceptors for at least one of these  $\beta$ 4GalTs. The application of NmLgtA in one-pot multienzyme (OPME) synthesis of two trisaccharides including GalNAc $\beta$ 1–3Gal $\beta$ 1–4Glc $\beta$ ProN<sub>3</sub> and Gal $\beta$ 1–3Gal $\beta$ 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.

\*Corresponding author. Tel: 530-754-6037; fax: 530-752-8995. xiichen@ucdavis.edu.

<sup>c</sup>Current address: Shandong Province Key Laboratory of Emergency and Critical Care Medicine, Shandong University, Jinan, Shandong 250100, China

<sup>d</sup>Current address: Department of Chemistry and Biochemistry, University of California, San Diego, CA 92093, USA

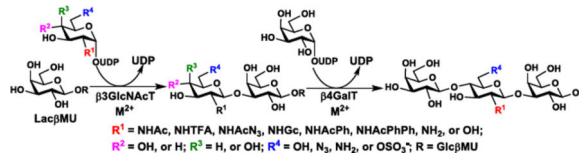
<sup>g</sup>Current address: Laboratory of Bacterial Polysaccharides, Food and Drug Administration, Bethesda, MD 20892, USA

<sup>h</sup>Current address: Children's National Medical Center, Washington DC, 20010, USA

<sup>†</sup>These authors contributed equally.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

## Graphical abstract



## Keywords

Carbohydrate; Glycosyltransferase; Donor substrate promiscuity; Acceptor substrate specificity; Enzymatic synthesis; One-pot multienzyme

## 1. Introduction

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

For  $\beta$ 3GlcNAcTs, mammalian enzymes were found to be quite specific on the acceptor substrates.<sup>1, 15, 16</sup> In comparison, recombinant  $\beta$ 3GlcNAcT from *Neisseria meningitidis* (NmLgtA) was shown to be more promiscuous and was able to use both  $\alpha$ - and  $\beta$ -galactosides as acceptors.<sup>15, 17</sup> It was found to have low<sup>15</sup> or no activity towards type 1 (Gal $\beta$ 1–3GlcNAc) or type 3 (Gal $\beta$ 1–3GalNAc $\alpha$ ) galactoside acceptors.<sup>1, 6</sup> Branched glycans or longer LacNAc repeats were less efficient acceptors for NmLgtA. More recently, a  $\beta$ 3GlcNAcT from *Helicobacter pylori* (Hp $\beta$ 3GlcNAcT) was cloned and characterized.<sup>1, 18</sup> Hp $\beta$ 3GlcNAcT was shown to be specific for the  $\beta$ -linked galactoside acceptors<sup>14</sup> and prefer type 2 (Gal $\beta$ 1–4GlcNAc) acceptors but can also use type 1 (Gal $\beta$ 1–3GlcNAc), type 3 (Gal $\beta$ 1–3GalNAc $\alpha$ ), and type 4 (Gal $\beta$ 1–3GalNAc $\beta$ ) acceptors on both linear and branched glycan structures.<sup>1</sup> Regarding donor substrate specificity, only a limited number of UDP-sugars have been tested for NmLgtA which was shown to tolerate uridine 5'-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.<sup>15, 19</sup> The donor substrate specificity of Hp $\beta$ 3GlcNAcT has not been explored in detail.

Bovine  $\beta$ 4GalT 1 (B $\beta$ 4GalT) is the most well studied mammalian glycosyltransferase that has been used for biosynthesis.<sup>20, 21</sup> In addition,  $\beta$ 4GalTs have been cloned from *Neisseria meningitidis*, *Neisseria gonorrhoeae*, and *Helicobacter pylori*.<sup>21–25</sup> *Neisseria meningitidis*  $\beta$ 4GalT (NmLgtB) and *Helicobacter pylori*  $\beta$ 4GalT (Hp $\beta$ 4GalT), and B $\beta$ 4GalT have been used broadly for enzymatic and chemoenzymatic synthesis of  $\beta$ 1–4-linked galactosides.<sup>13, 21, 26–30</sup> 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'-disphosphate glucose (UDP-Glc) as the starting material which can be converted to uridine 5'-disphosphate galactose (UDP-Gal) by StGalE for NmLgtB-catalyzed synthesis of  $\beta$ 1–4-linked galactosides.<sup>31</sup> Although the tolerance of the modifications on the acceptor substrates of  $\beta$ 4GalTs has been tested using monosaccharide derivatives,<sup>21, 32</sup> 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  $\beta$ 3GlcNAcTs and  $\beta$ 4GalTs, herein we report the donor substrate specificities of NmLgtA and Hp $\beta$ 3GlcNAcT using 4-methylumbelliferyl (MU)  $\beta$ -lactoside Gal $\beta$ 1–4Glc $\beta$ MU (or Lac $\beta$ MU) 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 Hp $\beta$ 3GlcNAcT-catalyzed reactions and have been used to investigate the acceptor substrate specificities of NmLgtB, Hp $\beta$ 4GalT, and B $\beta$ 4GalT (Scheme 1). We have also demonstrated the application of NmLgtA in one-pot multienzyme (OPME) preparative-scale synthesis of two trisaccharides including GalNAc $\beta$ 1–3Gal $\beta$ 1–4Glc $\beta$ ProN<sub>3</sub> and Gal $\beta$ 1–3Gal $\beta$ 1–4Glc 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.

## 2. Results and discussion

### 2.1. Cloning and expression of $\beta$ 3GlcNAcTs and $\beta$ 4GalTs

Both NmLgtA<sup>15</sup> and Hp $\beta$ 3GlcNAcT<sup>1</sup> 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 Hp $\beta$ 3GlcNAcT were cloned into pMAL-c2X and pMAL-c4X vectors, respectively, both as N-terminal maltose-binding protein (MBP)-fused and C-terminal hexa-histidine (His<sub>6</sub>)-tagged recombinant proteins. Soluble and active enzymes could be obtained by inducing *Escherichia coli* BL21(DE3) cells with 0.1 mM of isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) followed by incubation at 20 °C for 20 hours. Purification was achieved by one-step nickel-nitrilotriacetic acid (Ni<sup>2+</sup>-NTA) affinity chromatography. About 10 mg of NmLgtA and only about 1 mg of Hp $\beta$ 3GlcNAcT 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 Hp $\beta$ 3GlcNAcT (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 B $\beta$ 4GalT was usually expressed in *Escherichia coli* as inclusion bodies and active soluble enzyme was obtained after proper refolding procedures.<sup>33, 34</sup> We found that when cloned in vector plasmid pET15b as an N-His<sub>6</sub>-tagged fusion protein with an N-terminal 128 amino acid truncation, soluble and active

B $\beta$ 4GalT was readily obtained without refolding although its expression level ( $< 1 \text{ mg L}^{-1}$  culture) was low (Fig. S1). In comparison, NmLgtB and Hp $\beta$ 4GalT (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.<sup>21</sup>

## 2.2. pH Profile and metal effect of $\beta$ 3GlcNAcTs

When 4-methylumbelliferyl  $\beta$ -lactoside (Lac $\beta$ MU) was used as the acceptor, the observed optimal pH for both NmLgtA and Hp $\beta$ 3GlcNAcT was 8.0 (Fig. 2). The slight different optimal pH observed for NmLgtA compared to the previously reported pH 7.5<sup>15</sup> 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.

The effects of several divalent metal ions including Mn<sup>2+</sup>, Mg<sup>2+</sup>, and Ca<sup>2+</sup>, as well as ethylenediaminetetraacetic acid (EDTA) on the activity of both  $\beta$ 3GlcNAcTs were investigated. As shown in Fig. S2, both Hp $\beta$ 3GlcNAcT (Fig. S2A) NmLgtA (Fig. S2B) required a divalent metal ion for activity as adding 10 mM of EDTA completely abolished their activities. Hp $\beta$ 3GlcNAcT exhibited optimal activity in the presence of Mg<sup>2+</sup>. Increasing the concentration of Mg<sup>2+</sup> from 1 mM to 10 mM did not affect the activity of Hp $\beta$ 3GlcNAcT significantly, while its activity dropped slightly when the concentration of Mg<sup>2+</sup> increased from 10 mM to 20 mM. Mn<sup>2+</sup> was shown to be a less efficient cofactor for Hp $\beta$ 3GlcNAcT and increasing the concentration of Mn<sup>2+</sup> from 1 mM to 20 mM caused gradual decrease of Hp $\beta$ 3GlcNAcT activity. Ca<sup>2+</sup> was shown to be a less efficient cofactor for Hp $\beta$ 3GlcNAcT compared to Mg<sup>2+</sup> and its concentration ranging from 1–20 mM did not affect the Hp $\beta$ 3GlcNAcT activity significantly. Unlike Hp $\beta$ 3GlcNAcT, NmLgtA exhibited the best activity in the presence of Mn<sup>2+</sup> with an optimal activity observed in the presence of 10 mM of Mn<sup>2+</sup>. Increasing or decreasing the concentrations of Mn<sup>2+</sup> led to the decrease of NmLgtA activity. Although a less effective cofactor compared to Mn<sup>2+</sup> at concentrations at or lower than 10 mM, increasing the concentration of Mg<sup>2+</sup> from 1 mM to 20 mM led to a gradual increase of NmLgtA activity. Ca<sup>2+</sup> was not a suitable cofactor for NmLgtA. The results of the metal effect of NmLgtA were consistent to those reported previously.<sup>15</sup>

## 2.3. Donor Substrate Specificities of $\beta$ 3GlcNAcTs

Once the optimal conditions were identified for Hp $\beta$ 3GlcNAcT 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 Lac $\beta$ MU 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).

As shown in Table 1, both  $\beta$ 3GlcNAcTs exhibited promiscuous and somewhat complementary substrate specificities. NmLgtA was able to use 11 UDP-sugars and Hp $\beta$ 3GlcNAcT 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  $\beta$ 3GlcNAcTs could use 2'-modified UDP-GlcNAc derivatives including UDP-GlcNTFA (**2**), UDP-GlcNAcN<sub>3</sub> (**3**), and UDP-GlcNGc (**4**) as well as a 6'-modified UDP-GlcNAc derivative UDP-GlcNAc6NH<sub>2</sub> (**10**) as donor substrates. UDP-GlcNTFA (**2**) containing an *N*-trifluoroacetyl group was an excellent donor substrate for both  $\beta$ 3GlcNAcTs. In contrary to a previous report indicating UDP-GlcNAcN<sub>3</sub> (**3**) containing an *N*-azidoacetyl group as an intolerable donor for NmLgtA,<sup>19</sup> the compound was confirmed by both HRMS and CE assays to be an acceptable donor substrate for both  $\beta$ 3GlcNAcTs although the yields were not high.

Quite interestingly, two  $\beta$ 3GlcNAcTs 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-GlcNH<sub>2</sub> (**8**) with an amino group. However, these UDP-sugars were not acceptable donor substrates for Hp $\beta$ 3GlcNAcT. On the other hand, 6'-modifications of UDP-GlcNAc were more tolerable by Hp $\beta$ 3GlcNAcT compared to NmLgtA. For example, UDP-GlcNAc6NH<sub>2</sub> (**10**) containing a 6'-amino group was a better donor substrate for Hp $\beta$ 3GlcNAcT and the reaction was achieved with a higher yield than that with NmLgtA. In addition, UDP-GlcNAc6N<sub>3</sub> (**11**) containing a 6'-azido group and UDP-GlcNAc6S (**16**) containing a 6'-*O*-sulfate group were suitable donor substrates for Hp $\beta$ 3GlcNAcT, 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 Hp $\beta$ 3GlcNAcT. In consistent with previous reports,<sup>15, 19</sup> UDP-GalNAc (**23**) was an excellent donor substrate for NmLgtA. The compound, however, was not a donor for Hp $\beta$ 3GlcNAcT. To our surprise, unlike a previous report indicating both UDP-Glc (**21**) and UDP-Gal (**25**) were not donor substrate for NmLgtA,<sup>15</sup> our results showed that both of these UDP-sugars were suitable donor substrates for NmLgtA but not for Hp $\beta$ 3GlcNAcT.

Neither  $\beta$ 3GlcNAcTs 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 $\beta$ 3GlcNAcTs

Due to their potential in synthetic applications and chemical biological studies, three UDP-sugars including UDP-GlcNAc (**1**), UDP-GlcNTFA (**2**), and UDP-GlcNAcN<sub>3</sub> (**3**) that could be used by both  $\beta$ 3GlcNAcTs 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 Hp $\beta$ 3GlcNAcT no matter which substrate among the three was used. Compared to UDP-GlcNTFA (**2**), the catalytic efficiencies of both  $\beta$ 3GlcNAcTs 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  $\beta$ 3GlcNAcTs were lower (3–4-fold) when UDP-GlcNAcN<sub>3</sub> (**3**) was used as the donor substrate mainly due to the higher  $K_M$  values observed for UDP-GlcNAcN<sub>3</sub> (**3**).

## 2.5. Acceptor substrate specificities of $\beta$ 4GalTs

Thirteen trisaccharides could be obtained by reactions catalyzed by NmLgtA and/or Hp $\beta$ 3GlcNAcT. They were used to test the acceptor substrate specificities of three  $\beta$ 4GalTs including NmLgtB, Hp $\beta$ 4GalT, and B $\beta$ 4GalT 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 Hp $\beta$ 4GalT and NmLgtB, 6.2% for B $\beta$ 4GalT and Hp $\beta$ 4GalT, and 2.9% for NmLgtB and B $\beta$ 4GalT. The overall sequence similarity among the three was less than 16%.

Despite their low sequence homology, NmLgtB, Hp $\beta$ 4GalT, and B $\beta$ 4GalT are effective  $\beta$ 4GalTs and can use several trisaccharides and derivatives produced by  $\beta$ 3GlcNAcT-catalyzed reactions as acceptors. As shown in Table 3, ten of the 13 trisaccharides could be used as acceptors by either NmLgtB or Hp $\beta$ 4GalT or both. B $\beta$ 4GalT could use only four as acceptors (HRMS data are shown in Fig. S5). Besides GlcNAc-containing trisaccharide GlcNAc $\beta$ 1–3Lac $\beta$ MU (**1A**), all three  $\beta$ 4GalTs could use 2'-modified GlcNAc-containing trisaccharides GlcNTFA $\beta$ 1–3Lac $\beta$ MU (**2A**), GlcNAcN<sub>3</sub> $\beta$ 1–3Lac $\beta$ MU (**3A**), and GlcNAcPh $\beta$ 1–3Lac $\beta$ MU (**5A**) as acceptors. While GlcNTFA-containing trisaccharide GlcNTFA $\beta$ 1–3Lac $\beta$ MU (**2A**) was well tolerated by B $\beta$ 4GalT, other 2'-modifications on GlcNAc decrease its activity more significantly. 6'-Modification completely abolished its activity. While not acceptable by B $\beta$ 4GalT, GlcNGc $\beta$ 1–3Lac $\beta$ MU (**4A**) was an excellent acceptor for both NmLgtB and Hp $\beta$ 4GalT.

In consistent with those observed previously using monosaccharide derivatives as acceptors,<sup>21</sup> NmLgtB was more promiscuous towards acceptors containing a terminal 2'-modified GlcNAc, while Hp $\beta$ 4GalT was more tolerable towards acceptors containing a terminal 6'-modified GlcNAc. For example, trisaccharide GlcNAcPhPh $\beta$ 1–3Gal $\beta$ 1–4Glc $\beta$ MU (**6A**) containing a terminal *N*-diphenylacetylglucosamine (GlcNAcPhPh) was a suitable acceptor for NmLgtB but not for Hp $\beta$ 4GalT. On the other hand, GlcNAc6NH<sub>2</sub> $\beta$ 1–3Lac $\beta$ MU (**10A**) and GlcNAc6S $\beta$ 1–3Lac $\beta$ MU (**16A**) could be used as acceptors by Hp $\beta$ 4GalT but not NmLgtB. In addition, a high yield (100%) was achieved when GlcNAc6N<sub>3</sub> $\beta$ 1–3Lac $\beta$ MU (**11A**) was used as an acceptor for Hp $\beta$ 4GalT compared to a low yield (5–10%) for NmLgtB. Quite excitingly, the Glc $\beta$ 1–3Lac $\beta$ MU (**21A**) obtained by NmLgtA-catalyzed reaction can be extended further by NmLgtB for the formation of

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

## 2.6. Preparative-scale synthesis of GalNAc $\beta$ 1–3Gal $\beta$ 1–4Glc $\beta$ ProN<sub>3</sub> and Gal $\beta$ 1–3Gal $\beta$ 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 $\beta$ 1–3Gal $\beta$ 1–4Glc $\beta$ ProN<sub>3</sub> (**1**) (Scheme 2A) and Gal $\beta$ 1–3Gal $\beta$ 1–4Glc (**2**) (Scheme 2B) were carried out.

GalNAc $\beta$ 1–3Gal $\beta$ 1–4Glc $\beta$ OR component has been found in glycosphingolipids identified in rat granuloma, rat abdominal macrophage,<sup>35</sup> starfish *Asterias amurensis versicolor* Sladen.<sup>36</sup> As shown in Scheme 2A, GalNAc $\beta$ 1–3Gal $\beta$ 1–4Glc $\beta$ ProN<sub>3</sub> (**1**) was synthesized from Gal $\beta$ 1–4Glc $\beta$ ProN<sub>3</sub> and simple free monosaccharide GalNAc using a one-pot four-enzyme (OP4E) system containing *Bifidobacterium longum* ATCC55813 *N*-acetylhexosamine 1-kinase (NahK),<sup>37</sup> *Pasteurella multocida* *N*-acetylglucosamine- 1-phosphate uridylyltransferase (PmGlmU),<sup>38</sup> *Pasteurella multocida* inorganic pyrophosphatase (PmPpA),<sup>21</sup> 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 $\beta$ 1–3Gal $\beta$ 1–4Glc $\beta$ ProN<sub>3</sub> (**1**) and the yield was quantitative. Compared to a previous report,<sup>15</sup> 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 <sup>13</sup>C nuclear magnetic resonance (NMR) chemical shifts of the disaccharide Lac $\beta$ ProN<sub>3</sub> 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 $\beta$ ProN<sub>3</sub>. 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  $\beta$ -linkage. These results confirmed the structure of the trisaccharide product as GalNAc $\beta$ 1–3Gal $\beta$ 1–4Glc $\beta$ ProN<sub>3</sub> (**1**).

Gal $\beta$ 1–3Gal $\beta$ 1–4Glc (3'-galactosyllactose) has been found in human milk<sup>41</sup> and eastern quoll (*Dasyurus viverrinus*) milk<sup>42</sup> as well as in the colostrums of bovine,<sup>43</sup> ovine,<sup>44</sup> primate,<sup>45</sup> camel,<sup>46</sup> and horse.<sup>47</sup> It is also an abundant oligosaccharide in the milk of elephant<sup>48</sup> and marsupials such as brushtail possum (*Trichosurus vulpecula*),<sup>49</sup> tammar wallaby (*Macropus eugenii*) and grey kangaroo,<sup>50</sup> and koala (*Phascolarctos cinereus*).<sup>51</sup> As shown in Scheme 2B, Gal $\beta$ 1–3Gal $\beta$ 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),<sup>39</sup> a *Bifidobacterium longum* ATCC55813 UDP-sugar pyrophosphorylase (BLUSP),<sup>40</sup> PmPpA,<sup>21</sup> and NmLgtA. SpGalK, BLUSP, and PmPpA

were used previously for high efficient (85% yield) one-pot multienzyme (OPME) synthesis of UDP-Gal.<sup>40</sup> As UDP-Gal was a relatively weak substrate for NmLgtA, the synthetic yield (10%) of Gal $\beta$ 1–3Gal $\beta$ 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 $\beta$ MU) similar to that observed previously.<sup>15</sup> 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 <sup>13</sup>C 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  $\beta$ -linkage. These results confirmed the structure of the trisaccharide product as Gal $\beta$ 1–3Gal $\beta$ 1–4Glc (**2**).

### 3. Conclusions

The donor substrate specificity of two  $\beta$ 1–3-*N*-acetylglucosaminyltransferases (Hp $\beta$ 3GlcNAcT and NmLgtA) towards 39 sugar nucleotides and the acceptor substrate specificity of three  $\beta$ 1–4-galactosyltransferases (Hp $\beta$ 4GalT, NmLgtB, and B $\beta$ 4GalT) towards 13 trisaccharide acceptors formed by Hp $\beta$ 3GlcNAcT and/or NmLgtA-catalyzed reactions were investigated for the synthesis of diverse LacNAc-containing glycans or derivatives. Compared to Hp $\beta$ 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 NH<sub>2</sub>. Besides UDP-GlcNAc and its derivatives, UDP-GalNAc, UDP-Glc, and UDP-Gal were also tolerable donor substrates for NmLgtA but not for Hp $\beta$ 3GlcNAcT. However, for 6'-modified UDP-GlcNAc derivatives, Hp $\beta$ 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 $\beta$ 3GlcNAcT,  $\beta$ 4GalT from *H. pylori* (Hp $\beta$ 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 $\beta$ 4GalT. In addition, preparative-scale enzymatic synthesis of trisaccharides GalNAc $\beta$ 1–3Gal $\beta$ 1–4Glc $\beta$ ProN<sub>3</sub> and Gal $\beta$ 1–3Gal $\beta$ 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  $\beta$ 3GlcNAcTs and  $\beta$ 4GalTs will allow efficient chemoenzymatic synthesis of diverse derivatives of LacNAc-containing structures as probes for protein binding study and for developing potential therapeutics.





and 10 g L<sup>-1</sup> NaCl) containing 0.1 mg mL<sup>-1</sup> 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<sup>-1</sup> ampicillin and incubated at 37 °C. When the OD<sub>600 nm</sub> 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 Ni<sup>2+</sup>-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), MnCl<sub>2</sub> (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 (Bechman 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 (MgCl<sub>2</sub>, MnCl<sub>2</sub>, CaCl<sub>2</sub>) 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 Hp $\beta$ 3GlcNAcT and NmLgtA

All reactions were performed in duplicates. The enzymatic assay was carried out in 40  $\mu$ L reaction system for both  $\beta$ 3GlcNAcTs. For Hp $\beta$ 3GlcNAcT, the reactions were carried out in Tris-HCl (100 mM, pH 8.0) containing Lac $\beta$ MU (2 mM), UDP-GlcNAc or its derivative or other sugar nucleotide (2 mM), a metal ion (10 mM OF MgCl<sub>2</sub> was used for Hp $\beta$ 3GlcNAcT and 10 mM of MnCl<sub>2</sub> was used for NmLgtA), and Hp $\beta$ 3GlcNAcT or NmLgtA (10  $\mu$ g). 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 Lac $\beta$ MU and the trisaccharide product.

For HPLC analysis, aliquots of 1.0  $\mu$ L from above supernatants were withdrawn and added to ice-cold 18% acetonitrile (99  $\mu$ L) to make 100-fold dilution. The samples were kept on ice until aliquots of 10  $\mu$ L were injected and analyzed by a Shimadzu LC-2010A system equipped with a membrane online degasser, a temperature control unit, and a fluorescence detector. A reverse-phase Premier C18 column (250  $\times$  4.6 mm i.d., 5  $\mu$ m 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 Hp $\beta$ 4GalT, NmLgtB and B $\beta$ 4GalT

These assays were performed in duplicate in two steps using a sequential two-step process. An aliquot of 10  $\mu$ L of the supernatant of the reaction mixture for the donor substrate specificity of  $\beta$ 3GlcNAcT was drawn and added to a solution to give a final mixture of 20  $\mu$ L containing Tris-HCl buffer (pH 7.5, 200 mM), MgCl<sub>2</sub> (10 mM), UDP-Gal (1 mM), and Hp $\beta$ 4GalT (~2.0  $\mu$ g), NmLgtB (~2.0  $\mu$ g), or B $\beta$ 4GalT (~1.0  $\mu$ g). The reaction mixtures were allowed to proceed for 24 h at 37 °C.

#### 4.8. Kinetic study of Hp $\beta$ 3GlcNAcT and NmLgtA by capillary electrophoresis (CE)

Kinetic assays were carried out in duplicates in a 10  $\mu$ L system containing Tris-HCl buffer (100 mM, pH 8.0), MgCl<sub>2</sub> or MnCl<sub>2</sub> (10 mM), Lac $\beta$ MU (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 Hp $\beta$ 3GlcNAcT (0.7  $\mu$ g  $\mu$ L<sup>-1</sup>) or NmLgtA (1.0  $\mu$ g  $\mu$ L<sup>-1</sup>). 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 GalNAc $\beta$ 1–3Gal $\beta$ 1–4Glc $\beta$ ProN<sub>3</sub>

To prepare the trisaccharide, a reaction mixture in Tris-HCl buffer (100 mM, pH 7.5) in a total volume of 5 mL containing Lac $\beta$ ProN<sub>3</sub> (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), MgCl<sub>2</sub> (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:H<sub>2</sub>O: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:H<sub>2</sub>O = 5:2:1, by volume). Further purification by preparative HPLC using C18 column provided the trisaccharide GalNAcβ1-3Galβ1-4GlcβProN<sub>3</sub> as a white powder (58 mg, 98%). <sup>1</sup>H NMR (800 MHz, D<sub>2</sub>O) 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). <sup>13</sup>C NMR (200 MHz, D<sub>2</sub>O) 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 C<sub>23</sub>H<sub>40</sub>N<sub>4</sub>NaO<sub>16</sub> (M+Na) 651.2337, found 651.2335.

#### 4.10. One-pot four-enzyme preparative-scale synthesis of trisaccharide Galβ1-3Galβ1-4Glc

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 (Galβ1-4Glc, 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), MgCl<sub>2</sub> (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 (H<sub>2</sub>O, 0–7 min; 0–50% ACN in H<sub>2</sub>O, 7–35 min; 50% ACN in H<sub>2</sub>O, 35–40min). The pure trisaccharide Galβ1-3Galβ1-4Glc was obtained by H<sub>2</sub>O wash. The synthesis yield of this trisaccharide Galβ1-3Galβ1-4Glc was 10% (15 mg). <sup>1</sup>H NMR (800 MHz, D<sub>2</sub>O) 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); <sup>13</sup>C NMR (200 MHz, D<sub>2</sub>O) 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 C<sub>18</sub>H<sub>32</sub>NaO<sub>16</sub> (M+Na) 527.1588, found 527.1584.

### Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

This work was supported by National Science Foundation grant CHE-1300449 as well as, National Institutes of Health grants R01HD065122 and R01GM094523, National Basic Research Program of China 973 Program No. 2012CB822102, and the Key Grant Project of Chinese Ministry of Education No. 313033.

## References and notes

1. Peng W, Pranskevich J, Nycholat C, Gilbert M, Wakarchuk W, Paulson JC, Razi N. *Glycobiology*. 2012; 22:1453. [PubMed: 22786570]
2. Di Virgilio S, Glushka J, Moremen K, Pierce M. *Glycobiology*. 1999; 9:353. [PubMed: 10089209]
3. Naruchi K, Hamamoto T, Kurogochi M, Hinou H, Shimizu H, Matsushita T, Fujitani N, Kondo H, Nishimura S. *J. Org. Chem.* 2006; 71:9609. [PubMed: 17168577]
4. Nycholat CM, Peng W, McBride R, Antonopoulos A, de Vries RP, Polonskaya Z, Finn MG, Dell A, Haslam SM, Paulson JC. *J. Am. Chem. Soc.* 2013; 135:18280. [PubMed: 24256304]
5. Renkonen O. *Cell. Mol. Life Sci.* 2000; 57:1423. [PubMed: 11078021]
6. Vasiliu D, Razi N, Zhang Y, Jacobsen N, Allin K, Liu X, Hoffmann J, Bohorov O, Blixt O. *Carbohydr. Res.* 2006; 341:1447. [PubMed: 16650392]
7. Sauerzapfe B, Krenek K, Schmiedel J, Wakarchuk WW, Pelantova H, Kren V, Elling L. *Glycoconj. J.* 2009; 26:141. [PubMed: 18758940]
8. Blixt O, Razi N. *Methods Enzymol.* 2006; 415:137. [PubMed: 17116472]
9. Nycholat CM, McBride R, Ekiert DC, Xu R, Rangarajan J, Peng W, Razi N, Gilbert M, Wakarchuk W, Wilson IA, Paulson JC. *Angew. Chem. Int. Ed.* 2012; 51:4860.
10. Chien WT, Liang CF, Yu CC, Lin CH, Li SP, Primadona I, Chen YJ, Mong KK, Lin CC. *Chem. Commun.* 2014; 50:5786.
11. Johnson KF. *Glycoconj. J.* 1999; 16:141. [PubMed: 10612413]
12. Yu H, Lau K, Thon V, Autran CA, 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.
13. Priem B, Gilbert M, Wakarchuk WW, Heyraud A, Samain E. *Glycobiology*. 2002; 12:235. [PubMed: 12042246]
14. Gebus C, Cottin C, Randriantsoa M, Drouillard S, Samain E. *Carbohydr. Res.* 2012; 361:83. [PubMed: 23000215]
15. Blixt O, van Die I, Norberg T, van den Eijnden DH. *Glycobiology*. 1999; 9:1061. [PubMed: 10521543]
16. 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. [PubMed: 17890318]
17. Wakarchuk W, Martin A, Jennings MP, Moxon ER, Richards JC. *J. Biol. Chem.* 1996; 271:19166. [PubMed: 8702594]
18. Logan SM, Altman E, Mykytczuk O, Brisson JR, Chandan V, Schur MJ, St Michael F, Masson A, Leclerc S, Hiratsuka K, Smirnova N, Li J, Wu Y, Wakarchuk WW. *Glycobiology*. 2005; 15:721. [PubMed: 15814825]
19. Guan W, Ban L, Cai L, Li L, Chen W, Liu X, Mrksich M, Wang PG. *Bioorg. Med. Chem. Lett.* 2011; 21:5025. [PubMed: 21704524]
20. Palcic MM. *Curr. Opin. Chem. Biol.* 2011; 15:226. [PubMed: 21334964]
21. Lau K, Thon V, Yu H, Ding L, Chen Y, Muthana MM, Wong D, Huang R, Chen X. *Chem. Commun.* 2010; 46:6066.
22. Wakarchuk WW, Cunningham A, Watson DC, Young NM. *Protein Eng.* 1998; 11:295. [PubMed: 9680191]
23. Park JE, Lee KY, Do SI, Lee SS. *J. Biochem. Mol. Biol.* 2002; 35:330. [PubMed: 12297017]
24. Logan SM, Conlan JW, Monteiro MA, Wakarchuk WW, Altman E. *Mol. Microbiol.* 2000; 35:1156. [PubMed: 10712696]
25. Endo T, Koizumi S, Tabata K, Ozaki A. *Glycobiology*. 2000; 10:809. [PubMed: 10929007]

26. Namdjou DJ, Chen HM, Vinogradov E, Brochu D, Withers SG, Wakarchuk WW. *Chembiochem*. 2008; 9:1632. [PubMed: 18491328]
27. Dumon C, Bosso C, Utille JP, Heyraud A, Samain E. *Chembiochem*. 2006; 7:359. [PubMed: 16381046]
28. Wong CH, Wang R, Ichikawa Y. *J. Org. Chem*. 1992; 57:4343.
29. Palcic MM, Srivastava OP, Hindsgaul O. *Carbohydr. Res*. 1987; 159:315. [PubMed: 3105882]
30. Berliner LJ, Davis ME, Ebner KE, Beyer TA, Bell JE. *Mol. Cell. Biochem*. 1984; 62:37. [PubMed: 6429518]
31. Blixt O, Brown J, Schur MJ, Wakarchuk W, Paulson JC. *J. Org. Chem*. 2001; 66:2442. [PubMed: 11281786]
32. Brockhausen I, Benn M, Bhat S, Marone S, Riley JG, Montoya-Peleaz P, Vlahakis JZ, Paulsen H, Schutzbach JS, Szarek WA. *Glycoconj. J*. 2006; 23:525. [PubMed: 17006644]
33. Boeggeman EE, Balaji PV, Sethi N, Masibay AS, Qasba PK. *Protein Eng*. 1993; 6:779. [PubMed: 8248102]
34. Ramakrishnan B, Shah PS, Qasba PK. *J. Biol. Chem*. 2001; 276:37665. [PubMed: 11485999]
35. Hanada E, Handa S, Konno K, Yamakawa T. *J. Biochem*. 1978; 83:85. [PubMed: 564346]
36. Higuchi R, Inukai K, Zhou JX, Honda M, Komori T, Tsuji S, Nagai Y. *Liebigs Ann. Chem*. 1993:359.
37. Li Y, Yu H, Chen Y, Lau K, Cai L, Cao H, Tiwari VK, Qu J, Thon V, Wang PG, Chen X. *Molecules*. 2011; 16:6396. [PubMed: 21799473]
38. Chen Y, Thon V, Li Y, Yu H, Ding L, Lau K, Qu J, Hie L, Chen X. *Chem. Commun*. 2011; 47:10815.
39. Chen M, Chen LL, Zou Y, Xue M, Liang M, Jin L, Guan WY, Shen J, Wang W, Wang L, Liu J, Wang PG. *Carbohydr. Res*. 2011; 346:2421. [PubMed: 21903203]
40. Muthana MM, Qu J, Li Y, Zhang L, Yu H, Ding L, Malekan H, Chen X. *Chem. Commun*. 2012; 48:2728.
41. Donald AS, Feeny J. *Carbohydr. Res*. 1988; 178:79. [PubMed: 3274083]
42. 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. [PubMed: 26047593]
43. Saito T, Itoh T, Adachi S. *Carbohydr. Res*. 1987; 165:43. [PubMed: 3664520]
44. Urashima T, Saito T, Nishimura J, Ariga H. *Bioch. Biophys. Acta*. 1989; 992:375.
45. Urashima T, Kawai Y, Nakamura T, Arai I, Saito T, Namiki M, Yamaoka K, Kawahawa K, Messer M. *Comp. Biochem. Physiol. C Pharmacol. Toxicol. Endocrinol*. 1999; 124:295. [PubMed: 10661722]
46. Fukuda K, Yamamoto A, Ganzorig K, Khuukhenbaatar J, Senda A, Saito T, Urashima T. *J. Dairy Sci*. 2010; 93:5572. [PubMed: 21094729]
47. Difilippo E, Willems HA, Vendrig JC, Fink-Gremmels J, Gruppen H, Schols HA. *J. Agric. Food Chem*. 2015; 63:4805. [PubMed: 25924866]
48. Kunz C, Rudloff S, Schad W, Braun D. *Br. J. Nutr*. 1999; 82:391. [PubMed: 10673912]
49. Urashima T, Fujita S, Fukuda K, Nakamura T, Saito T, Cowan P, Messer M. *Glycoconj. J*. 2014; 31:387. [PubMed: 24906475]
50. Messer M, Trifonoff E, Stern W, Collins JG, Bradbury JH. *Carbohydr. Res*. 1980; 83:327. [PubMed: 7407802]
51. Urashima T, Taufik E, Fukuda R, Nakamura T, Fukuda K, Saito T, Messer M. *Glycoconj. J*. 2013; 30:801. [PubMed: 23824565]
52. Zhang L, Muthana MM, Yu H, McArthur JB, Qu J, Chen X. *Carbohydr. Res*. 2016; 419:18. [PubMed: 26598987]
53. Muthana MM, Qu J, Xue M, Klyuchnik T, Siu A, Li Y, Zhang L, Yu H, Li L, Wang PG, Chen X. *Chem. Commun*. 2015; 51:4595.
54. Yu H, Yu H, Karpel R, Chen X. *Bioorg. Med. Chem*. 2004; 12:6427. [PubMed: 15556760]
55. Zhang L, Lau K, Cheng J, Yu H, Li Y, Sugiarto G, Huang S, Ding L, Thon V, Wang PG, Chen X. *Glycobiology*. 2010; 20:1077. [PubMed: 20466652]

### Highlights

- *N. meningitidis* and *H. pylori*  $\beta$ 3GlcNAcTs were actively expressed in *E. coli*.
- Bovine  $\beta$ 4GalT 1 was expressed in *E. coli* as a soluble active enzyme.
- Thirteen UDP-sugars could be used as donor substrates for NmLgtA and Hp $\beta$ 3GlcNAcT.
- Ten trisaccharides could be used as acceptors for one of the three  $\beta$ 4GalTs.
- GalNAc $\beta$ 1–3Lac $\beta$ ProN3 and Gal $\beta$ 1–3Lac were produced by OPME reaction containing NmLgtA.

**A.**

```

NmLgtA      : -----MQPIVSVLIGRYNVERYFAQSIAA : 24
Hpβ3GlcNAcT : MTSASSHSFKEQDFHIPIAFADFKNYLIPAGACIYSLESIAKRNKKIRSTLHALVV : 57

NmLgtA      : VVNQTRWN-----LDILIVDDGSTDGTIAIAQRFQEQDGRIRILAQPRNSGLI : 72
Hpβ3GlcNAcT : GLNEEDKTKLNQITEPEPFKFAVLEVKDIEPELDTIPNPFDELETKRESKMLVVKYEL : 114

NmLgtA      : PSINIGLDELAKSGGGGEYIARTDADDDIAAPDWIEK-IVGEMEKDRSIIAMGAWLEV : 128
Hpβ3GlcNAcT : ADLFPKYSKMVWSDVEVIFGNEFSALFLNIKENDENYFYGVLEVEKHHMMECFLECN : 171

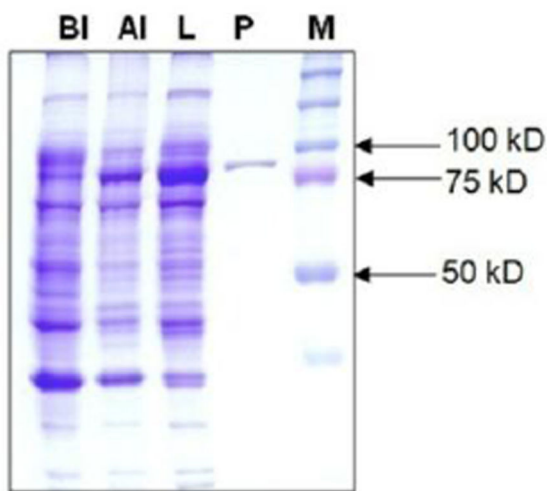
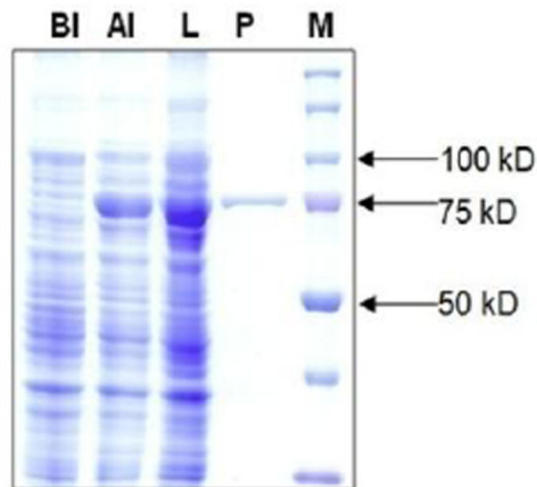
NmLgtA      : LSEEKDGNRLARHHEHGK-----IWKKPTRHEDIADEFEGNPIHNNTM : 172
Hpβ3GlcNAcT : LDYQRKKNETLRMHBLIKGNEARGELDFTKWCWPNMKALGIEYCVFYYVTIKDFSN : 228

NmLgtA      : IMRRSVIDGGRLRYNTERDWAEDYQFWYDVS KLGRLAY-YPEADVKYRPHANQVSSKY : 228
Hpβ3GlcNAcT : AYLNENYKKTILEARENPTIIHYDAWGWAVKPWDYPFGLKALWLNPAKTPFMSDY : 285

NmLgtA      : SIRCHEIAQG-ICKTFRNDFLOSMG-----FKTRFDSLEYRQIKAVAYELLEKHL P : 278
Hpβ3GlcNAcT : TKKMTNESFYTTKNAEQHYFSVKSSKEIVFKAPYLFFKSYLFFVVEKERKIHRSVF : 342

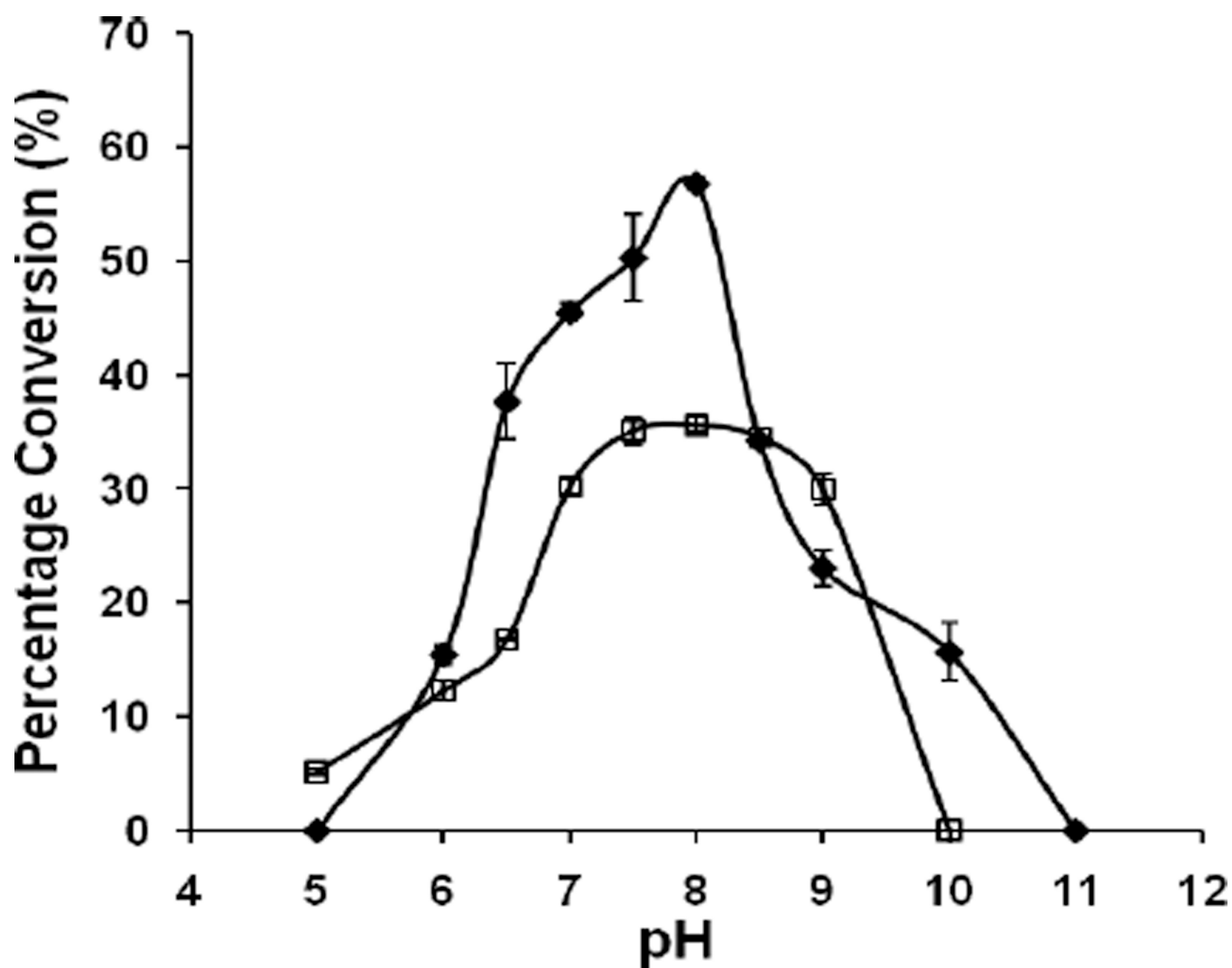
NmLgtA      : EEDFELARRFLYQCFFKRTDTLEAGAWLDFAADG-ERMRRIFETLRQYEGILHRLLRNR : 333
Hpβ3GlcNAcT : ELTCNIVKKEFNKLIYEVFLMEKALVRAVSKILVVLGHHGIVKKIILKLRLLRKG : 398

```

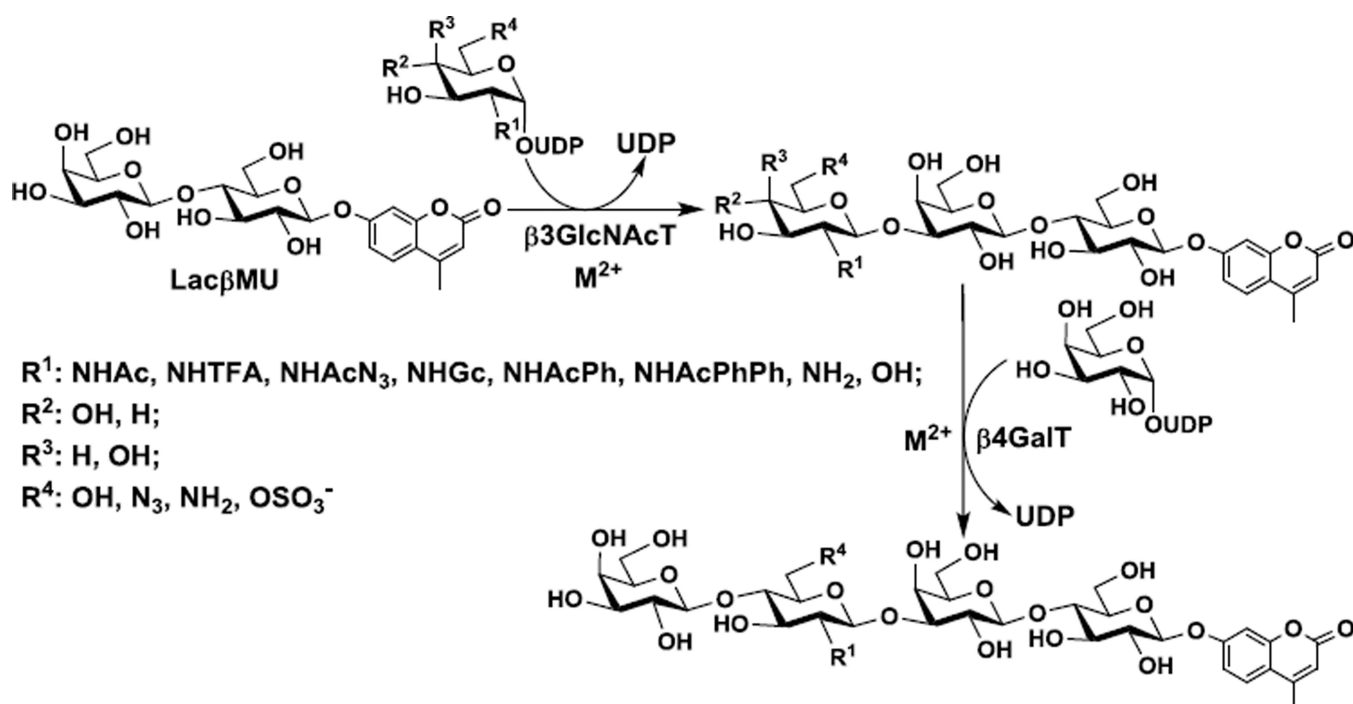
**B.****C.****Figure 1.**

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, Ni<sup>2+</sup>-NTA column purified protein; M, protein markers (Bio-Rad precision Plus Protein Standards, 10–250 kDa).



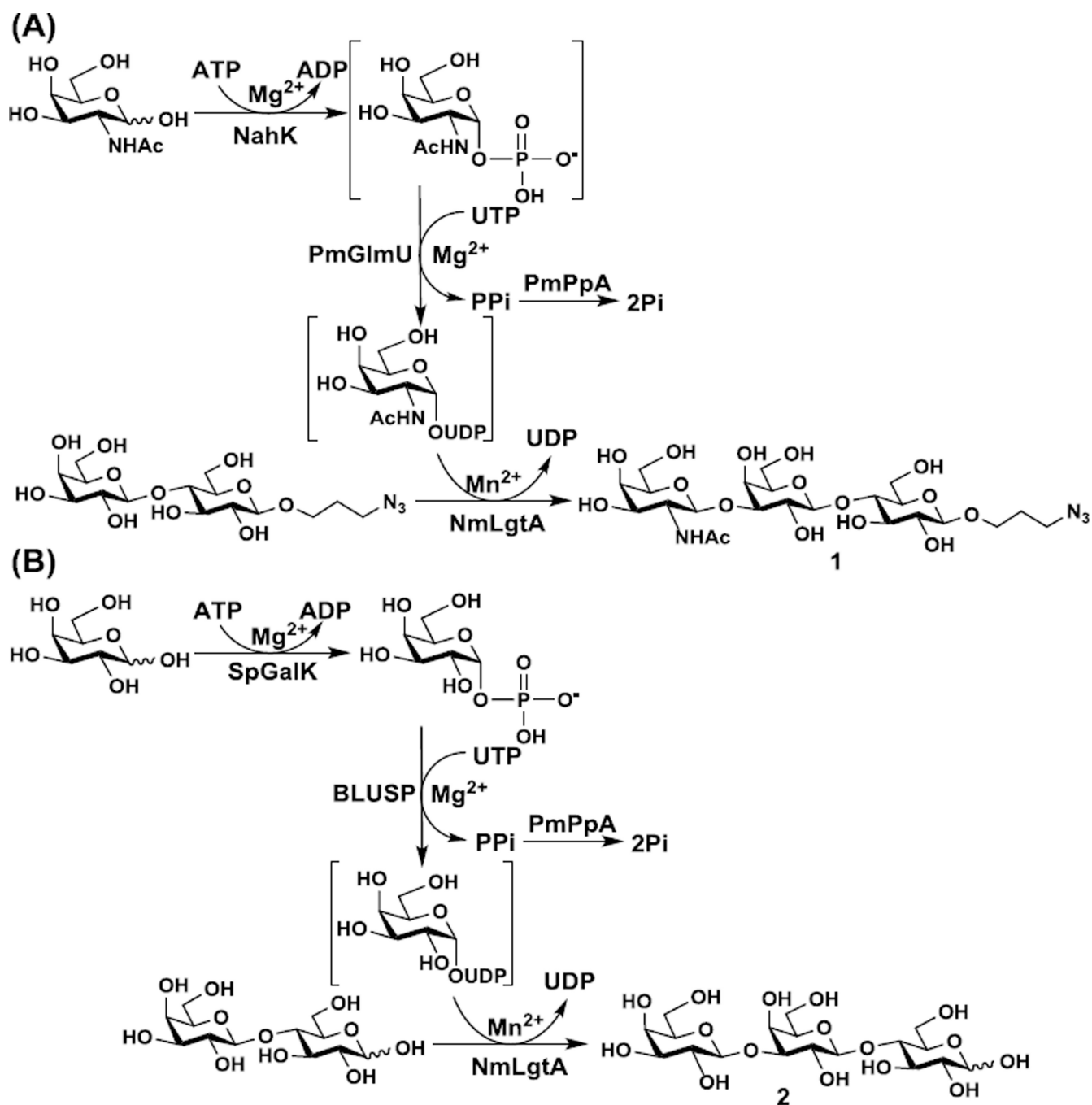


**Figure 2.** pH profiles of NmLgtA (filled diamond) and Hpβ3GlcNAcT (open square). Buffers used were: MES, pH 5.0–6.5; Tris-HCl, pH 7.0–9.0; CAPS, pH 10.0–11.0.



**Scheme 1.**

Donor substrate specificity and acceptor substrate specificity studies of  $\beta 3\text{GlcNAcTs}$  and  $\beta 4\text{GalTs}$ , respectively. Lac $\beta\text{MU}$  was used as an acceptor for  $\beta 3\text{GlcNAcTs}$  (NmLgtA or Hp $\beta 3\text{GlcNAcT}$ ) and the resulting trisaccharide products were used as acceptor substrates for  $\beta 4\text{GalTs}$  (NmLgtB, Hp $\beta 4\text{GalT}$ , or B $\beta 4\text{GalT}$ ).

**Scheme 2.**

One-pot four-enzyme synthesis of GalNAc $\beta$ 1-3Gal $\beta$ 1-4Glc $\beta$ ProN<sub>3</sub> (**23A**) (A) and Gal $\beta$ 1-3Gal $\beta$ 1-4Glc (**25A**) (B). Enzyme abbreviations: NahK, *Bifidobacterium longum* ATCC55813 *N*-acetylhexosamine 1-kinase;<sup>37</sup> PmGlmU, *Pasteurella multocida* *N*-acetylglucosamine-1-phosphate uridylyltransferase;<sup>38</sup> PmPpA, *Pasteurella multocida* inorganic pyrophosphatase;<sup>21</sup> SpGalK, *Streptococcus pneumoniae* TIGR4 galactokinase;<sup>39</sup> BLUSP, *Bifidobacterium longum* UDP-sugar pyrophosphorylase.<sup>40</sup>

Donor substrate specificities of Hpb3GlcNAcT 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.

Table 1

#	Donors	NmLgtA		Hpb3GlcNAcT	
		HRMS <sup>a</sup>	CE <sup>b,c</sup> or HPLC <sup>d</sup>	HRMS <sup>a</sup>	CE <sup>b</sup> or HPLC <sup>d</sup>
1	UDP-GlcNAc	90–95%	81.2±0.8% <sup>b</sup>	55–60%	66.2±1.3% <sup>b</sup>
2	UDP-GlcNTFA	100%	100±0% <sup>b</sup>	90–95%	92.9±0.3% <sup>b</sup>
3	UDP-GlcNAcN <sub>3</sub>	25–35%	29.1±0.1% <sup>b</sup>	20–30%	26.3±0.4% <sup>b</sup>
4	UDP-GlcNGc	30–40%	36.1±0.8% <sup>b</sup>	55–60%	77.5±0.1% <sup>b</sup>
5	UDP-GlcNAcPh	20–25%	12.9±0.5% <sup>b</sup>	0	0
6	UDP-GlcNAcPhPh	<5%	N.D.	0	0
8	UDP-GlcNH <sub>2</sub>	20–25%	peaks together <sup>d</sup>	0	0
10	UDP-GlcNAc6NH <sub>2</sub>	20–25%	peaks together <sup>d</sup>	70–75%	peaks together <sup>d</sup>
11	UDP-GlcNAc6N <sub>3</sub>	0	0	30–35%	27.8±1.1% <sup>b</sup>
16	UDP-GlcNAc6S	0	0	65–70%	peaks together <sup>d</sup>
21	UDP-Glc	20–25%	23.2±2.6% <sup>b</sup>	0	0
23	UDP-GalNAc	80–90%	peaks together <sup>d</sup>	0	0
25	UDP-Gal	40–45%	39.6±0.6% <sup>c</sup>	0	0

<sup>a</sup> Yields estimated by HRMS.

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

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

<sup>d</sup> 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;

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 2**

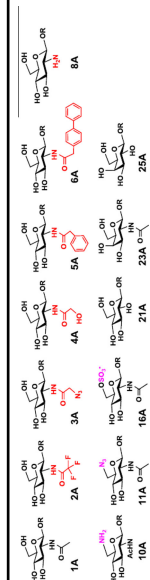
Apparent kinetic data for Hp $\beta$ 3GlcNAcT and NmLgtA.

Substrate	NmLgtA			Hp $\beta$ 3GlcNAcT		
	$k_{cat}$ (s <sup>-1</sup> )	$K_M$ (mM)	$k_{cat}/K_M$ (s <sup>-1</sup> mM <sup>-1</sup> )	$k_{cat}$ (s <sup>-1</sup> )	$K_M$ (mM)	$k_{cat}/K_M$ (s <sup>-1</sup> mM <sup>-1</sup> )
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 <sub>3</sub> (3)	0.15 ± 0.03	1.34 ± 0.24	0.11	0.29 ± 0.07	4.74 ± 1.09	0.062

Table 3

Acceptor substrate specificities of  $\beta$ 4GalTs using a sequential two-step process.<sup>a</sup>

#	Acceptor	$\beta$ 3GlcNAcT <sup>*</sup>	NmLgtB	H $\beta$ 4GalT	B $\beta$ 4GalT
1A	GlcNAc $\beta$ 1-3Lac $\beta$ MU	90–95%	NmLgtA	100%	90–95%
2A	GlcNTFA $\beta$ 1-3Lac $\beta$ MU	100%	NmLgtA	100%	60–65%
3A	GlcNAc $\beta$ 1-3Lac $\beta$ MU	25–35%	NmLgtA	100%	35–40%
4A	GlcNGc $\beta$ 1-3Lac $\beta$ MU	55–60%	H $\beta$ 3GlcNAcT	75–80%	0
5A	GlcNAcPh $\beta$ 1-3Lac $\beta$ MU	20–25%	NmLgtA	50%	20–25%
6A	GlcNAcPhPh $\beta$ 1-3Lac $\beta$ MU	<5%	NmLgtA	0	0
8A	GlcNH $\beta$ 1-3Lac $\beta$ MU	20–25%	NmLgtA	0	0
10A	GlcNAc6NH $\beta$ 1-3Lac $\beta$ MU	70–75%	H $\beta$ 3GlcNAcT	20%	0
11A	GlcNAc6N $\beta$ 1-3Lac $\beta$ MU	30–35%	H $\beta$ 3GlcNAcT	100%	0
16A	GlcNAc6S $\beta$ 1-3Lac $\beta$ MU	65–70%	H $\beta$ 3GlcNAcT	100%	0
21A	Glc $\beta$ 1-3Lac $\beta$ MU	20–25%	NmLgtA	0	0
23A	GalNAc $\beta$ 1-3Lac $\beta$ MU	80–90%	NmLgtA	0	0
25A	Gal $\beta$ 1-3Lac $\beta$ MU	40–45%	NmLgtA	0	0



<sup>a</sup>Yields estimated by HRMS. R= Lac $\beta$ MU.

\* Lac $\beta$ MU was used as the acceptor for  $\beta$ 3GlcNAcT-catalyzed reactions. The trisaccharide producing reactions with a higher yield (catalyzed by either H $\beta$ 3GlcNAcT or NmLgtA) were used for the acceptor substrate specificity studies of  $\beta$ -4GalTs.