ACADEMIC PRESS

Biochemical and Biophysical Research Communications 295 (2002) 1–8

BBRC

www.academicpress.com

Overexpression and biochemical characterization of β-1,3-*N*-acetylgalactosaminyltransferase LgtD from *Haemophilus influenzae* strain Rd

Jun Shao, Jianbo Zhang, Przemyslaw Kowal, Yuquan Lu, and Peng George Wang*

Department of Chemistry, Wayne State University, Detroit, MI 48202, USA

Abstract

The lipopolysaccharide of capsule deficient *Haemophilus influenzae* strain Rd contains an *N*-acetylgalactosamine residue attached to the terminal globotriose moiety in the Hex5 glycoform. Genome analysis identified an open reading frame HI1578, referred to as *lgtD*, whose amino acid sequence shows significant level of similarity to a number of bacterial glycosyltransferases involved in lipopolysaccharide biosynthesis. To investigate its function, overexpression and biochemical characterization were performed. Most of the protein was obtained in a highly soluble and active form. By using standard glycosyltransferase assay and HPLC, we show that LgtD is an *N*-acetylgalactosaminyltransferase with high donor substrate specificity and globotriose is a highly preferred acceptor substrate for the enzyme. The K_m for UDP-GalNAc and globotriose are 58 μM and 8.6 mM, respectively. The amino acid sequence of the enzyme shows the conserved features of family II glycosyltransferases. This is the first *N*-acetylgalactosaminyltransferase identified from *H. influenzae*, which shows potential application in large-scale synthesis of globo-series oligosaccharides.

Keywords: *Haemophilus influenzae*; Lipopolysaccharide; *N*-Acetylgalactosaminyltransferase; UDP-GalNAc; Globotetraose

Haemophilus influenzae is a gram-negative human pathogen that routinely colonizes the upper respiratory tract. Serotype b capsular strains are associated with invasive diseases such as meningitis, septicemia, epiglottitis, pneumonia, and empysema, particularly in infants. Acapsular or nontypable strains are a common cause of otitis media, sinusitis, and lower respiratory infections, which account for many millions of childhood deaths in developing countries [1,2]. Like other gram-negative bacteria [3,4], the outer surface of *H. influenzae* outer membrane consists predominantly of lipopolysaccharide (LPS). LPS provides the organism with a permeability barrier to certain antibacterial agents and is implicated as a major virulence factor. Both capsular and acapsular strains of *H. influenzae* express heterogeneous populations of low-molecular-mass LPS which exhibit extensive antigenic diversity. LPS can mimic host glycolipids and has a propensity for reversible switching of expression of terminal epitopes (phase variation) of the oligosaccharide portion, which is advantageous for the survival of this bacterium when it is confronted by the differing microenvironments and immune responses presented by the host during the pathogenesis of disease [5,6]. The structural diversity of *H. influenzae* LPS arising from phase variation has complicated the study of the molecular features of these molecules and the understanding of their role in pathogenic interactions with the host.

Amino sugars are recognized as very important carbohydrates in many biological systems. They are constituents of structures found in all kinds of organisms, mostly on the surface of cells and in the spaces between them, forming the substance which binds cells together, membranes which envelope them, and protective layers which cover them. Pathogenic bacteria such as *Neisseria meningitides* [7–9], *Helicobacter pylori* [10–12], and *Campylobacter jejuni* [13–16] have been found to carry amino sugars in their lipopolysaccharides. The presence of *N*-acetylglucosamine as a minor component of LPS has been reported in *H. influenzae* type b strain A2 [17]

* Corresponding author. Fax: +1-313-577-2554.
E-mail address: pwang@chem.wayne.edu (P. George Wang).

0006-291X/02/$ - see front matter
PII: S0006-291X(02)00615-0

and in the related species *Haemophilus ducreyi* [18]. In a capsule deficient *H. influenzae* strain Rd, structural analysis of LPS epitopes [19,20] has shown *N*-acetylgalactosamine to be attached to a terminal globotriose moiety in the Hex5 glycoform, giving the globotetraose unit, GalNAcβ1–3Galα1–4Galβ1–4Glc, that corresponds to the P antigen expressed in mammalian globo series of glycolipids. The structural data provide the first evidence demonstrating the involvement of *N*-acetylgalactosaminyltransferase activity in the synthesis of *H. influenzae* LPS.

Haemophilus influenzae LPS is composed of a membrane-anchoring lipid A portion linked by a single 2-keto-3-deoxyoctulosonic acid (KDO) molecule to heterogeneous oligosaccharide composed of neutral heptose and hexose sugars [7]. The biosynthesis of this complex macromolecule includes many processes such as the activation of sugars, initiation of glycolipid, elongation of oligosaccharides, and transportation of the mature glycolipid. Each process requires diverse functions of numerous different proteins. The availability of complete genome sequence of *H. influenzae* strain Rd [21] has facilitated significant progress in the comprehensive study of cell surface lipopolysaccharides in *H. influenzae*. Based on homologous gene comparisons, open reading frame HI1578 has previously been suggested to encode a putative glycosyltransferase that is involved in LPS biosynthesis, but the construction of strains with a mutant gene has proved extremely difficult [22]. Recently, a fingerprinting strategy is employed to establish the structure of LPS from strains with mutated putative glycosyltransferase genes [19] and HI1578 gene, *lgtD*, appears to code for an *N*-acetylgalactosaminyltransferase involved in LPS extension. To determine its enzymatic function and explore whether the *lgtD* gene product could be useful in the in vitro oligosaccharide synthesis, we cloned the *lgtD* gene from *H. influenzae* strain Rd and overexpressed it in *Escherichia coli*. Biochemical characterization of the recombinant protein confirms that the gene encodes a β-1,3-*N*-acetylgalactosaminyltransferase (GalNAcT) which is responsible for the addition of *N*-acetylgalactosamine to terminal globotriose moiety in the LPS of *H. influenzae* strain Rd. Moreover, determination of the biochemical properties shows the potential application of this bacterial glycosyltransferase in large-scale synthesis of globo-series oligosaccharides.

Materials and methods

Materials. Restriction enzymes, T4 DNA ligase and IPTG were obtained from Promega (Madison, WI). Vent DNA polymerase was purchased from New England Biolabs (Beverly, MA). Ni^{2+}-NTA agarose, PCR purification kit, QIAEX II gel extraction kit, and DNA miniprep spin kit were from Qiagen (Santa Clarita, CA). Low range protein standard was from Bio-Rad. UDP-D-[1-^{3}H(N)]GalNAc, UDP-D-GlcNAc, UDP-D-GalNAc, and ampicillin were obtained from Sigma Chemical (St. Louis, MO). UDP-D-[6-^{3}H(N)]GlcNAc, UDP-D-[6-^{3}H]galactose, and UDP-D-[6-^{3}H]glucose were from Amersham Pharmacia Biotech. Other reagents were of analytical or higher grade. All kits or enzymes were used following manufacturer's instructions.

Bacterial strains and plasmids. Escherichia coli competent cell DH5α [*lacΔM15 hsdR recA*] was from Gibco-BRL Life Technology. Plasmid vector pET15b and *E. coli* competent cell BL21 (DE3) [F^- *ompT hsdS*$_B$ ($r_B^- m_B^-$) *gal dcm* (DE3)] were from Novagen (Madison, WI). *H. influenzae* type d strain RM118 [KW-20] chromosomal DNA (ATCC 51907D) was purchased from ATCC.

Cloning and expression of N-acetylgalactosaminyltransferase from H. influenzae. HI1578 gene was cloned into the *Nde*I and *Xho*I sites of pET15b plasmid with N-terminal histidine tag. A pair of primers: P_1 5′CGCCATATGGAAAATTGTCCATTAGTATCG3′/P_2 5′CCGCTCGAGCTAAAATATAACATTTATTTT3′, with *Nde*I and *Xho*I restriction sites (underlined) incorporated into P_1 and P_2 respectively, were used for polymerase chain reaction (PCR) amplification of the putative gene from the chromosomal DNA of *H. influenzae*. PCR consisted of 50 ng genomic DNA, 200 nM of each primer, 0.2 mM of each dNTP, 15 mM $MgCl_2$, 0.5 U vent DNA polymerase, and 1 × buffer in a total of 50 μl. After heating at 94 °C for 2 min, 25 cycles were carried out including 30 s at 94 °C, 1 min at 55 °C, and 90 s at 72 °C. A final 10 min elongation was performed at 72 °C. The DNA fragment obtained was digested with *Nde*I/*Xho*I and inserted into pET15b linearized by the same restriction enzymes. The construct was subsequently transformed into the *E. coli* DH5α cloning host strain and then the BL21 (DE3) expression strain with ampicillin (100 μg/ml) selection. Selected clones were characterized by restriction mapping.

For protein expression, *E. coli* BL21 (DE3) harboring the recombinant plasmid was grown in LB medium at 37 °C. When $A_{600\,nm}$ reached 0.8, isopropyl-β-D-thiogalactoside (IPTG) was added to a final concentration of 0.2 mM and expression was allowed to proceed for 8 h at 30 °C. Cells were harvested by centrifugation at 4 °C and washed twice with 50 mM Tris–HCl buffer, pH 7.0. The cell pellet was resuspended in chilled lysis buffer (50 mM Tris–HCl, pH 7.0, 0.1 M NaCl, 0.5% (w/v) Triton X-100, 10% glycerin (w/v), and 10 mM 2-mercaptoethanol) and disrupted by brief sonication (Branson Sonifier 450, VWR Scientific) on ice. The lysate was cleared by centrifugation (12,000*g*, 20 min, 4 °C) and loaded at a flow rate of 2 ml/min onto a Ni^{2+}-NTA agarose column (Qiagen, CA). Protein eluted from the column was nearly homogenous as seen by SDS–PAGE analysis. Further purification was carried out by ÄKTA FPLC system using HiLoad_16/60_Superdex 200 Column (Amersham Pharmacia Biotech), equilibrated, and eluted with 50 mM Tris–HCl buffer (pH 7.0) containing 20% glycerin, 10 mM 2-mercaptoethanol, and 150 mM NaCl.

Enzymatic assays. The standard *N*-acetylgalactosaminyltransferasereactions were performed at 30 °C for 20 min in a final volume of 100 μl containing 50 mM Tris–HCl (pH 7.5), 10 mM $MnCl_2$, 0.1% bovine serum albumin, 1 mM dithiothreitol, 0.3 mM UDP-D-[1-^{3}H]GalNAc (final specific activity of 1000 cpm/nmol), 3 mM acceptor globotriose or lactose, and various amounts of purified enzyme. Acceptor was omitted for blank. Globotriose was prepared using a whole-cell reaction system as previously described [23]. The reaction was terminated by adding 100 μl ice-cold 0.1 M EDTA. Dowex 1 × 8–200 chloride anion exchange resin was then added in a water suspension (0.8 ml, 1:1 (v/v)). After centrifugation at 10,000*g* for 5 min, supernatant (0.5 ml) was collected in a 20 ml plastic vial and 5 ml ScintiVerse BD was added. The vial was vortexed thoroughly before the radioactivity of the mixture was counted in a liquid scintillation counter (Beckmann LS-3801 counter). One unit of activity is defined as the amount of enzyme that catalyzes the transfer of 1 μmol GalNAc from UDP-GalNAc to acceptor/min at 30 °C.

In evaluating k_{cat} and K_m in steady-state kinetic studies, every effort was made to examine substrate concentrations having values ranging from approximately 0.5 to 4 times the K_m. In all cases, apparent

steady-state parameters for a given substrate were determined in the presence of saturating amounts of other substrates, usually at concentrations between 5 and 10 times the K_m.

Protein concentration was determined according to the method of Bradford [24]. Protein isoelectric point (pI) and theoretical molecular weight (MW) were calculated with ProtParam tool from ExPASy proteomics server.

HPLC assays. These assays were performed by ProStar HPLC system (Varian, CA) to identify the products synthesized by the purified enzyme. The reaction mixture contained 50 mM Tris–HCl (pH 7.5), 10 mM $MnCl_2$, 3 mM UDP-GalNAc, 3 mM globotriose or lactose, and purified enzyme in a final volume of 0.3 ml. HPLC sample was prepared by adding 0.3 ml acetonitrile to the reaction mixture. The mixture was cleared by centrifugation (12,000*g*, 10 min). Then 20 μl sample was injected into a Microsorb NH_2 100 Å column (4.6 × 250 mm, Varian, CA), with an eluent formed by 65% acetonitrile and 35% H_2O at a flow rate of 1.0 ml/min. Oligosaccharide peaks were recorded with a Star 9040 Refractive Index Detector (Varian, CA).

Preparative synthesis using purified N-acetylgalactosaminyltransferase. To obtain sufficient material for structural analysis, a preparative reaction was conducted with globotriose as acceptor. The reaction was optimized so that the substrate, UDP-GalNAc, was converted completely to the product. The high yield simplified the isolation of oligosaccharide product for further analysis. Briefly, the reaction was conducted at 30 °C in a final volume of 1.5 ml containing 50 mM Tris–HCl (pH 7.5), 10 mM $MnCl_2$, 1 mM dithiothreitol, 10 mM UDP-GalNAc, and 12 mM acceptor globotriose. The reaction was initiated by addition of 15 mU purified GalNAcT. The progress of the reaction was monitored by thin-layer chromatography [i-PrOH/H_2O/NH_4OH = 7:3:2 (v/v/v)] conducted on Baker Si250F silica gel TLC plates with a fluorescent indicator. After complete conversion of donor substrate, protein was removed by brief boiling, followed by centrifugation (12,000*g*, 20 min). Then oligosaccharide products were separated and purified by Bio-Gel P-2 gel filtration (Bio-Rad, CA) with water as mobile phase. The desired fractions were pooled, lyophilized, and finally stored at −20 °C.

Mass spectrometry and NMR spectroscopy. Mass spectra (ESI) were run in the negative model at the mass spectrometry facility of Wayne State University. Product structure was identified by ^1H and ^{13}C NMR spectroscopy using 400-MHz Mercury and 500-MHz Varian NMR spectrometer. Chemical shifts (ppm) were given downfield from the methyl resonance of acetone set at 2.225 ppm for ^1H and 30.7 ppm for ^{13}C, respectively.

Computational analysis of protein sequences. Amino acid sequence homology searches were carried out in the National Center for Biotechnology Information databases with the advanced BLAST 2.0 (basic logical alignment search tool) software. Multiple sequence alignments were performed with Clustalw 1.8 by Thompson et al. [25]. The alignment was rendered using BioEdit 5.0.6 software by Tom Hall.

Results

Preparation and properties of recombinant protein

Amplification by PCR of the *lgtD* gene from *H. influenzae* type d strain RM118 [KW-20] and its insertion into the expression vector pET15b yielded the plasmid pET15b-LgtD. The gene product was overexpressed in *E. coli* BL21 (DE3) with a N-terminal His_6-tag. When expression was carried out with low concentration IPTG (0.2 mM) induction at 30 °C, the expression level was high (about 45% of total cellular proteins) and most of the protein was in a soluble form (Fig. 1). It was easily purified to 85–90% by one-step nickel affinity agarose column. Most of the contaminants were further eliminated by HiLoad_16/60_Superdex 200 column to produce up to 95% pure protein. Therefore, the protein was purified 2.5-fold to reach homogeneity (Table 1). The recombinant protein has high isoelectric point (pI = 7.28) and an apparent molecular weight of 40 kDa estimated by SDS–PAGE, similar to the theoretical value (39843.9 kDa) calculated from its predicted amino acid sequence.

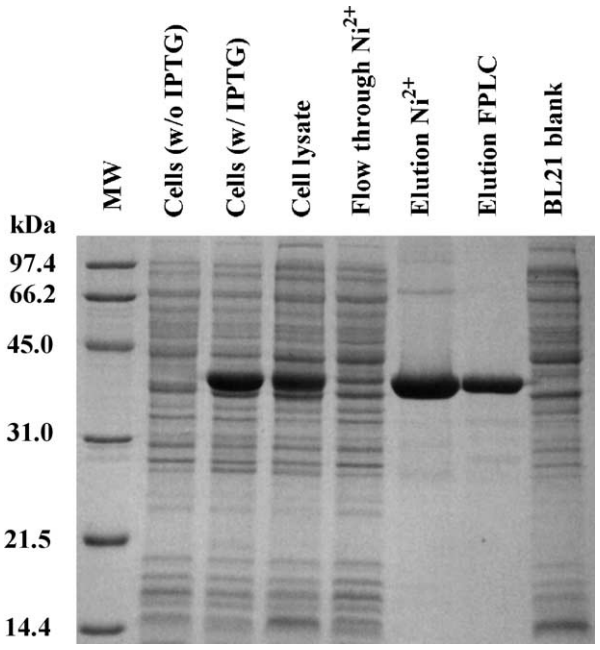

Fig. 1. SDS–PAGE analysis for expression and purification of recombinant protein. Ten μl aliquots were withdrawn at each step of the purification and loaded on a 12% SDS–PAGE gel in Mini Protein III cell gel electrophoresis unit (Bio-Rad). The detection was performed with Coomassie blue staining. MW, low range (14–98 kDa) molecular weight markers (Bio-Rad).

Identification and characterization of N-acetylgalactosaminyltransferase activity

To investigate the function of *lgtD* gene from *H. influenzae* strain Rd, the purified recombinant enzyme was used to test *N*-acetylgalactosaminyltransferase activity with either globotriose (Gb_3) or lactose as an acceptor. The enzyme was capable of catalyzing the transfer of a GalNAc residue from UDP-GalNAc to both of the acceptors. Although lactose (Galβ1–4Glc) unit is abundantly expressed in bacterial LPS and is a highly favored substrate for some glycosyltransferases [26], this enzyme showed a 5-fold higher reaction rate with globotriose versus lactose (Fig. 2). The acceptor pattern is in agreement with the biosynthesis of the LPS oligosaccharide epitopes of the *H. influenzae* strain Rd, in which a β-GalNAc residue is 1–3 linked to a globotriose unit.

Table 1
Purification of *N*-acetylgalactosaminyltransferase from *H. influenzae* strain Rd

Stages	Total protein (mg)	Recovery (%)	Total activity[a] (units)	Specific activity (units/mg)	Purification (fold)
Homogenate[b]	38.6	100	27	0.7	1
Cell lysate[b]	27.8	72	25	0.9	1.28
Nickel affinity	8.5	22	12.7	1.5	2.14
Gel filtration	5.8	15	10.4	1.8	2.57

[a] The reactions were performed using 10 μl cell or enzyme extract in a total volume of 100 μl under the standard assay conditions.
[b] Homogenate and cell lysate from *E. coli* BL21 (DE3) strain harboring the empty pET15b plasmid were used as controls for these enzymatic activity assays.

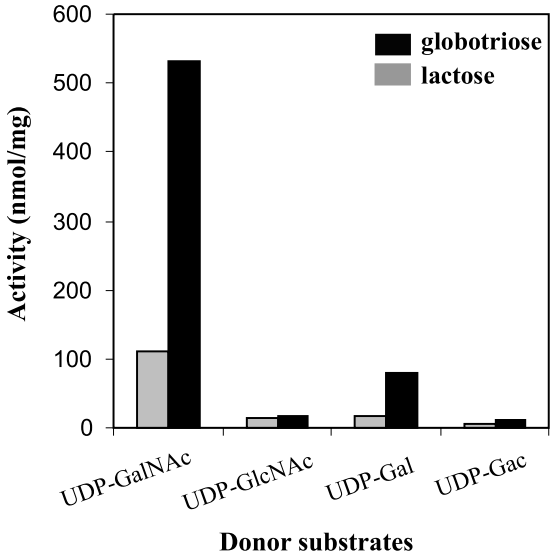

Fig. 2. Substrate specificity of recombinant GalNAcT from *H. influenzae* Rd. Various nucleotide sugar donor substrates were tested for enzymatic activity under standard conditions with either globotriose or lactose as acceptor. The values under 50 nmol/mg protein are not significant therefore, the substrates are not proper donors under these conditions.

To investigate the donor specificity, several other UDP-sugars were tested as donor substrates for the recombinant GalNAcT using standard activity assay (Fig. 2). When globotriose was used as acceptor, UDP-Glc and UDP-GlcNAc were not donors. UDP-Gal, on the other hand, could be a poor donor with reaction efficiency 9-fold lower than that of UDP-GalNAc. However, none of these UDP-sugars were suitable as donors if lactose was used as an acceptor.

The enzyme activity required divalent cations, primarily Mn^{2+} or Mg^{2+}, and was abolished with 10 mM EDTA, as shown in Fig. 3A. Since Mn^{2+} was the most effective activator, it was used in standard assays. The recombinant enzyme showed a broad pH range, but it was most active in the pH range of 7.0–7.5 (Fig. 3B). The optimal reaction temperature for the enzyme was between 28 and 30 °C (data not shown). The purified enzyme appeared stable for at least two months when stored at 4 °C in 20% (w/v) glycerin in the presence of 10 mM 2-mercaptoethanol.

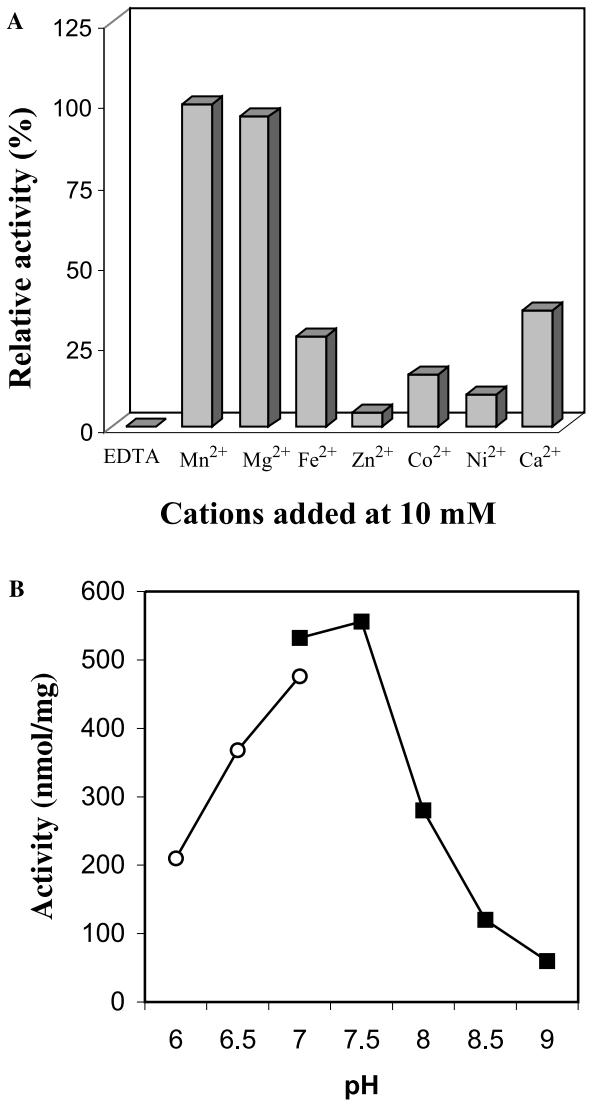

Fig. 3. Effects of cations (A) and pH (B) on the activity of recombinant GalNAcT from *H. influenzae* Rd. (A) Assays were done under standard conditions with different divalent cation chlorides or EDTA at final concentrations of 10 mM. (B) The pH dependency of β3GalNAcT activity was measured using MES buffer (open circles) and Tris–HCl buffer (solid squares) at a final concentration of 50 mM.

Table 2 depicts the results of preliminary kinetic analysis of *N*-acetylgalactosaminyltransferase activity. UDP-GalNAc and globotriose are apparently the preferred substrates for the recombinant GalNAcT,

Table 2
Preliminary kinetic parameters of GalNAcT from *H. influenzae* strain Rd

Substrates	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($\mu M^{-1}\ s^{-1}$)
UDP-GalNAc	58	7.8	1.34×10^{-1}
UDP-Gal	1.1×10^3	1.5	1.36×10^{-3}
Globotriose	8.6×10^3	18.6	2.16×10^{-3}
Lactose	3.3×10^4	1.7	5.15×10^{-5}

displaying K_m values of 58 μM and 8.6 mM, respectively, and high turnover numbers. UDP-Gal appears to be a reasonable substrate for the enzyme with a k_{cat} of 1.5 s^{-1}, nearly 20% of that of UDP-GalNAc. However, its K_m value (1.1 mM) is much higher and overall the recombinant GalNAcT shows a 98-fold kinetic preference of UDP-GalNAc to UDP-Gal.

HPLC analysis of the GalNAcT reaction products

Reaction products synthesized from different substrates by the purified GalNAcT were further characterized by HPLC with refractive index detection. Under analytical conditions, globotetraose and globotriose are well resolved with peaks at 14.8 and 13.3 min, respectively (Fig. 4). These results confirmed the data from the activity assay using UDP-D-[1-^{3}H(N)]GalNAc as donor.

Preparative synthesis and structure analysis

With the purified GalNAcT protein we could use globotriose as the acceptor for 10 mg scale reaction. The reaction was conducted at 30 °C for 20 h. Oligosaccharide products were separated and purified by Bio-Gel P-2 gel filtration. Total 9.6 mg oligosaccharide product was collected to allow structure analysis by ESI-MS spectrometry and NMR spectroscopy. The product was repeatedly exchanged in D_2O (Aldrich Chemical) with intermediate lyophilization before ^{1}H and ^{13}C NMR spectra were recorded at 298 K in a 5 mm tube.

The ESI-MS spectrum in the negative mode is shown in Fig. 5. The structure of the proposed globotetraose product is shown in the inset. The prominent peak at *m/z* 705.83 is interpreted as the $(M–H)^-$ of the parent compound, given the predicted molecular weight of 706.8 for globotetraose. The result confirmed that only one GalNAc residue was attached to globotriose by the recombinant GalNAc transferase.

In the ^{1}H and ^{13}C NMR spectra (data not shown), signals were found for a newly introduced GalNAc residue. The chemical shift of the new anomeric hydrogen at 4.54 ppm and *N*-acetyl group at 1.95 ppm, as long as the chemical shift of 22.40 ppm and 175.37, 183.92 ppm for carbon in *N*-acetyl group indicated clearly that a GalNAc residue was attached to both

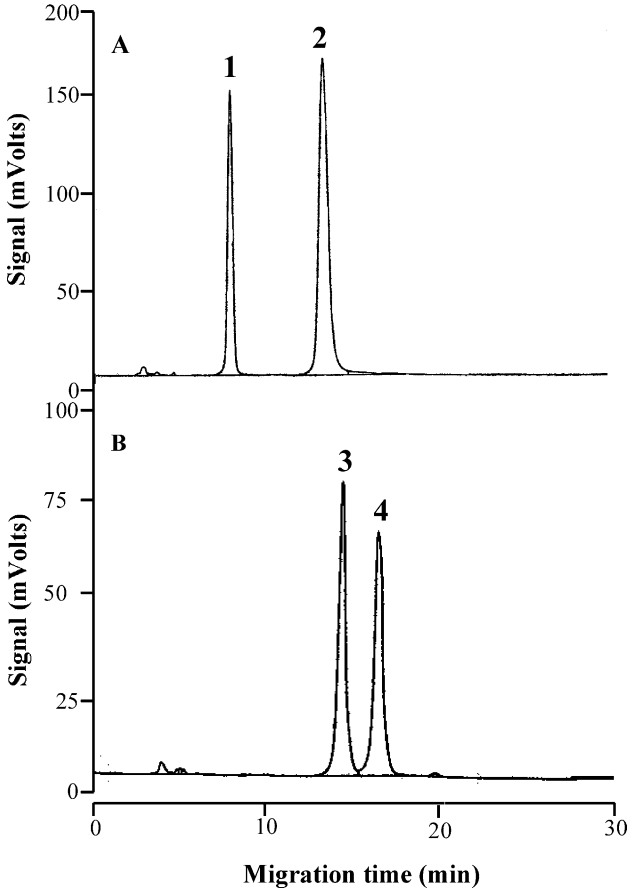

Fig. 4. Identification of the reaction products of GalNAcT from *H. influenzae* Rd by HPLC. The reactions were carried out as described in Materials and methods. (A) Reaction with lactose as acceptor: 1, lactose; 2, GalNAcβ1,3Lac. (B) Reaction with globotriose as acceptor: 3, globotriose; 4, globotetraose.

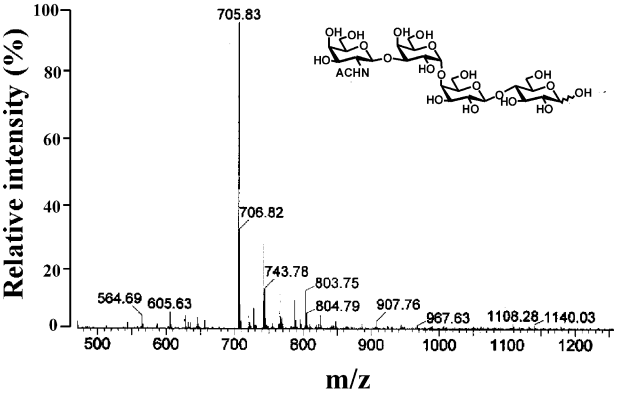

Fig. 5. ESI mass spectrometry of the product generated by the recombinant GalNAcT. Spectra were acquired in the negative mode. The molecular weight of globotetraose is 706.8. The ion peak at *m/z* 705.83 is attributed to the molecular ion $[M–H]^-$ of globotetraose.

anomers of the acceptor molecule. A coupling constant of 8.5 Hz for the H-1 resonance and the chemical shift of 103.49 ppm for the C-1 resonance of GalNAc showed

that this residue was in β-anomeric configuration. The significant downfield increments in the chemical shifts of the C-3″ resonance of Gal ($\Delta\delta = 9.7$ ppm) of the acceptor, but not in the other resonances, indicated that the GalNAc residue had been introduced to C-3 of Gal. This linkage was further confirmed by the crosspeaks between the anomeric carbon of GalNAc and the H-3 (δ 3.50) of the Gal residue in HMBC spectrum (data not shown). The β-anomeric configuration of GalNAc residue was also evident from the crosspeaks between the anomeric proton and the anomeric carbon of GalNAc in the HMQC spectrum (data not shown) and the crosspeaks between the anomeric proton of GalNAc (H-1, δ 4.54) and the H-3 and H-4 (δ 3.80) protons of the accepting Gal residue in the NOESY spectrum (data not shown). Hence, the structure of the product was assigned to be GalNAcβ1–3Galα1–4Galβ1–4GlcOH.

Discussion

Lipopolysaccharide is a major component of the cell wall of human pathogen *H. influenzae*, whose capacity to express heterogeneous glycoforms is important for the pathogenesis of disease in its host [27]. By using classical genetic analysis or genome sequence search, several glycosyltransferases involved in *H. influenzae* LPS biosynthesis have been characterized such as α-2,3-sialyltransferase Lic3A [28], heptosyltransferase II RfaF [29], galactosyltransferase LgtC [22], and Kdo transferase WaaA [30]. HI1578 is an open reading frame of *H. influenzae* strain Rd genome. Hood et al. [22] previously postulated that HI1578 gene might be involved in the attachment of *N*-acetylglucosamine residues to the outer portion of the lipid A molecule because of its sequence similarity to the *lgtA* gene from *Neisseria gonorrhoeae*. However, no GlcNAc residue has ever been found in LPS structure analysis of *H. influenzae* strain Rd. The GalNAc residue attached to a globotriose epitope in the Hex5 glycoform of LPS [20] was the only amino sugar found in this strain. Overexpression and biochemical characterization performed in this study demonstrate that HI1578 is an *N*-acetylgalactosaminyltransferase structural gene involved in LPS oligosaccharide biosynthesis. The in vitro synthesis of globotetraose and NMR structure analysis demonstrate that the enzyme forms a β glycosidic linkage between C1 of GalNAc and C3 of the terminal Gal of globotriose acceptor. Although the recombinant enzyme can transfer a galactosyl residue from UDP-Gal to globotriose in vitro, the poor efficiency of catalysis indicates that this reaction is unlikely to happen in vivo. Site-directed mutagenesis and protein structure analysis will help interpret the molecular basis for substrate specificity of this enzyme.

The known glycosyltransferases have been classified into 47 different families, based on both sequence similarity and substrate/product stereochemistry [31–33]. The GalNAc transferase encoded by HI1578 gene belongs to family II, whose members range from the polysaccharide synthases responsible for the production of polymers, such as cellulose and chitin, to enzymes involved in bacterial cell surface glycosylation. These enzymes are expected to use NDP-α-D-sugars to generate β-linked products, and the reaction mechanism is

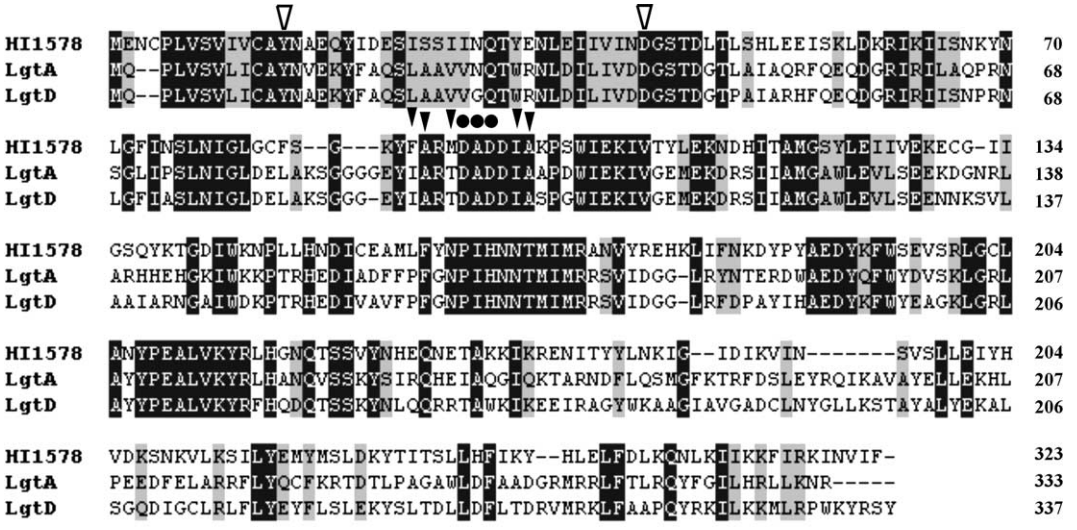

Fig. 6. Multiple amino acid sequence alignment of *H. influenzae* GalNAcT and family II homologous glycosyltransferases. The identical residues are boxed in black, whereas conserved residues are in gray boxes. Introduced gaps are shown as *hyphens*. Sequences used for the alignment are: HI1578 from *H. influenzae* (H64130), LgtA from *N. meningitides* (2204376A) and LgtD from *N. gonorrhoeae* (AAA68012). The D × D Mn^{2+} ion-binding motif is indicated by closed black circles. Flanking hydrophobic residues are indicated by black triangles. Potential uracil base binding residues Tyr^{14} and Asp^{42} are displayed with white triangles.

widely believed to be S_N2 with an oxocarbenium ion-like character in the transition state. The amino acid sequence of HI1578 shows significant similarity to several LPS synthesis related glycosyltransferases of this family including GalNAc transferase LgtD (45% identity) from *N. gonorrhoeae* and GlcNAc transferase LgtA (41% identity) from *N. meningitides*. Multiple amino acid sequence alignment (Fig. 6) reveals that they share a highly conserved N-terminal glycosyltransferase region containing the so-called D × D motif, which is thought to stabilize the Mn^{2+} ion and thus indirectly bind the diphosphate moiety of UDP. Nonpolar residues usually flank the motif and both acidic residues have been proved to be essential for activity in bovine α1,3-galactosyltransferase [34] and yeast α1,3-mannosyltransferase [35]. In *H. influenzae* HI1578, the motif is present in the form of ^{93}DAD^{95} with flanking F^{89}, A^{90}, M^{92}, I^{97}, and A^{98} to form the possible hydrophobic clusters. Recently, the X-ray structures of several glycosyltransferases have been determined including SpsA [36], a family II inverting GlcNAc transferase, which is involved in the synthesis of the spore coat of *Bacillus subtilis*. There are three sequence clusters of SpsA seemingly conserved in family II glycosyltransferases, all of which involve residues in direct contact with the nucleotide diphosphate donor [36]. Y^{11} is involved in stacking with uracil base, D^{39} makes a hydrogen bond to N3 of the base, and D^{99} sits adjacent to the distal phosphate where it coordinates the leaving group Mn^{2+} ion. All these three residues are well conserved in β-1,3-*N*-acetylgalactosaminyltransferase from *H. influenzae* (Y^{14}, D^{42}, and D^{95}, respectively). In conclusion, the amino acid sequence of HI1578 is similar to those of well-characterized NDP-sugar transferase. The majority of the sequence similarity, including residues that are essential for donor binding and catalytic function, lies in its N-terminal domain.

Globoside is the most prominent neutral glycosphingolipid in human erythrocytes and is an essential structure of blood group P antigen. It has been suggested that globoside is an adhesion molecule on epithelial cells to various bacteria such as uropathogenic *E. coli* [37] and a receptor for pig edema disease toxin [38]. Hence, the oligosaccharide structure of globoside, globotetraose, is a potential therapeutic agent that is difficult to make by traditional chemical means. Compared to chemical synthesis, enzymatic approach utilizing glycosyltransferases has been proven to be more efficient in the production of complex carbohydrates. Recently, human β3GalNAcT, a key enzyme responsible for the synthesis of globoside, has been identified and expressed in mouse fibroblast L cell [39]. However, eukaryotic host expression involves expensive tissue culture media and low yield of protein. Expression of mammalian glycosyltransferases in *E. coli*, on the other hand, normally results in the formation of inactive inclusion body. Since bacteria have been shown to synthesize oligosaccharide structures identical to those in mammals, bacterial glycosyltransferases would be better prospects for recombinant protein expression in *E. coli* [40]. We now demonstrate that β-1,3-*N*-acetylgalactosaminyltransferase from *H. influenzae* strain Rd is involved in the biosynthesis of the globotetraose-like lipopolysaccharide. High-level expression of the active enzyme in *E. coli* would facilitate large-scale enzymatic synthesis of the P blood group antigen-like oligossaccharides, thereby promoting researches related to their physiological function and potential therapeutic application.

Acknowledgments

We are grateful to Dr. Ziye Liu and Mr. Yingxin Zhang for support and helpful discussion throughout this work.

References

[1] A.A. Campagnari, M.R. Gupta, K.C. Dudas, T.F. Murphy, M.A. Apicella, Antigenic diversity of lipooligosaccharides of nontypable *Haemophilus influenzae*, Infect. Immun. 55 (1987) 882–887.

[2] F. Shan, M. Grattan, S. Germer, V. Linnsman, D. Hazlem, R. Payne, Aetiology of pneumonia in children in Goroke hospital. Papua New Guinea, Lancet ii (1984) 537–541.

[3] H. Nikaido, in: F.C. Neidhardt (Ed.), *Escherichia coli* and *Salmonella*: Cellular and Molecular Biology, vol. 1, second ed., American Society for Microbiology, Washington, DC, 1996, pp. 29–47.

[4] C.R.H. Raetz, in: F.C. Neidhardt (Ed.), *Escherichia coli* and *Salmonella*: Cellular and Molecular Biology, vol. 1, second ed., American Society for Microbiology, Washington, DC, 1996, pp. 1035–1063.

[5] A. Kimura, E.J. Hansen, Antigenic and phenotypic variations of *Haemophilus influenzae* type b lipopolysaccharede and their relationship to virulence, Infect. Immun. 51 (1986) 69–79.

[6] J.N. Weiser, D.J. Maskell, P.D. Butler, A.A. Lindberg, E.R. Moxon, Characterization of repetitive sequences controlling phase variation of *Haemophilus influenzae* lipopolysaccharide, J. Bacteriol. 172 (1990) 3304–3309.

[7] C.M. Kahler, R.W. Carlson, M.M. Rahman, L.E. Martin, D.S. Stephens, Inner core biosynthesis of lipooligosaccharide (LOS) in *Neisseria meningitidis* serogroup B: identification and role in LOS assembly of the α1,2 *N*-acetylglucosamine transferase (RfaK), J. Bacteriol. 178 (1996) 1265–1273.

[8] P. Van der Ley, M. Kramer, A. Martin, J.C. Richards, J.T. Poolman, Analysis of the *icsBA* locus required for biosynthesis of the inner core region from *Neisseria meningitidis* lipopolysaccharide, FEMS Microbiol. Lett. 146 (1997) 247–253.

[9] D. Zhou, D.S. Stephens, B.W. Gibson, J.J. Engstrom, C.F. McAllister, F.K.N. Lee, M.A. Apicella, Lipooligosaccharide biosynthesis in pathogenic *Neisseria*. Cloning, identification, and characterization of the phosphoglucomutase gene, J. Biol. Chem. 269 (1994) 11162–11169.

[10] G.O. Aspinall, M.A. Monteiro, Lipopolysaccharides of *Helicobacter pylori* strains P466 and MO19: structures of the O antigen and core oligosaccharide regions, Biochemistry 35 (1996) 2498–2504.

[11] G.O. Aspinall, M.A. Monteiro, H. Pang, E.J. Walsh, A.P. Moran, Lipopolysaccharide of *Helicobacter pylori* type strain NCTC

11637 (ATCC 43504): structure of the O-antigen chain and core oligosaccharide regions, Biochemistry 35 (1996) 2489–2497.

[12] M.A. Monteiro, K.H.N. Chan, D.A. Rasko, et al., Simultaneous expression of type 1 and type 2 Lewis blood group by *Helicobacter pylori* lipopolysaccharides, J. Biol. Chem. 273 (1998) 11533–11543.

[13] M. Gilbert, J.R. Brisson, M.F. Karwaski, J. Michniewicz, A.M. Cunningham, Y. Wu, N.M. Young, W.W. Wakarchuk, Biosynthesis of ganglioside mimics in *Campylobacter jejuni* OH4384. Identification of the glycosyltransferase genes, enzymatic synthesis of model compounds, and characterization of nanomole amounts by 600-MHz (1)H and (13)C NMR analysis, J. Biol. Chem. 275 (2000) 3896–3906.

[14] A.P. Moran, M.M. Prendergast, Molecular mimicry in *Campylobacter jejuni* and *Helicobacter pylori* lipopolysaccharides: contribution of gastrointestinal infections to autoimmunity, J. Autoimmun. 16 (2001) 241–256.

[15] J.L. Penner, G.O. Aspinall, Diversity of lipopolysaccharide structures in *Campylobacter jejuni*, J. Infect. Dis. 176 (1997) S135–S138.

[16] A.C. Wood, N.J. Oldfield, C.A. O'Dwyer, J.M. Ketley, Cloning, mutation and distribution of a putative lipopolysaccharide biosynthesis locus in *Campylobacter jejuni*, Microbiology 145 (1999) 379–388.

[17] N.J. Philips, M.A. Apicella, J.M. Griffiss, B.W. Gibson, Structural studies of the lipopolysaccharides from *Haemophilus influenzae* type b strain A2, Biochemistry 32 (1993) 2003–2012.

[18] E.K.H. Schweda, A.C. Sundstrom, L.M. Eriksson, A.A. Lindberg, Structural studies of cell envelope lipopolysaccharides from *Haemophilus ducreyi* strains ITM2665 and ITM4747, J. Biol. Chem. 16 (1994) 12040–12048.

[19] D.W. Hood, A.D. Cox, W.W. Wakarchuk, M. Schur, E.K. Schweda, S.L. Walsh, M.E. Deadman, A.A. Martin, E.R. Moxon, J.C. Richards, Genetic basis for expression of the major globotetraose-containing lipopolysaccharide from *H. influenzae* strain Rd (RM118), Glycobiology 11 (2001) 957–967.

[20] A. Risberg, H. Masoud, A. Martin, J.C. Richards, E.R. Moxon, E.K.H. Schweda, Structural analysis of the lipopolysaccharide oligosaccharide epitopes expressed by a capsule-deficient strain of *Haemophilus influenzae* Rd, Eur. J. Biochem. 261 (1999) 171–180.

[21] R.D. Fleischmann, M.D. Adams, O.R.A. White, E.F. Clayton, A.R. Kirkness, C.J. Kerlavage, J.F. Bult, B.A. Tomb, J.M. Dougherty, K. Merrick, et al., Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd, Science 269 (1995) 496–512.

[22] D.W. Hood, M.E. Deadman, T. Allen, H. Masoud, A. Martin, J.R. Brisson, R. Fleischmann, J.C. Venter, J.C. Richards, E.R. Moxon, Use of the complete genome sequence information of *Haemophilus influenzae* strain Rd to investigate lipopolysaccharide biosynthesis, Mol. Microbiol. 22 (1996) 951–965.

[23] X. Chen, J.-B. Zhang, P. Kowal, P.R. Andreana, P.G. Wang, Transferring a biosynthetic cycle into a productive *E. coli* strain: large-scale synthesis of galactosides, J. Am. Chem. Soc. 123 (2001) 8866–8867.

[24] M.M. Bradford, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, Anal. Biochem. 72 (1976) 248–254.

[25] J.D. Thompson, D.G. Higgins, T.J. Gibson, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting positions-specific gap penalties and weight matrix choice, Nucleic Acids Res. 22 (1994) 4673–4680.

[26] O. Blixt, I. van Die, T. Norberg, T.D.H. van den Eijnden, High-level expression of the *Neisseria meningitidis* lgtA gene in *Escherichia coli* and characterization of the encoded *N*-acetylglucosaminyltransferase as a useful catalyst in the synthesis of GlcNAc beta1 → 3Gal and GalNAc beta1 → 3Gal linkages, Glycobiology 9 (1999) 1061–1071.

[27] J.N. Weiser, N. Pan, Adaptation of *Haemophilus influenzae* to acquired and innate humoral immunity based on phase variation of lipopolysaccharide, Mol. Microbiol. 30 (1998) 767–775.

[28] D.W. Hood, A.D. Cox, M. Gilbert, W.W. Wakarchuk, et al., Identification of a lipopolysaccharide alpha-2,3-sialyltranfrom from *Haemophilus influenzae*, Mol. Microbiol. 39 (2001) 341–350.

[29] W.A. Nichols, B.W. Gibson, W. Melaugh, N.G. Lee, M. Sunshine, M.A. Apicella, Identification of the ADP-L-glycero-D-manno-heptose-6-epimerase (rfaD) and heptosyltransferase II (rfaF) biosynthesis genes from nontypeable *Haemophilus influenzae* 2019, Infect. Immun. 65 (1987) 1377–1386.

[30] W. Brabetz, S. Muller-Loennies, H. Brade, 3-Deoxy-D-manno-oct-2-ulosonic acid (Kdo) transferase (WaaA) and kdo kinase (KdkA) of *Haemophilus influenzae* are both required to complement a waaA knockout mutation of *Escherichia coli*, J. Biol. Chem. 275 (2000) 34954–34962.

[31] J.A. Campbell, G.J. Davies, V. Bulone, B. Henrissat, A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities, Biochem. J. 326 (1997) 929–939.

[32] J.A. Campbell, G.J. Davies, V. Bulone, B. Henrissat, A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities, Biochem. J. 329 (1998) 719.

[33] P.M. Coutinho, B. Henrissat, Carbohydrate-active enzymes server. http://afmb.cnrs-mrs.fr/~pedro/cazy/db.html 1999.

[34] L.N. Gastinel, C. Bignon, A.K. Misra, O. Hindsgaul, J.H. Shaper, D.H. Joziasse, Bovine alpha1,3-galactosyltcatalytic catalytic domain structure and its relationship with ABO histo-blood group and glycosphingolipid glycosyltransferases, EMBO J. 20 (2001) 638–649.

[35] C.A.R. Wiggins, S. Munro, Activity of the yeast MNN1 α1,3-mannosyltransferase requires a motif conserved in many other families of glycosyltransferases, Proc. Natl. Acad. Sci. USA 95 (1998) 7945–7950.

[36] S.J. Charnock, G.J. Davies, Structure of the nucleotide-diphospho-sugar transferase, SpsA from *Bacillus subtilis*, in native and nucleotide-complexed forms, Biochemistry 38 (1999) 6380–6385.

[37] B. Lund, F.P. Lindberg, M. Baga, S. Normark, Globoside specific adhesins of uropathogenic *Escherichia coli* are encoded by similar trans-complementable gene clusters, J. Bacteriol. 162 (1985) 1293–1301.

[38] G.J. Tyrrell, K. Ramotar, B. Toye, B. Boyd, C.A. Lingwood, J.L. Brunton, Alteration of the carbohydrate binding specificity of verotoxins from Gal alpha 1–4Gal to GalNAc beta 1–3Gal alpha 1–4Gal and vice versa by site-directed mutagenesis of the binding subunit, Proc. Natl. Acad. Sci. USA 89 (1992) 524–528.

[39] T. Okajima, Y. Nakamura, M. Uchikawa, D.B. Haslam, S.I. Numata, K. Furukawa, T. Urano, K. Furukawa, Expression cloning of human globoside synthase cDNAs. Identification of beta 3Gal-T3 as UDP-*N*-acetylgalactosamine: globotriaosylceramide beta 1,3-*N*-acetylgalactosaminyltransferase, J. Biol. Chem. 275 (2001) 40498–40503.

[40] W.W. Wakarchuk, A. Cunningham, D.C. Watson, N.M. Young, Role of paired basic residues in the expression of active recombinant galactosyltransferases from the bacterial pathogen *Neisseria meningitidis*, Protein Eng. 11 (1998) 295–302.