Complete chemoenzymatic synthesis of the Forssman antigen using novel glycosyltransferases identified in Campylobacter jejuni and Pasteurella multocida

Houliston, R. Scott; Bernatchez, Stephane; Karwaski, Marie-France; Mandrell, Robert E.; Jarrell, Harold C.; Wakarchuk, Warren W.; Gilbert, Michel

Publisher's version / Version de l'éditeur:
https://doi.org/10.1093/glycob/cwn117
Glycobiology, 19, 2, pp. 153-159, 2008-10-25

Glycobiology vol. 19 no. 2 pp. 153–159, 2009
doi:10.1093/glycob/cwn117
Advance Access publication on October 25, 2008

Complete chemoenzymatic synthesis of the Forssman antigen using novel glycosyltransferases identified in *Campylobacter jejuni* and *Pasteurella multocida*

R Scott Houliston[2], Stéphane Bernatchez[2], Marie-France Karwaski[2], Robert E Mandrell[3], Harold C Jarrell[2], Warren W Wakarchuk[2], and Michel Gilbert[1,2]

[2]Institute for Biological Sciences, National Research Council Canada, Ottawa, Ontario, K1A 0R6, Canada; and [3]US Department of Agriculture, Agricultural Research Service, Produce Safety and Microbiology Research Unit, Albany, CA, USA

Received on July 1, 2008; revised on October 21, 2008; accepted on October 22, 2008

We have identified an α1,4-galactosyltransferase (CgtD) and a β1,3-*N*-acetylgalactosaminyltransferase (CgtE) in the lipooligosaccharide (LOS) locus of *Campylobacter jejuni* LIO87. Strains that carry these genes may have the capability of synthesizing mimics of the P blood group antigens of the globoseries glycolipids. We have also identified an α1,3-*N*-acetylgalactosaminyltransferase (Pm1138) from *Pasteurella multocida* Pm70, which is involved in the synthesis of an LOS-bound Forssman antigen mimic and represents the only known bacterial glycosyltransferase with this specificity. The genes encoding the three enzymes were cloned and expressed in *Escherichia coli* as soluble recombinant proteins that can be used to chemoenzymatically synthesize the Forssman antigen, and its biosynthetic precursors, in high yields.

Keywords: Campylobacter jejuni/Forssman antigen/glycosyltransferase/*Pasteurella multocida*

Introduction

The Forssman glycolipid is a glycosylceramide possessing a neutral pentasaccharide head group, referred to as the Forssman antigen (Fa) and is a member of the globoseries glycolipid family. The Fa has been identified in a number of mammals and exhibits heterogeneity with respect to developmental and cell-type expression among species. Some studies have reported the presence of the Fa in certain human embryonic and tumor cells (Yokota et al. 1981; Ono et al. 1994). However, the human Forssman synthetase gene has been shown to encode an inactive enzyme (Xu et al. 1999; Elliott et al. 2003) which suggests that humans are not capable of synthesizing the Forssman glycolipid. Immunological methods used to detect the presence of the Fa may not effectively discriminate between authentic Fa and similar glycoconjugates.

The P^k and P blood group antigens, known as globotriose (Gb_3) and globotetraose (Gb_4), respectively, are biosynthetic precursors of the Fa (Figure 1). The glycosphingolipid Gb_4 is synthesized from Gb_3 through the addition of an *N*-acetylgalactosaminyl (GalNAc) residue catalyzed by a β1,3-GalNAc-transferase. The Forssman glycolipid is synthesized from Gb_4 by adding a terminal αGalNAc residue, catalyzed by an α1,3-GalNAc-transferase. The gene encoding the latter enzyme, referred to as the Forssman synthetase, was first identified in canine kidney cells (Haslam and Baenziger 1996). Its amino acid (aa) sequence is homologous with α1,3-Gal/GalNAc-transferases that are involved in the synthesis of related glycans, such as the AB histo-blood group antigens (Yamamoto et al. 1990; Keusch et al. 2000).

Bacteria often synthesize glycans found at the surface of host cells as a method to evade immune targeting. For instance, *Campylobacter jejuni* has been shown to synthesize ganglioside mimics as part of its lipooligosaccharide (LOS) (Yuki et al. 1993); *Haemophilus influenzae* can incorporate Gb_4 units in its LOS (Risberg et al. 1999); and *Neisseria meningitidis* and *Escherichia coli* synthesize capsular polysaccharides similar to polysialic acid found in mammals (Finne et al. 1983; Troy et al. 1975). Several genes encoding glycosyltransferases involved in the synthesis of these molecular mimics have been cloned and expressed in *E. coli*, and the enzymes have been successfully used to synthesize these glycans (Antoine et al. 2005; Blixt et al. 2005; Willis et al. 2008). This in turn has provided investigators in both academia and industry with readily available oligosaccharides that have been traditionally difficult to synthesize and/or purify. It is now generally recognized that bacterial glycosyltransferases are optimally suited for chemoenzymatic glycan synthesis because of the ease with which their genes can be cloned and expressed and their improved solubility over eukaryotic glycosyltransferases.

We have identified an α1,4-Gal-transferase and a β1,3-GalNAc-transferase from *C. jejuni* LIO87, and an α1,3-GalNAc-transferase from *Pasteurella multocida* Pm70. The three enzymes were produced as soluble recombinant proteins and used sequentially to synthesize the Fa in multimilligram quantities starting from *p*-nitrophenyl lactose.

Results

The LOS locus of C. jejuni LIO87 contains α1,4-Gal-transferase and β1,3-GalNAc-transferase genes

The core glycan structure from *C. jejuni* is known to be highly variable primarily due to significant genetic heterogeneity within its LOS biosynthesis locus. The loci from over 70 strains have been sequenced to date and grouped into 19 classes based on their component genes (Gilbert et al. 2008). Strains belonging to classes "A," "B," and "C" have received considerable attention because within their loci are enzymes that direct the

[1]To whom correspondence should be addressed: Tel: +1-613-991-9956; Fax: +1-613-952-9092; e-mail: michel.gilbert@nrc-cnrc.gc.ca

synthesis of LOS-bound ganglioside mimics, which have been implicated in triggering autoimmune disease (Godschalk et al. 2004). Strain LIO87, which belongs to class "D," does not carry sialyltransferase or sialic acid biosynthesis genes in its LOS locus. Five genes among its 10 open reading frames (ORFs) are likely involved in the synthesis of the inner core, while the remaining five are putative outer core glycosyltransferases of unknown specificity (GenBank accession number AF400669).

We cloned the five putative class "D" outer core glycosyltransferases into *E. coli* and performed enzymatic activity screens. This entails adding different combinations of activated sugar donors (UDP-Glc, UDP-Gal, UDP-GlcNAc, and UDP-GalNAc) with fluorescently tagged acceptors (GalNAcα-, Glcβ-, GlcNAcβ-, Lac-, LacNAc-, and Gb_3-FCHASE) to lysates of transformed cells expressing the heterologous protein. Capillary-electrophoresis (CE) is then used to detect transferase activity. Two genes were found to be active in our screen. CgtD (*C. jejuni* glycosyltransferase D, protein sequence AAM90647) was found to transfer a galactosyl residue (from UDP-Gal) to both lactose (Lac) and *N*-acetyl-lactosamine (LacNAc) acceptors. A β1,3-*N*-acetylgalactosaminyltransferase (CgtE; protein sequence AAM90646) was found to transfer GalNAc (from UDP-GalNAc) to an αGal(1–4)-Lac substrate (i.e., Gb_3). This indicates that in vivo the product of the CgtD enzymatic addition serves as the target for CgtE (Figure 1).

In order to scale up their production, both genes were cloned into a maltose-binding protein (MalE) expression vector (pCWori+), as C-terminal fusion constructs. Only CgtD gave rise to strong activity as a MalE–CgtD fusion (200–300 units per liter) and an apparent inducible band on sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) (Figure 2). The analogous MalE–CgtE fusion did not express well (10–15 units per liter), indicating that the fusion with MalE appears to inhibit CgtE. CgtE expressed better without a fusion partner with a production of 100–150 units per liter of culture. In addition, we did not observe an IPTG inducible band on SDS–PAGE for either CgtE or MalE–CgtE. It is unclear why recombinant CgtE does not accumulate in the *E. coli* cytoplasm. It is possible that it is sensitive to proteolytic degradation and cannot accumulate to give a visible band by SDS–PAGE. Both CgtD and CgtE were found to require divalent cations, with the presence of Mn^{2+} yielding higher transferase levels than with Mg^{2+} (Table I). Using *p*-nitrophenyl-Lac (*p*NP-Lac) as an acceptor, CgtD was found to synthesize the Gb_3 antigen, while Gb_4 is produced by CgtE with *p*NP-Gb_3 as its substrate. The structures of both derivatives were confirmed by NMR spectroscopy (Figure 3 and Table II), establishing that CgtD is an α1,4-Gal-transferase and CgtE a β1,3-GalNAc-transferase. *p*NP-LacNAc was also found to be a viable acceptor for CgtD, to produce *p*NP-αGal(1–4)LacNAc (Table I); this was also confirmed by NMR spectroscopy (supplementary Table I).

Surprisingly, the aa sequence of CgtD does not show significant homology with any characterized or putative glycosyltransferases, not even with LgtC, an α1,4-Gal-transferase grouped into GT8 of the CAZy database (Coutinho et al. 2003) that is found in *H. influenzae*, *N. meningitidis*, and *N. gonorrhoeae*. It shares significant aa sequence identity (57–97%) with hypothetical proteins in the LOS locus of other *C. jejuni* strains. CgtE is homologous (47–48% identity) with CgtB, a β1,3-Gal-transferase present in several *C. jejuni* strains (Bernatchez et al. 2007). However, the CgtE sequence shares very low sequence

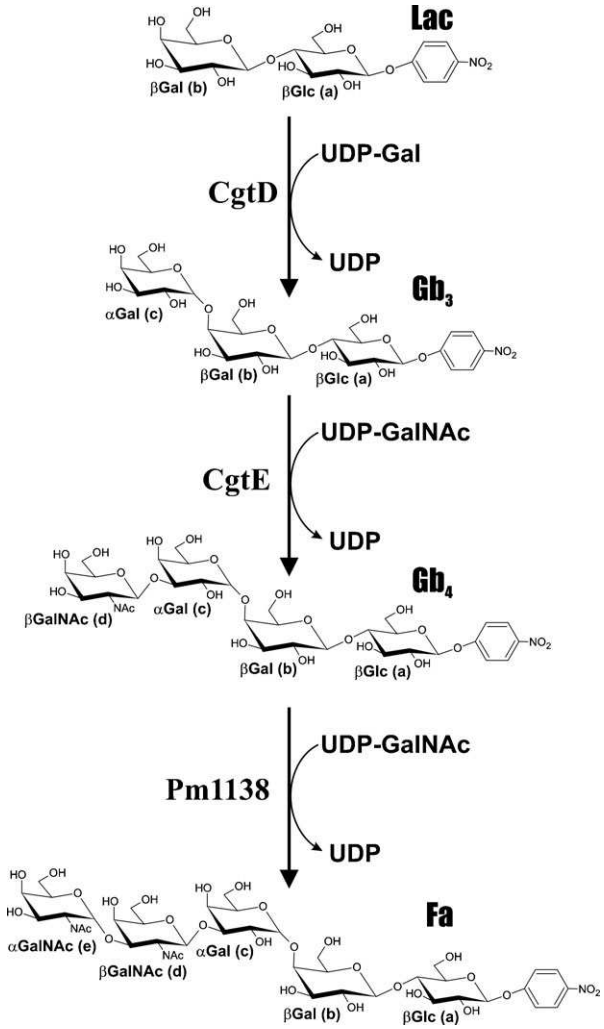

Fig. 1. Three-step synthesis of the Forssman antigen. *p*NP-Lac serves as the acceptor for CgtD, an α1,4-Gal-transferase, to synthesize *p*NP-Gb_3. *p*NP-Gb_4 is synthesized using the β1,3-GalNAc-transferase CgtE. Finally, the *p*NP derivative of the Fa is synthesized using Pm1138, an α1,3-GalNAc-transferase.

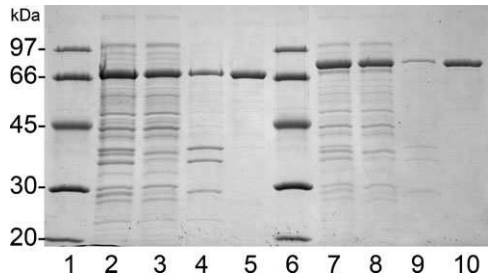

Fig. 2. SDS–PAGE (10%) analysis of purified MalE–CgtD (construct CJL-99) and MalE–Pm1138 (construct PML-01). Lanes 1 and 6: molecular mass markers (kDa); lane 2: total extract of *E. coli* AD202/CJL-99; lane 3: supernatant (26,892 × *g*) of *E. coli* AD202/CJL-99; lane 4: pellet (26,892 × *g*) of *E. coli* AD202/CJL-99; lane 5: purified MalE–CgtD (2 μg); lane 7: total extract of *E. coli* AD202/PML-01; lane 8: supernatant (26,892 × *g*) of *E. coli* AD202/PML-01; lane 9: pellet (26,892 × *g*) of *E. coli* AD202/PML-01; lane 10: purified MalE–Pm1138 (2 μg).

Table I. Characteristics of the three enzymes used to synthesize the Forssman antigen from lactosyl derivatives

Enzyme	Activity/Host organism	Functional characteristics	Sequence and homology[a]
CgtD	α1,4-galactosyltransferase cloned from *C.jejuni* LIO87, a strain possessing a type "D" LOS locus[b]	Donor: UDP-Gal Activity measurements[c]: Lac-FCHASE: 860 mU/mg LacNAc-FCHASE: 1,212 mU/mg Syntheses: *p*NP-Gb_3 (41.2 mg) *p*NP-αGal(1–4)-LacNAc (19.2 mg) pH optimum: 7–8 Metal requirement: Mn^{2+} or Mg^{2+} (15% less active with Mg^{2+})	Protein sequence: AAM90647 No apparent sequence homology with other glycosyltransferases
CgtE	β1,3-*N*-acetylgalactosaminyltransferase cloned from C. *jejuni* LIO87, a strain possessing a type "D" LOS locus[d]	Donor[c,d]: UDP-GalNAc: 5,540 mU/mL UDP-Gal: 1,950 mU/mL Activity measurements[c,e]: Gb_3-FCHASE: 5,540 mU/mL αGal-FCHASE: 610 mU/mL βGal-FCHASE: 52 mU/mL Syntheses: *p*NP-Gb_4 (31.6 mg) *p*NP-βGalNAc(1–3)αGal(1–4)LacNAc (6.9 mg) pH optimum: 7– 7.5 Metal requirement: Mn^{2+} or Mg^{2+} (50% less active with Mg^{2+})	Protein sequence: AAM90646 GT2 CAZy member
Pm1138	α1,3-*N*-acetylgalactosaminyltransferase cloned from P. *multocida* Pm70	Donor: UDP-GalNAc Activity measurements: βGalNAc-FCHASE: 186 mU/mg αGalNAc-FCHASE: 2 mU/mg Syntheses: αGalNAc(1,3)βGalNAc-FCHASE (1 mg) *p*NP-Fa (13.9 mg) pH optimum: 8 Metal requirement: Mn^{2+} or Mg^{2+}	Protein sequence: NP_246075 GT4 CAZy member

[a]CAZy database classification described in Coutinho et al. (2003).
[b]From Gilbert et al. (2008).
[c]Activity measurements: in mU/mg of pure protein for CgtD and Pm1138 and in mU/mL of extract for CgtE. Donor concentration: 1 mM. Acceptor concentration: 0.5 mM.
[d]The donor specificity of CgtE was probed using Gb_3-FCHASE as an acceptor.
[e]The acceptor specificity of CgtE was probed using UDP-GalNAc as a donor.

identity with glycosyltransferases possessing similar activities previously identified in other bacteria; it possesses only 20% identity with the β1,3-GalNAc-transferase LgtD found in *H. influenzae* and *N. gonorrhoeae*, both of which have been assigned to family GT2 in the CAZy database. Clearly, activity assays were essential to assign the functionality of both CgtD and CgtE.

P. multocida Pm70 carries an α1,3-GalNAc-transferase associated with LOS biosynthesis

The glycan component of the LOS from *P. multocida* Pm70 was found to incorporate the complete Fa at its nonreducing end (St. Michael et al. 2005). This observation correlates with the fact that this pathogen is frequently isolated from mammals that are known to express the Forssman glycolipid. Several putative LOS-associated glycosyltransferases are closely clustered within the *P. multocida* Pm70 genome. One of those, designated Pm1138, was hypothesized to be an α1,3-GalNAc-transferase that adds the terminal α-GalNAc residue to the LOS outer core (St. Michael et al. 2005). Pm1138 shares aa sequence similarity with several putative Gal-transferases including an α1,3-Gal-transferase (GenBank accession number AF143904) from *Actinobacillus pleuropneumoniae* (Ramjeet et al. 2005). Pm1138 also shows weak homology with PglH and PglJ, two α1,4-GalNAc-transferases involved in *C. jejuni* protein glycosylation (Glover et al. 2005). Pm1138 belongs to CAZy family GT4 and does not possess sequence similarity with mammalian Forssman synthetases which belong to CAZy family GT6. Clearly, sequence homology is not sufficient to confirm the proposed specificity of Pm1138.

The gene encoding Pm1138 was cloned into a maltose-binding protein (MalE) expression vector (pCWori+) and expressed as a C-terminal fusion construct in *E. coli* (Figure 2). Subsequent activity screening indicated that it possesses transferase activity in the presence of an UDP-GalNAc donor and a βGalNAc-FCHASE substrate. Pm1138 has very low activity

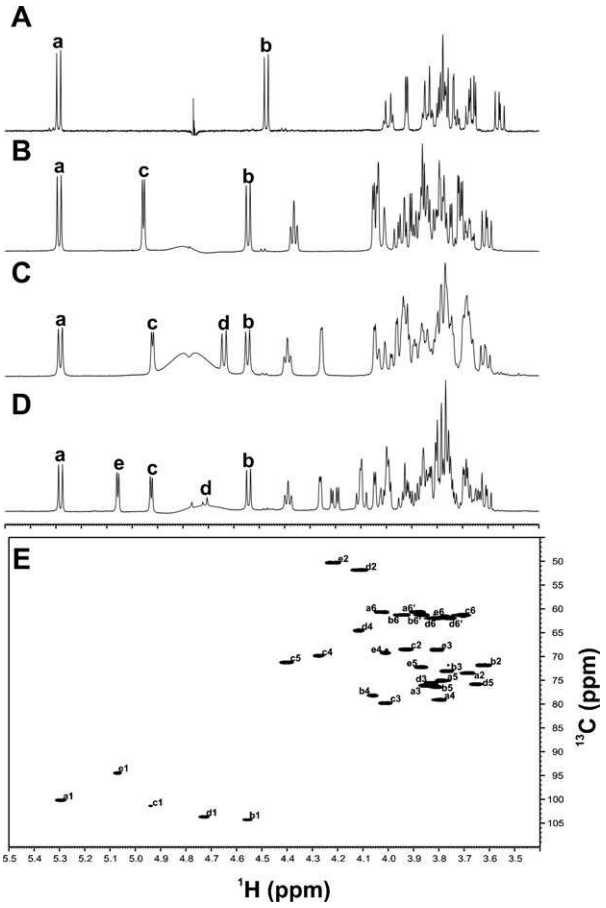

Fig. 3. NMR spectra of the Forssman antigen and its synthetic precursors. Displayed is an overlay of 1D-^{1}H spectra in the region between 3.5 and 5.5 ppm for *p*NP derivatives of Lac (**A**), Gb_3 (**B**), Gb_4 (**C**), and the Fa (**D**). Anomeric protons for constituent monosaccharides in the glycan derivatives are labeled alphabetically according to the scheme in Figure 1. (**E**) Presented is an ^{1}H–^{13}C HSQC spectrum of *p*NP-Fa with each resonance labeled alphanumerically based on the structure presented in Figure 1.

on αGalNAc-FCHASE (Table I). The purified MalE–Pm1138 fusion was used to synthesize preparative quantities of the GalNAc-βGalNAc-FCHASE product. The reaction product was analyzed by mass spectrometry; the masses of the singly charged (985.5 *m/z*) and doubly charged (492.3 *m/z*) molecular ions are consistent with the expected molecular weight of a GalNAc-GalNAc-FCHASE molecule (data not shown). Analysis by NMR spectroscopy confirmed that the product of Pm1138 was αGalNAc(1–3)βGalNAc-FCHASE (supplementary Table II), which confirms that Pm1138 is an α1,3-GalNAc-transferase. It would appear that its role in vivo is to add the terminal αGalNAc residue to complete the LOS-bound Fa mimic produced by *P. multocida* Pm70, and currently it is the only known bacterial enzyme possessing this specificity.

Chemoenzymatic synthesis of the Fa pentasaccharide

The availability of recombinant and soluble forms of an α1,4-Gal-transferase, a β1,3-GalNAc-transferase, and an α1,3-GalNAc-transferase provides the means to synthesize the complete Fa glycan starting with a lactosyl derivative. A one-pot synthesis was not attempted because the β1,3-GalNAc-transferase (CgtE) can transfer both GalNAc and Gal which would result in a mixed product, although the activity with UDP-GalNAc is higher than with UDP-Gal (Table I). Starting with 107 μmol (50 mg) of *p*NP-Lac, we synthesized and isolated 66 μmol (61.7% yield) of *p*NP-Gb_3. Analysis of the reaction by thin layer chromatography (TLC) suggested that the conversion of *p*NP-Lac to *p*NP-Gb_3 went to completion, but some material was lost during purification on the C18 SepPak column. Starting with 40 μmol (25 mg) of *p*NP-Gb_3, we synthesized and isolated 38 μmol (95% yield) of *p*NP-Gb_4. In the final step of the synthesis, we synthesized and isolated 13.5 μmol of *p*NP-Fa from 15 μmol (12.4 mg) of *p*NP-Gb_4 (90% yield). Both of the latter two reactions resulted in complete conversion of the substrates as judged by TLC analysis, which was consistent with the excellent recovery yield for the final two steps of the synthesis. The structure of the final product (*p*NP-Fa) was confirmed by NMR spectroscopy (Figure 3, Table II).

Discussion

The Fa is expressed in a variety of animals in a species-specific pattern, usually as a complex carbohydrate associated with cell-surface antigens. Gb_4, the precursor of the Fa, serves as an attachment site for bacteria, viruses, and toxins in humans and other species. The conversion of a Gb_4 precursor into the Fa will have an impact on the adherence of pathogenic organisms, directly affecting microbial ecology and modifying host susceptibility to infectious diseases in some species (Elliot et al. 2003). Understanding binding interactions of the Fa with proteins such as toxins, antibodies, and selectins would be facilitated by the availability of large amounts of soluble Fa derivatives. Although the Fa pentasaccharide has been chemically synthesized (Nilsson et al. 1994), we believe that the glycosyltransferases described in this work could be usefully applied to enable rapid and efficient synthesis of this cell-surface antigen for future investigations.

A recombinant α1,4-Gal-transferase (LgtC) from *N. meningitidis* has been used previously in high-yield syntheses of the Gb_3 trisaccharide (Zhang et al. 2002; Antoine et al. 2005) and of the P1 trisaccharide (αGal(1–4)βGal(1–4)βGlcNAc) (Liu et al. 2003). Zhang et al. (2002) obtained the production of 300 units per liter of culture for the recombinant LgtC α1,4-Gal-transferase, while we report the production of 200–300 units per liter for the recombinant CgtD. It is difficult, however, to compare directly these production levels because the assay conditions and acceptor labeling groups were different. Both enzymes achieved quantitative conversion of lactosyl derivatives to the corresponding products and are potentially useful in large-scale syntheses of Gb_3 derivatives.

Shao et al. (2002) and Antoine et al. (2005) have used a recombinant β1,3-GalNAc-transferase (LgtD) from *H. influenzae* in high-yield syntheses of the Gb_4 tetrasaccharide. The production of LgtD in units per liter was not reported in these two publications and it is difficult to determine if the recombinant CgtE represents an improvement as a source of β1,3-GalNAc-transferase activity. The recombinant LgtD was produced as a highly soluble form at a high level by Shao et al. (2002). An IPTG inducible band was clearly visible by SDS–PAGE and a purification protocol has been developed for LgtD. Considering that we did not observe an inducible band with CgtE and

Table II. ^{1}H and ^{13}C assignments for the Forssman antigen and its synthetic precursors. All reported values are based on spectra acquired at 25°C in 100% D_2O. The residues are labeled based on the synthesis scheme presented in Figure 1

		*p*NP-Lac		*p*NP-Gb_3		*p*NP-Gb_4		*p*NP-Fa	
Residue	Atom (type)	δ_H (ppm)	δ_C (ppm)	δ_H (ppm)	δ_C (ppm)	δ_H (ppm)	δ_C (ppm)	δ_H (ppm)	δ_C (ppm)
βGlc (a)	1 (CH)	5.29	100.2	5.29	100.2	5.28	100.2	5.29	100.2
	2 (CH)	3.66	73.4	3.67	73.4	3.67	73.5	3.68	73.5
	3 (CH)	3.83	76.1	3.84	76.1	3.84	76.1	3.85	76.1
	4 (CH)	3.77	78.8	3.78	79.1	3.78	79.1	3.79	79.0
	5 (CH)	3.78	75.0	3.78	75.1	3.78	75.1	3.79	75.0
	6 (CH_2)	3.84, 4.00	60.7	3.87, 4.02	60.7	3.87, 4.02	60.7	3.87, 4.02	60.6
βGal (b)	1 (CH)	4.47	103.9	4.54	104.3	4.54	104.3	4.55	104.2
	2 (CH)	3.55	71.9	3.60	71.9	3.61	71.9	3.62	71.8
	3 (CH)	3.66	73.4	3.76	73.1	3.76	73.1	3.77	73.0
	4 (CH)	3.92	69.5	4.05	78.3	4.05	78.2	4.06	78.2
	5 (CH)	3.72	76.4	3.80	76.4	3.80	76.4	3.81	76.4
	6 (CH_2)	3.77	62.0	3.86, 3.95	61.3	3.86, 3.94	61.3	3.86, 3.94	61.3
αGal (c)	1 (CH)			4.96	101.2	4.92	101.4	4.94	101.3
	2 (CH)			3.85	69.5	3.90	68.6	3.92	68.5
	3 (CH)			3.91	70.1	3.97	79.7	4.01	79.7
	4 (CH)			4.04	69.9	4.26	69.9	4.27	69.8
	5 (CH)			4.36	71.8	4.31	71.2	4.40	71.2
	6 (CH_2)			3.71	61.4	3.69	61.3	3.70	61.3
βGalNAc (d)	1 (CH)					4.64	104.3	4.73	103.7
	2 (CH)					3.94	53.6	4.11	51.8
	3 (CH)					3.76	73.1	3.83	75.6
	4 (CH)					3.93	68.7	4.11	64.5
	5 (CH)					3.67	75.9	3.65	75.8
	6 (CH_2)					3.76, 3.79	62.0	3.76, 3.81	62.0
	NAc (CH_3)					2.04	23.2	2.04	22.9
αGalNac (e)	1 (CH)							5.07	94.5
	2 (CH)							4.22	50.3
	3 (CH)							3.81	68.6
	4 (CH)							4.00	69.2
	5 (CH)							3.87	72.2
	6 (CH_2)							3.77	61.8
	NAc (CH_3)							2.07	23.1

could not develop a purification protocol, it is likely that LgtD has more potential as a reagent for large-scale syntheses of Gb_4 analogs.

The complete structure of the LOS core glycan of *C. jejuni* LIO87 has not been determined, although we know that it does not contain a Gb_3 or Gb_4 mimic. Mass spectrometry analysis has shown that it possesses a truncated LOS outer core with only a HexNAc residue attached to the inner core (supplementary Table III). The truncated LOS of *C. jejuni* LIO87 indicates that the activities of CgtD and CgtE are limited, probably by an inactive and yet unidentified glycosyltransferase that prevents the production of a glycan template that can serve as a substrate for further elongation by CgtD and CgtE. In contrast, the LOS core glycan of *C. jejuni* RM1221, which has identical copies of *cgtD* and *cgtE* (NCBI protein sequences AAW35603 and AAW35602, respectively), has been elucidated recently and found to contain a Gb_4 mimic (Mandrell et al. 2007). Both *cgtD* and *cgtE* from *C. jejuni* RM1221 were found to contain homopolymeric G-tracts, which lead to the phase-variable expression of these genes. As a consequence, the LOS outer core structure in *C. jejuni* RM1221 possesses a heterogeneous mixture of glycans including P (Gb_4), P1, and possibly P^k (Gb_3) antigen mimics (Mandrell et al. 2007). Strains carrying *cgtD* and *cgtE*, therefore, have the potential to synthesize LOS-bound P blood group antigen mimics, which represents another glycan antigen mimicry strategy to evade the immune response, similar to the well-described ganglioside mimicry (Aspinall et al. 1994; Moran et al. 1996).

Material and methods

Bacterial cells

The genes encoding CgtD (α1,4-Gal-transferase) and CgtE (β1,3-GalNAc-transferase) were cloned from *C. jejuni* strain LIO87. Cells were grown on Mueller–Hinton agar (Becton Dickinson, Sparks, MD) under microaerophilic conditions at 37°C. Pm1138 (α1,3-GalNAc-transferase) was cloned from *P. multocida* strain Pm70 which was grown on chocolate agar at 37°C.

Gene cloning and transformation

Genomic DNA from both *C. jejuni* and *P. multocida* was obtained using a DNeasy Tissue Kit (QIAGEN, Mississauga, Canada). The *cgtE* gene was amplified from *C. jejuni* LIO87 genomic DNA using *Pwo* polymerase (Roche Applied Science, Laval, Canada) and the following two primers containing *Nde*I and *Sal*I restriction sites (*italicized*), respectively: CJ-640 (5′-TTTAAGAAAA*CATATG* CGTAAAA TTTCAATCATC-3′ 34-mer) and CJ-641 (5′-GGTAATCTA *GTCGAC*AATTATAACACATTC-3′). Similarly, Pm1138 was amplified from *P. multocida* Pm70 genomic DNA using the

primers PML-1 (5′-GCTACTCTT*CATATG*AATATTCTATTTGTACATAAAAGCCTTG-3′ 44-mer) and PML-2 (5′-CTT AGC*GTCGAC*TTAACTATTGAATTTTTG TAAATGAGA-3′ 40-mer). The gene encoding CgtD in *C. jejuni* LIO87 was amplified in two stages in order to introduce silent mutations to an heterogeneous homopolymeric G-tract and to obtain an in-frame ORF. Two separate PCR reactions were performed to generate two overlapping gene fragments containing the silent mutations due to either the 5′ or the 3′ primers (underlined bases in primers CJ-637 and CJ-638). Two PCR products were obtained using either primers CJ-636 (5′-TAAAAGGCTA*CATATG* ACTGAAATTTCAAGTTTTTGG-3′ 37-mer) and CJ-637 (5′-CCCATACGCCTCCTCTGAGATA AAGTAG-3′) or primers CJ-638 (5′-CTACTTTATCTCAGAG GAGGCGTATGGG-3′) and CJ-639 (5′-GGCAAGATGATT *GTCGAC*TTAGGCATTGTTTTTC-3′ 34-mer). The two PCR products were then used with primers CJ-636 and CJ-639, containing *Nde*I and *Sal*I restriction sites, to amplify the full-length *cgtD* gene containing an in-frame homogeneous ORF. The PCR products were digested with *Nde*I and *Sal*I and cloned into pCWori+(−*lacZ*) and in a version of pCWori+(−*lacZ*) containing the sequence encoding the *E. coli* maltose-binding protein (without the leader peptide) and the thrombin cleavage site. The resulting plasmids were electroporated into *E. coli* AD202.

Enzyme expression and purification

The strains carrying the constructs for the expression of the glycosyltransferases were grown in 2 YT medium containing 150 μg/mL ampicillin and 2 g/L glucose. The cultures were incubated at 37°C until $A_{600} = 0.35$, induced with 1 mM isopropyl-1-thio-β-D-galactopyranoside, and then incubated overnight at 20°C. The cells were broken using an Avestin C5 Emulsiflex cell disruptor (Avestin, Ottawa, Canada). The fusion proteins MalE–CgtD and MalE–Pm1138 were purified by affinity chromatography on amylose resin following the manufacturer's instructions (New England Biolabs, Beverly, MA). Since CgtE could not be produced with a purification tag, we centrifuged the cell extract at 27,000 × *g* and used the supernatant directly in syntheses of Gb_4 derivatives.

Enzyme assays

FCHASE-labeled oligosaccharides (Wakarchuk et al. 1996) were incubated with cell lysates prepared by sonication or with recombinant enzymes purified as described above. Enzyme screening assays were performed at 37°C for 5–30 min using a 1 mM donor and a 0.5 mM acceptor. All reactions were stopped with 10 μL of 50% acetonitrile, 10 mM EDTA, and 1% SDS and diluted with H_2O to obtain 10–15 μM of the FCHASE-labeled compounds. The samples were analyzed by CE as described previously (Wakarchuk and Cunningham 2003) except that a P/ACE MDQ CE System equipped with a Laser module 488 (Beckman Coulter, Fullerton, CA) was used. Quantification of activity was performed by integration of the trace peaks using the MDQ 32 Karat software. Transferase activity to acceptor glycans was also verified by CE-mass spectrometry using previously described methods (Li et al. 2004).

Chemoenzymatic synthesis of pNP-Gb_3, pNP-Gb_4, and pNP-Fa

For the synthesis of *p*NP-Gb_3, 50 mg (107 μmol) of *p*NP-Lac was dissolved in 10 mL of H_2O. The reaction was started in a volume of 14.7 mL containing 50 mM Hepes, pH 7.5, 10 mM $MnCl_2$, 195 μmol of UDP-Gal, and five units of MalE-CgtD (α1,4-Gal-transferase). The reaction was >95% complete after 2 h of incubation at 37°C, as judged by TLC analysis. The reaction was supplemented with 0.85 unit of MalE–CgtD and 20 μmol of UDP-Gal, and incubated for another 1 h at 37°C. The conversion of *p*NP-Lac to *p*NP-Gb_3 was complete after a total of 3 h of incubation at 37°C. The reaction was diluted with 120 mL of H_2O and loaded on a C18 SepPak column (5 g) equilibrated with H_2O. Hydrophilic material was washed off with water and the product was eluted with 100% methanol. Appropriate fractions were collected and evaporated to give 41.2 mg (66 μmol) of *p*NP-Gb_3.

For the synthesis of *p*NP-Gb_4, 25 mg (40 μmol) of *p*NP-Gb_3 was dissolved in 0.8 mL of H_2O. The reaction was performed in an 8 mL volume containing 50 mM Hepes, pH 7.5, 10 mM $MnCl_2$, 80 μmol of UDP-GlcNAc, 15 units of UDP-GlcNAc 4-epimerase (Bernatchez et al. 2005), and 1.7 units of CgtE (β1,3-GalNAc-transferase). The reaction was complete after 2 h of incubation at 37°C, as judged by TLC analysis. The reaction was diluted with 100 mL of H_2O and purified on a C18 SepPak column (5 g) as described above, except that the product was eluted with 60% methanol. Appropriate fractions were collected and evaporated to give 31.6 mg (38 μmol) of *p*NP-Gb_4.

For the synthesis of *p*NP-Fa, 12.4 mg (15 μmol) of *p*NP-Gb_4 was dissolved in 0.3 mL of H_2O. The reaction was performed in 3 mL containing 50 mM Hepes, pH 8, 10 mM $MgCl_2$, 30 μmol of UDP-GlcNAc, 5.3 units of UDP-GlcNAc 4-epimerase (Bernatchez et al. 2005), and 0.12 units of MalE-Pm1138 (α1,3-GalNAc-transferase). The reaction was complete after 60 min of incubation at 37°C, as judged by TLC analysis. The reaction was diluted with 35 mL of H_2O and loaded on two C18 SepPak columns (1 g) equilibrated with H_2O. Hydrophilic material was washed off with water and the product was eluted with 60% methanol. Appropriate fractions were collected and evaporated to give 13.5 μmol of *p*NP-Fa.

Structural characterization of the glycan derivatives

NMR spectroscopy was used to confirm the composition and structure of CgtD, CgtE, and Pm1138 enzymatic products, which included αGalNAc(1–3)βGalNAc-FCHASE, *p*NP-αGal(1–4)LacNAc, *p*NP-βGal(1–3)αGal(1–4)LacNAc as well as *p*NP-Gb_3, *p*NP-Gb_4, and *p*NP-Fa. The spectra of the compounds were acquired in 100% D_2O at 25°C at concentrations of 3–6 mM on Varian spectrometers operating at 500 and 600 MHz (Varian, Palo Alto, CA). Standard ^1H–^1H COSY, TOCSY and NOESY spectra and ^1H-^{13}C HSQC and HMBC spectra were acquired as previously described (Gilbert et al. 2000), to make resonance assignments and to verify the linkage position between component monosaccharides. Chemical shifts were referenced with respect to the methyl group of an internal acetone standard appearing at 2.23 and 31.1 ppm for ^1H and ^{13}C, respectively.

Supplementary data

Supplementary data for this article is available online at http://glycob.oxfordjournals.org/.

Funding

The Human Frontier Science Program (RGP 38/2003) and NEOSE Technologies, Horsham, PA, USA.

Acknowledgements

We thank Dr. Jianjun Li for the mass spectrometry analysis of αGalNAc(1–3)βGalNAc-FCHASE.

Conflict of interest statement

None declared.

Abbreviations

Aa, amino acid; CE, capillary electrophoresis; Fa, Forssman antigen; FCHASE, 6-(fluorescein-5-carboxamido)-hexanoic acid; Gal, galactose; GalNAc, *N*-acetylgalactosamine; Gb_3, globotriose; Gb_4, globotetraose; Lac, lactose; LOS, lipooligosaccharide; ORFs, open reading frames; *p*NP, *p*-nitrophenol; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; TLC, thin layer chromatography.

References

Antoine T, Bosso C, Heyraud A, Samain E. 2005. Large scale in vivo synthesis of globotriose and globotetraose by high cell density culture of metabolically engineered *Escherichia coli*. *Biochimie*. 87:197–203.

Aspinall GO, Fujimoto S, McDonald AG, Pang H, Kurjanczyk LA, Penner JL. 1994. Lipopolysaccharides from *Campylobacter jejuni* associated with Guillain–Barré syndrome patients mimic human gangliosides in structure. *Infect Immun.* 62:2122–2125.

Bernatchez S, Gilbert M, Blanchard MC, Karwaski MF, Li J, Defrees S, Wakarchuk WW. 2007. Variants of the β1,3-galactosyltransferase CgtB from the bacterium *Campylobacter jejuni* have distinct acceptor specificities. *Glycobiology.* 17:1333–1343.

Bernatchez S, Szymanski CM, Ishiyama N, Li J, Jarrell HC, Lau PC, Berghuis AM, Young NM, Wakarchuk WW. 2005. A single bifunctional UDP-GlcNAc/Glc 4-epimerase supports the synthesis of three cell surface glycoconjugates in *Campylobacter jejuni*. *J Biol Chem.* 280:4792–4802.

Blixt O, Vasiliu D, Allin K, Jacobsen N, Warnock D, Razi N, Paulson JC, Bernatchez S, Gilbert M, Wakarchuk W. 2005. Chemoenzymatic synthesis of 2-azidoethyl-ganglio-oligosaccharides GD3, GT3, GM2, GD2, GT2, GM1, and GD1a. *Carbohydr Res.* 340:1963–1972.

Coutinho PM, Deleury E, Davies GJ, Henrissat B. 2003. An evolving hierarchical family classification for glycosyltransferases. *J Mol Biol.* 328:307–317.

Elliott SP, Yu M, Xu H, Haslam DB. 2003. Forssman synthetase expression results in diminished Shiga toxin susceptibility: A role for glycolipids in determining host–microbe interactions. *Infect Immun.* 71:6543–6552.

Finne J, Leinonen M, Makela PH. 1983. Antigenic similarities between brain components and bacteria causing meningitis. Implications for vaccine development and pathogenesis. *Lancet.* 322:355–357.

Gilbert M, Brisson JR, Karwaski MF, Michniewicz J, Cunningham AM, Wu Y, Young NM, Wakarchuk WW. 2000. 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.* 275:3896–3906.

Gilbert M, Parker CT, Moran AP. 2008. *Campylobacter jejuni* lipooligosaccharides: Structures and biosynthesis. In: *Campylobacter*. 3rd ed. Washington (DC): ASMPress. p. 483–504.

Glover KJ, Weerapana E, Imperiali B. 2005. In vitro assembly of the undecaprenylpyrophosphate-linked heptasaccharide for prokaryotic *N*-linked glycosylation. *Proc Natl Acad Sci USA.* 102:14255–14259.

Godschalk PC, Heikema AP, Gilbert M, Komagamine T, Ang CW, Glerum J, Brochu D, Li J, Yuki N, Jacobs BC, et al. 2004. The crucial role of *Campylobacter jejuni* genes in anti-ganglioside antibody induction in Guillain–Barré syndrome. *J Clin Invest.* 114:1659–1665.

Haslam DB, Baenziger JU. 1996. Expression cloning of Forssman glycolipid synthetase: A novel member of the histo-blood group ABO gene family. *Proc Natl Acad Sci USA.* 93:10697–10702.

Keusch JJ, Manzella SM, Nyame KA, Cummings RD, Baenziger JU. 2000. Expression cloning of a new member of the ABO blood group glycosyltransferases, iGb3 synthase, that directs the synthesis of isoglobo-glycosphingolipids. *J Biol Chem.* 275:25308–25314.

Li J, Cox AD, Hood D, Moxon ER, Richards JC. 2004. Application of capillary electrophoresis-electrospray-mass spectrometry to the separation and characterization of isomeric lipopolysaccharides of *Neisseria meningitidis*. *Electrophoresis.* 25:2017–2025.

Liu Z, Lu Y, Zhang J, Pardee K, Wang PG. 2003. P1 trisaccharide (Galα1–4Galβ1–4GlcNAc) synthesis by enzyme glycosylation reactions using recombinant *Escherichia coli*. *Appl Environ Microbiol.* 69:2110–2115.

Mandrell RE, Partker CT, Bates AH, Vinogradov E, Li J, St. Michael F, Gilbert M. 2007. *Campylobacter jejuni* strain RM1221 lipooligosaccharide mimics globo- and lactoneo-series P blood group antigens. 107th General Meeting of the American Society for Microbiology. D-027.

Moran AP, Prendergast MM, Appelmelk BJ. 1996. Molecular mimicry of host structures by bacterial lipopolysaccharides and its contribution to disease. *FEMS Immunol Med Microbiol.* 16:105–115.

Nilsson U, Ray AK, Magnusson G. 1994. Synthesis of the Forssman pentasaccharide and terminal tetra-, tri-, and di-saccharide fragments. *Carbohydr Res.* 252:137–148.

Ono K, Hattori H, Uemura K, Nakayama J, Ota H, Katsuyama T. 1994. Expression of Forssman antigen in human large intestine. *J Histochem Cytochem.* 42:659–665.

Ramjeet M, Deslandes V, St. Michael F, Cox AD, Kobisch M, Gottschalk M, Jacques M. 2005. Truncation of the lipopolysaccharide outer core affects susceptibility to antimicrobial peptides and virulence of *Actinobacillus pleuropneumoniae* serotype 1. *J Biol Chem.* 280:39104–39114.

Risberg A, Masoud H, Martin A, Richards JC, Moxon ER, Schweda EK. 1999. Structural analysis of the lipopolysaccharide oligosaccharide epitopes expressed by a capsule-deficient strain of *Haemophilus influenzae* Rd. *Eur J Biochem.* 261:171–180.

Shao J, Zhang J, Kowal P, Lu Y, Wang PG. 2002. Overexpression and biochemical characterization of β-1,3-*N*-acetylgalactosaminyltransferase LgtD from *Haemophilus influenzae* strain Rd. *Biochem Biophys Res Commun.* 295:1–8.

St. Michael F, Vinogradov E, Li J, Cox AD. 2005. Structural analysis of the lipopolysaccharide from *Pasteurella multocida* genome strain Pm70 and identification of the putative lipopolysaccharide glycosyltransferases. *Glycobiology.* 15:323–333.

Troy FA, Vijay IK, Tesche N. 1975. Role of undecaprenyl phosphate in synthesis of polymers containing sialic acid in *Escherichia coli*. *J Biol Chem.* 250:156–163.

Wakarchuk W, Martin A, Jennings MP, Moxon ER, Richards JC. 1996. Functional relationships of the genetic locus encoding the glycosyltransferase enzymes involved in expression of the lacto-*N*-neotetraose terminal lipopolysaccharide structure in *Neisseria meningitidis*. *J Biol Chem.* 271:19166–19173.

Wakarchuk WW, Cunningham AM. 2003. Capillary electrophoresis as an assay method for monitoring glycosyltransferase activity. *Methods Mol Biol.* 213:263–274.

Willis LM, Gilbert M, Karwaski MF, Blanchard MC, Wakarchuk WW. 2008. Characterization of the α-2,8-polysialyltransferase from *Neisseria meningitidis* with synthetic acceptors, and the development of a self-priming polysialyltransferase fusion enzyme. *Glycobiology.* 18:177–186.

Xu H, Storch T, Yu M, Elliott SP, Haslam DB. 1999. Characterization of the human Forssman synthetase gene. An evolving association between glycolipid synthesis and host-microbial interactions. *J Biol Chem.* 274:29390–29398.

Yamamoto F, Clausen H, White T, Marken J, Hakomori S. 1990. Molecular genetic basis of the histo-blood group ABO system. *Nature.* 345:229–233.

Yokota M, Warner GA, Hakomori S. 1981. Blood group A-like glycolipid and a novel Forssman antigen in the hepatocarcinoma of a blood group O individual. *Cancer Res.* 41:4185–4190.

Yuki N, Taki T, Inagaki F, Kasama T, Takahashi M, Saito K, Handa S, Miyatake T. 1993. A bacterium lipopolysaccharide that elicits Guillain–Barré syndrome has a GM1 ganglioside-like structure. *J Exp Med.* 178:1771–1775.

Zhang J, Kowal P, Fang J, Andreana P, Wang PG. 2002. Efficient chemoenzymatic synthesis of globotriose and its derivatives with a recombinant α-(1→4)-galactosyltransferase. *Carbohydr Res* 337:969–976.