Susan M. Logan¹, Eleonora Altman, Oksana Mykytczuk, Jean-Robert Brisson, Vandana Chandan, Frank St. Michael, Amara Masson, Sonia Leclerc, Koji Hiratsuka, Natalia Smirnova, Jianjun Li, Yuyang Wu, and Warren W. Wakarchuk

Institute for Biological Sciences, National Research Council, Ottawa, Ontario, Canada K1A OR6

© The Author 2005. Published by Oxford University Press. All rights reserved. For permissions, please e-mail: journals.permissions@oupjournals.org

Activity screening and insertional inactivation of lipopolysaccharide (LPS) biosynthetic genes in Helicobacter pylori have led to the successful characterization of two key enzymes encoded by HP0159 (JHP0147) and HP1105 (JHP1032) open reading frames (ORFs) which are members of the large and diverse carbohydrate active enzymes (CAZY) GT-8 family of glycosyltransferases. Activity screening of a genomic library led to the identification of the enzyme involved in the biosynthesis of the type 2 N-acetyl-lactosamine O-chain backbone, the β-1,3-N-acetyl-glucosaminyl transferase. In addition, the activity screening approach led to the identification and characterization of a key core biosynthetic enzyme responsible for the biosynthesis of the α-1,6-glucan polymer. This α-1,6-glucosyltransferase protein is encoded by the HP0159 ORF. Both enzymes play an integral part in the biosynthesis of LPS, and insertional inactivation leads to the production of a truncated LPS molecule on the bacterial cell surface. The LPS structures were determined by mass spectrometry and chemical analyses. The linkage specificity of each glycosyltransferase was determined by nuclear magnetic resonance (NMR) analysis of model compounds synthesized in vitro. A cryogenic probe was used to structurally characterize nanomole amounts of the product of the HP1105 (JHP1032) enzyme. In contrast to the HP0159 enzyme, which displays the GT-8-predicted retaining stereochemistry for the reaction product, HP1105 (JHP1032) is the first member of this GT-8 family to have been shown to have an inverting stereochemistry in its reaction products.

Key words: CAZY 8 family/glycosyltransferases/ Helicobacter pylori/lipopolysaccharide

Introduction

Helicobacter pylori infection affects more than half of the world’s population, and the clinical spectrum of disease ranges from asymptomatic gastritis to peptic ulcer and gastric cancer (Dunn et al., 1997; Suerbaum and Michetti, 2002). As a major cell-surface component, lipopolysaccharide (LPS) is well situated to selectively interact with surface components of the host, and as a consequence, considerable structural analysis of H. pylori LPS has been described (Monteiro, 2001). These findings have revealed that the LPS of H. pylori appears to be unique both in structure and in function.

The structure of both the lipid A and polysaccharide regions differs from that of other gram-negative bacteria, and they have been shown to be immunologically and functionally distinct (Ogawa et al., 1997; Suda et al., 2001). The lipid A of H. pylori displays low biological (endotoxic) activity, and the O chain is unique in displaying structural homology to mammalian histo-blood group antigens. Although considerable effort has been invested in identifying the genetic determinants of O-chain biosynthesis (Cie et al., 1997; Wang et al., 1999; Logan et al., 2000; Rasko et al., 2000a,b), including the extensive characterization of the fucosyltransferase biosynthetic genes, a key biosynthetic target responsible for type 2 N-acetyllactosamine (LacNAc) backbone biosynthesis, the β-1,3-N-acetylglucosaminyl transferase, has yet to be identified. In addition, the genes responsible for unique components of the complex core structure of this organism remain elusive.

Glycosyltransferases display remarkable diversity in their biosynthetic capacity, and it is this diversity which results in the production of a multitude of complex carbohydrates and polysaccharides which are recognized as key players in numerous biological interactions (Allen and Kisailus, 1992). To meet the challenge of the postgenomic era and to facilitate the functional characterization of the increasing number of sequences revealed in sequenced genomes to be involved in glycosidic bond formation, researchers have established an evolving hierarchical family classification scheme (Coutinho et al., 2003). This scheme has identified 20 ORFs in the genome of H. pylori as potentially involved in some form of glycan biosynthesis within the cell. To date, supporting functional data for only four of these H. pylori carbohydrate biosynthetic targets have been provided in the literature.

We present here the functional characterization of two novel H. pylori glycosyltransferases which are CAZY GT-8 (rfaJ homologs) and define the role of each in LPS biosynthesis. We show that HP0159 is responsible for the synthesis of the polymeric α-1,6-glucan found attached to the outer core backbone, whereas, more surprisingly, the second rfaJ homolog, JHP1032, is responsible for the addition of β-1,3-GlcNAc to galactose (Gal) in the O-chain backbone.

Results

Glycosyltransferase activity in H. pylori

The complete genome sequences of two strains of H. pylori are currently available (Tomb et al., 1997; Alm et al., 1999).
Utilizing the CAZY classification system, only 22 \textit{H. pylori} ORFs have been assigned to 11 of the current 73 glycosyltransferase families (Table I). Structural analysis of \textit{H. pylori} LPS (Figure 1) has demonstrated that both the 1,4-linked Gal and 1,3-linked GlcNAc of the O-chain backbone are \(\beta\)-linked glycosidic bonds, which indicates that these glycosyltransferases use a reaction mechanism with inverting stereochemistry. In the core structure, only a single glycosidic linkage would be formed by an inverting glycosyltransferase (Gal \(\beta\)-1,7-Hep), whereas all the other linkages would be formed by enzymes with retaining stereochemistry.

To facilitate the functional characterization of glycosyltransferase genes identified by the CAZY classification system, we performed enzymatic assays with cell-free extracts of both genome-sequenced strains. The assays were performed with synthetic acceptor substrates which have previously been used for a variety of LPS biosynthetic enzymes (Gilbert et al., 1996, 1997, 2000). The assays were followed by thin layer chromatographic (TLC) analysis of the reaction mixtures as previously described and are presented in Figure 2 (Gilbert et al., 1996). We were able to detect transfer of glucose (Glc) from uridine-5’-diphospho (UDP)-Glc to 6-(5-fluorescein-carboxamido)-hexanoic acid succinimidyl ester-\(\alpha\)-Glc (FCHASE-\(\alpha\)-Glc), Gal from UDP-Gal to FCHASE-\(\beta\)-GlcNAc, and \(\alpha\)-N-acetyl-glucosaminyltransferase (GlcNAc) from UDP-GlcNAc to FCHASE-\(\beta\)-LacNAc activity in both 26695 and J99 lysates. In addition, in the J99 lysate, we observed a product of transfer of Gal from UDP-Gal to FCHASE-\(\alpha\)-Man. As indicated above, we have previously shown that the gene encoding the UDP-Gal to FCHASE-\(\beta\)-GlcNAc activity is HP0826, a member of GT-25 (Logan et al., 2000).

These analyses clearly demonstrated that both 26695 and J99 have glycosyltransferase enzymatic activities for which the corresponding gene function is yet to be assigned. These genes would be responsible for the transfer of \(\beta\)-GlcNAc to LacNAc and \(\alpha\)-Glc to Glc and \(\alpha\)-Gal to Man and are likely key LPS biosynthetic enzymes.

Table I. CAZY family assignment of \textit{Helicobacter pylori} glycosyltransferases

<table>
<thead>
<tr>
<th>CAZY family (mechanism)</th>
<th>Number of family members</th>
<th>26695 ORF (J99 ORF)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GT-2 (inverting)</td>
<td>3097</td>
<td>HP0102 (HP0094)</td>
</tr>
<tr>
<td>GT-4 (retaining)</td>
<td>2052</td>
<td>HP0421 (HP0963)</td>
</tr>
<tr>
<td>GT-8 (retaining)</td>
<td>309</td>
<td>HP0159 (HP0147, HP0208, HP0194), HP1416 (HP1311), HP1105 (HP1032/1031), HP1378, HP0820*</td>
</tr>
<tr>
<td>GT-9 (inverting)</td>
<td>257</td>
<td>HP1191 (waaF), HP0276 (waaC), HP1283, HP1284 (HP1204)</td>
</tr>
<tr>
<td>GT-10 (inverting)</td>
<td>139</td>
<td>HP0379 (fucU), HP0651 (fucT)</td>
</tr>
<tr>
<td>GT-11 (inverting)</td>
<td>128</td>
<td>HP0093.:HP0094 (HP0086)</td>
</tr>
<tr>
<td>GT-19 (inverting)</td>
<td>104</td>
<td>HP0867 (lpxB)</td>
</tr>
<tr>
<td>GT-25 (inverting)</td>
<td>141</td>
<td>HP0805 (HP0741) HP0826 (HP0765), HP0619 (HP0563/0562)</td>
</tr>
<tr>
<td>GT-28 (inverting)</td>
<td>227</td>
<td>HP1155 (murG)</td>
</tr>
<tr>
<td>GT-30 (inverting)</td>
<td>121</td>
<td>HP0957 (WaaA)</td>
</tr>
<tr>
<td>GT-51 (inverting)</td>
<td>653</td>
<td>HP0597 (HP0544)</td>
</tr>
</tbody>
</table>

\(^a\) probable gene duplication; \(^\_\) frame shift to single ORF.

\(^a\)No homolog in other genome sequenced strain.

\(N\)-Acetylglucosaminyl transferase activity

As we had no clear indication through bioinformatic analysis of the likely identity of the glycosyltransferase responsible for the addition of \(\beta\)-GlcNAc to LacNAc, we decided to identify the gene responsible by using an activity-based screening strategy (Gilbert et al., 1996; Endo et al., 2000). A plasmid library of \textit{H. pylori} J99 was constructed by using 1–2-Kb random sheared, gel purified, and end repaired DNA fragments. These fragments were ligated into pTrueBlue to produce 6000 white colonies that were picked to form 30 pools of 200 colonies. Each pool was screened for \(\beta\)-1,4-galactosyltransferase activity (UDP-Gal to FCHASE-GlcNAc) and for \(\beta\)-N-acetyl-glucosaminyltransferase activity [\(\beta\)-GlcNAc transferase (UDP-GlcNAc to FCHASE-LacNAc)], and pools with \(\beta\)-GlcNAc transferase activity were plated for single colonies and rescreened to identify individual clones displaying the same activity. The genomic library was used to successfully identify a 1831-bp clone with \(\beta\)-GlcNAc transferase activity. Sequence analysis revealed that the clone contained 230 base pairs of sequence 5’ to the JHP1032 gene (HP1105), the complete JHP1032 ORF, 138 bp of intergenic region, and the initial 265 bp of the JHP1031 ORF. Comparative analysis of this region in 26695 and J99 indicated that the JHP1032 and JHP1031 genes have likely arisen by gene duplication (60% identity), whereas in 26695 genome, there resides a single ORF (HP1105) which shows strong conservation at the protein level to both JHP1032 (73% identity) and JHP1031 (64% identity). To confirm that the JHP1032 ORF encodes the GlcNAc transferase activity, we subcloned the ORF into the expression vector pCW and measured enzyme activity following the induction of expression with isopropyl thio-\(\beta\)-D galactoside (IPTG) (data not shown). The product of the enzyme reaction using FCHASE-Lac as an acceptor (1 mg) was purified through a C-18 Sep-Pak (Millipore, Billerica, MA) and eluted with 50% acetonitrile. Part (0.2 mg) of the dried sample was used in an alditol acetate analysis. Gas liquid chromatography mass spectrometry (GLC-MS) of alditol acetates revealed Glc, Gal, which are both constituents of lactose, and GlcNAc, in the...
Functional characterization of rfaJ homologs from *H. pylori*

...ratio of ~1:1:1. Three major linkage types 4-Glc, 3-Gal, and terminal GlcNAc were observed by GLC-MS. Nuclear magnetic resonance (NMR) was used to confirm the linkage and determine that the GlcNAc was β-1,3 linked to Gal (see section NMR analysis of HPO159 and JHP1032 reaction products).

**Glucosyltransferase activity in strains 26695 and J99**

The second strong enzymatic activity detected in the screens of 26695 and J99 lysates was that of glucosyltransferase activity. A more sensitive capillary electrophoresis (CE) analysis of the products suggested that the product made by each strain contained unique linkages. In the case of 26695, multiple product peaks were obvious by both TLC and CE analysis. Assays with J99 lysates showed glucosyltransferase activity under the same reaction conditions, but only a single product was seen by TLC and CE. Methylation analysis of this product confirmed that no 1,6-linked Glc was present.

Structural studies of the *H. pylori* LPS core have revealed that along with the relatively rare incorporation of DD-α-Hep as a unique outer core branched structure, the LPS core also contains a second novel polymeric structure, an α-1,6-polymeric glucan (Figure 1). Fine structural analyses have shown this to be produced by most *H. pylori* strains (Monteiro et al., 2000, 2001) and with an average reported glucan chain corresponding to three to four glucose residues. In addition, some nontypable strains are capable of producing a longer α-1,6-glucan chain (Altman et al., 2003). To determine the level of heterogeneity present on *H. pylori* strains with respect to the α-1,6-glucan polymer, we completed a structural analysis of core LPS from many strains (Table II). Methylation analysis performed on the intact LPS or bacterial cells from strains 26695, SS1, PJ1, PJ2, O:3, and M6 confirmed the presence of 6-linked Glc in these strains, and the length of the respective glucan polymer was determined by electrospray mass spectrometry (ES-MS) of delipidated LPS (PJ1, PJ2), MS analysis (SS1), or methylation analysis (O:3, 26695). In comparison with other strains examined, SS1 appears to add only a single Glc residue in α-1,6 linkage to an adjacent Glc residue, which in turn is α-1,2 linked to branched DD-heptose in the outer core backbone of *H. pylori* LPS. J99, the second strain for which a complete genome sequence is available, appeared to be unable to add even a single residue of Glc to an adjacent Glc residue in α-1,6 linkage. This result correlated with the activity screening described above where no Glc polymer was made in the J99 lysate.

**Insertional mutagenesis of HP0159 and HP1416 CAZY family 8 members**

To facilitate the identification of the α-1,6 glycosyltransferase activity found in the 26695 lysate, we examined the...
CAZY 8 family rfaJ α-1,2-glucosyltransferase homologs from Helicobacter (Table I). Only two Helicobacter homologs are common to both strains and appear to produce a similar full-length product, HP0159 (JHP0147) and HP1416 (JHP1311). Helicobacter pylori mutants carrying disrupted genes were constructed by allelic exchange in 26695. Attempts to make the equivalent disrupted genes were constructed by allelic exchange in HP1416 (JHP1311).

To produce a similar full-length product, HP0159 (JHP0147) and HP1416 (JHP1311) were established by complete assignment of their proton and 13C bonds (Gilbert et al., 2000). The analysis was performed on a conventional 3-mm probe (Figure 4B). This increased sensitivity made possible the acquisition of data on a dilute sample that would otherwise be impractical. Hence, it was possible to acquire an overnight heteronuclear single quantum coherence (HSQC) and overnight HMBC spectra. For the FEX-glycosides, more material was available and experiments were done using a conventional 5-mm NMR probe. The samples were dissolved in 600 µL of D2O and run at high temperature. The doublet of the anomeric resonances (J1,2 = 3.5 Hz) was not resolved because of linewidths of ~5 Hz, possibly owing to aggregation. However, it was possible to obtain an HSQC spectrum overnight. As can be assessed from the proton spectra in Figure 4C and D, all compounds were pure and impurities or degradation products that were present, because of instability of the FCHASE or FEX residue, did not interfere with the NMR analysis which was performed as previously described (Gilbert et al., 2000).

The enzymatic product of the Lac-FCHASE acceptor was identified as βGlcNAc(1-3)βGal(1-4)βGlc-FCHASE. The proton resonances of the Glc unit were assigned from a selective total correlation spectroscopy (TOCSY) experiment on the H-1 resonance of Glc (Figure 5A and B). The ring resonances for the Gal residue were assigned from a selective TOCSY on the H-4b resonance (Figure 5C). The resonances for the residue c were located using a selective TOCSY on the H-1c resonance (Figure 5D). Selective TOCSY experiments for the H-5a, H-5b, and H-5c resonances were used to assign their respective H6 and H6′ resonances.

**Expression of α-1,6-glucosyltransferase**

Analysis of the HP0159 gene from the 26695 strain cloned as a fusion protein with the maltose-binding protein into the expression vector pCW in Escherichia coli also revealed multiple reaction products as detected by TLC similar to those observed with the H. pylori lysates. The more sensitive CE analysis of the reaction mixture revealed that these products represent the sequential addition of Glc molecules in a processive reaction at the nonreducing end of up to six molecules (Figure 3). NMR analysis of a disaccharide and trisaccharide produced was used to determine the nature of the linkage.

**NMR analysis of HP0159 and JHP1032 reaction products**

The linkage specificity for each glycosyltransferase was established by complete assignment of their proton and 13C resonances (Table III), comparison of chemical shifts and JH,H with those of model compounds, and detection of nuclear Overhauser effect (NOE) or heteronuclear multiple bond coherence (HMBC) correlations across the glycosidic bonds (Gilbert et al., 2000). The analysis was performed on nanomole amounts of purified reaction product to reduce the time needed for the purification of the enzymatic products. FCHASE and fluorescein-5-EX succinimidyl ester (FEX) are identical fluorophores and differ only in the spacer between the aminophenyl glycoside and the fluorophore. They are used interchangeably for labeling acceptor sugars, and no influence on the activity of the enzymes used in this work was noted.

In Figure 4A, the proton NMR spectra of FCHASE-glycoside compound (I) from the JHP1032 reaction are shown. The compound was soluble with linewidths of a few Hz because the H-1 anomic doublets (J1,2 = 8 Hz) were well resolved. Owing to the limited quantity of material available (0.3 mg), experiments were carried out on a cryogenically cooled probe, which offers the greatest sensitivity. An improvement by a factor of 6 was obtained when compared with the same sample acquired under similar conditions on a 500-MHz spectrometer by using a conventional 3-mm probe (Figure 4B). This increased sensitivity made possible the acquisition of data on a dilute sample that would otherwise be impractical. Hence, it was possible to acquire an overnight heteronuclear single quantum coherence (HSQC) and overnight HMBC spectra. For the FEX-glycosides, more material was available and experiments were done using a conventional 5-mm NMR probe. The samples were dissolved in 600 µL of D2O and run at high temperature. The doublet of the anomeric resonances (J1,2 = 3.5 Hz) was not resolved because of linewidths of ~5 Hz, possibly owing to aggregation. However, it was possible to obtain an HSQC spectrum overnight. As can be assessed from the proton spectra in Figure 4C and D, all compounds were pure and impurities or degradation products that were present, because of instability of the FCHASE or FEX residue, did not interfere with the NMR analysis which was performed as previously described (Gilbert et al., 2000).

The enzymatic product of the Lac-FCHASE acceptor was identified as βGlcNAc(1-3)βGal(1-4)βGlc-FCHASE. The proton resonances of the Glc unit were assigned from a selective total correlation spectroscopy (TOCSY) experiment on the H-1 resonance of Glc (Figure 5A and B). The ring resonances for the Gal residue were assigned from a selective TOCSY on the H-4b resonance (Figure 5C). The resonances for the residue c were located using a selective TOCSY on the H-1c resonance (Figure 5D). Selective TOCSY experiments for the H-5a, H-5b, and H-5c resonances were used to assign their respective H6 and H6′ resonances.
The linkage site was at Gal C-3 because of its large downfield shift of 9.2 ppm upon glycosidation (Gilbert et al., 2000).

The α-1,6-glucosyltransferase activity was demonstrated by analysis of products with αGlc-FEX as an acceptor. The resulting disaccharide and trisaccharide products were analyzed by correlated spectroscopy (COSY), TOCSY, NOESY, and HSQC experiments, as described previously (Gilbert et al., 2000). The products were found to be αGlc(1–6)αGlc-FEX and αGlc(1–6)αGlc-FEX (Table III). Their 13C chemical shifts were in agreement with those of the αGlc(1–6)αGlc(1–6)Glc oligosaccharide (Bock et al., 1984).

**Galactosyltransferase activity in J99**

We attempted to measure Gal transfer to heptose by using a synthetic mannose acceptor as a surrogate acceptor. Previously, we have shown that an LPS α-GlcNAc transferase which uses heptose as an acceptor would transfer to a mannose-based acceptor (Wakarchuk et al., 1994). Characterization of the reaction products by methylation and glycosidase digestion showed them to be a mixture of disaccharides with Glc (not Gal) in β linkage (data not shown), a product possibly representative of the periplasmic glucans which are commonly found in many gram-negative bacteria and likely not a product of an LPS biosynthetic enzyme (Bohin, 2000).

### Table III. Proton (δH) and 13C (δC) chemical shifts for the glycose units of βGlcNAc(1–3)βGal(1–4)βGlc-FCHASE (I), αGlc(1–6)αGlc-FEX (II), and αGlc–6αGlc(1–6)αGlc-FEX (III)

<table>
<thead>
<tr>
<th>Residue</th>
<th>Chemical shift (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>H-1</td>
</tr>
<tr>
<td>−4)βGlc (I-a)</td>
<td>4.57</td>
</tr>
<tr>
<td>-3)βGal (I-b)</td>
<td>99.5</td>
</tr>
<tr>
<td>βGlcNAc (I-c)</td>
<td>103.9</td>
</tr>
<tr>
<td>βGlc (II-b)</td>
<td>4.72</td>
</tr>
<tr>
<td>αGlc (II-b)</td>
<td>103.6</td>
</tr>
<tr>
<td>−6)αGlc (III-a)</td>
<td>5.35</td>
</tr>
<tr>
<td>αGlc (III-b)</td>
<td>98.6</td>
</tr>
<tr>
<td>−6)αGlc (III-b)</td>
<td>4.75</td>
</tr>
<tr>
<td>αGlc (III-c)</td>
<td>99</td>
</tr>
<tr>
<td>−6)αGlc (III-b)</td>
<td>5.35</td>
</tr>
<tr>
<td>αGlc (III-b)</td>
<td>98.3</td>
</tr>
<tr>
<td>αGlc (III-c)</td>
<td>98.3</td>
</tr>
<tr>
<td>αGlc (III-c)</td>
<td>98.8</td>
</tr>
</tbody>
</table>

Data obtained from the 600-MHz heteronuclear single quantum coherence (HSQC) spectra of the samples in D2O at 20°C for I, 50°C for II, 40°C for III. Chemical shifts are referenced to the methyl resonance of acetone set at δH 2.225 ppm and δC 31.07 ppm. The error is ±0.02 ppm for δH and ±0.2 for δC. For βGlcNAc, the NAc–CH3 resonance is at δH 2.06, δC 23.1 and the NAc–CO resonance is at δC 175.8 from the heteronuclear multiple bond coherence (HMBC) spectrum.
Fig. 5. Nuclear magnetic resonance (NMR) experiments for resonance assignments of β-GlcNAc(1–3)β-Gal(1–4)β-Glc-FCHASE. Spectra were obtained at 600 MHz and 20°C by using a cold probe. (A) Proton spectrum showing the glycoside resonances from 3.1 to 4.8 ppm (F2). Selective total correlation spectroscopy (TOCSY) spectra for the H-1a resonance (B), the H-4b resonance (C), and the H-1c resonance (D) acquired with a mixing time of 80 ms. Selective nuclear Overhauser enhancement spectroscopy (NOESY) spectra for the H-1a resonance (E), the H-1b resonance (F), and the H-1c resonance (G) acquired with a mixing time of 800 ms. (H) Heteronuclear single quantum coherence (HSQC) spectrum was used to assign the 13C resonances.

Enzymatic characterization of JHP1032 and HP0159

To enhance the expression and solubility for functional characterization, we expressed both enzymes as C-terminal fusions to the E. coli MalE protein (Wakarchuk et al., 2004). The JHP1032 fusion protein was inherently unstable and was predisposed to form aggregates such that the activity level fluctuated, preventing a comprehensive kinetic study. The MalE fusion protein for HP0159 could be purified and was predisposed to form aggregates such that the activity was observed. All sugars were present in the pyranose form.

To determine the role of HP1105 (JHP1032) ORF in LPS biosynthesis, we constructed an H. pylori mutant carrying a disrupted gene by allelic exchange mutagenesis in 26695 as described above for HP0159. PCR analysis was used to confirm that the Km6 cassette was inserted into the chromosomal copy of HP1105 and that no wild-type copy of the gene was present in the cell (data not shown). Enzymatic analysis of mutant lysate confirmed the loss of HP1105 activity when compared with the parent strain (data not shown).

Structural characterization of H. pylori LPS mutant 26695::HP0159kan

Growth of bacterial strains was carried out as described previously (Logan et al., 2000), and LPS was isolated by phenol–water extraction procedure (Westphal and Jann, 1965). Crude aqueous-phase soluble LPS was subjected to further purification by ultracentrifugation. Sugar analysis of the intact LPS of H. pylori 26695::0159kan as alditol acetates revealed the presence of L-fucose (L-Fuc), D-glucose (D-Glc), D-galactose (D-Gal), N-acetyl-D-glucosamine (D-GlcNAc), D-glycero-D-manno-heptose (DD-Hep), and L-glycero-D-manno-heptose (LD-Hep) with the following molar ratios 0.5:6.3:3.0:8.7:2.8:1.0, respectively, and showed significant reduction in L-Fuc and D-Gal, when compared with the parent LPS, indicating the presence of the structure devoid of O chain. Methylation analysis of the intact LPS from 26695::0159kan showed the presence of 3-substituted L-Fuc, terminal D-Glc, 3- and 4-substituted D-Gal, terminal 2-, 7-, and 2,7-substituted DD-Hep, 2- and 3-substituted LD-Hep, and terminal 3-substituted D-GlcNAc. No 3,4-substituted GlcNAc, terminal Fuc, and 2-linked Gal, characteristic of the O chain containing Lewis x (Le^x) and Lewis y (Le^y) epitopes, were detected (Table IV). These conclusions were further confirmed by western analysis of 0159 mutants with anti-Le^x and anti-Le^y-specific monoclonal antibodies (lanes 4 and 6 in Figure 6A).

In addition, the methylation analysis of LPS from 26695::HP0159kan and O:3::HP0159kan revealed the presence of 4-substituted D-Glc, and no 6-substituted D-Glc was observed. All sugars were present in the pyranose form. NMR analysis of a high-molecular mass fraction, isolated by gel filtration chromatography from a partially delipitated LPS from 26695::HP0159kan, indicated it to contain β-1,4-linked glucan, a contaminant produced by some strains of H. pylori (Knirel et al., 1999).

To deduce the sequence information on the outer extremities of the LPS molecule, we subjected methylated intact LPS to the fast atom bombardment mass spectrometric analysis in the positive mode. A-Type primary glycosyl oxonium ions containing Lewis blood group-related Fuc and GlcNAc residues were observed at m/z 260 [GlcNAc]^+ and m/z 682 [Fuc, GlcNAc, Hep]^+. No higher mass ions representing a glucosylated DD-Hept residue were detected.
This evidence together with the absence of 6-substituted Glc in methylation analysis indicated this LPS mutant to be deficient in the biosynthesis of α(1-6)-glucan present in 26695 parent strain (Table IV). Fast atom bombardment mass spectroscopy (FAB-MS) spectra of methylated LPS from\textit{H. pylori} 26695::HP0159kan showed the primary ion at \( m/z \) 668 and its corresponding secondary ion at \( m/z \) 228, indicative of the type 1 linear B blood group [Gal(1-3)Gal(1-3)GlcNAc] antigen, a blood group antigen found in the LPS of 26695 and SS1 (Monteiro et al., 2000). Other Lewis blood group-related secondary ions were observed at \( m/z \) 228 (260-32) [GlcNAc]\(^+\), 402 (434-32) [Fuc, GlcNAc]\(^+\), 576 (608-32) [Fuc(1-3)Fuc(1-4)GlcNAc]\(^+\), as described previously (Monteiro et al., 1998; Logan et al., 2000). The LPS structures identified in 26695::HP0159 mutant strain are shown in Figure 7A.

Table IV. Methylation analysis of the intact aqueous-phase lipopolysaccharide (LPS) from strains 26695, 26695::HP0159kan, and 26695::HP1105kan showing approximate molar ratios (based on the detector response, total ion count)

<table>
<thead>
<tr>
<th>Sugar linkage</th>
<th>26695</th>
<th>26695::HP0159kan</th>
<th>26695::HP1105kan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Terminal Fuc</td>
<td>2.6</td>
<td>—</td>
<td>0.1</td>
</tr>
<tr>
<td>3-linked Fuc</td>
<td>0.7</td>
<td>0.2</td>
<td>0.8</td>
</tr>
<tr>
<td>Terminal Glc</td>
<td>1.3</td>
<td>0.4</td>
<td>1.1</td>
</tr>
<tr>
<td>Terminal Gal</td>
<td>1.0</td>
<td>—</td>
<td>0.9</td>
</tr>
<tr>
<td>3-Linked Glc</td>
<td>1.5</td>
<td>—</td>
<td>1.3</td>
</tr>
<tr>
<td>4-Linked Gal</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>4-Linked Glc</td>
<td>2.7</td>
<td>0.1</td>
<td>—</td>
</tr>
<tr>
<td>6-Linked Glc</td>
<td>3.0</td>
<td>—</td>
<td>3.4</td>
</tr>
<tr>
<td>2-Linked Gal</td>
<td>0.4</td>
<td>—</td>
<td>0.1</td>
</tr>
<tr>
<td>3-Linked Gal</td>
<td>8.0</td>
<td>0.3</td>
<td>—</td>
</tr>
<tr>
<td>Terminal DD-Hep</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>2-Linked DD-Hep</td>
<td>2.5</td>
<td>0.2</td>
<td>1.4</td>
</tr>
<tr>
<td>6-Linked DD-Hep</td>
<td>1.4</td>
<td>—</td>
<td>1.1</td>
</tr>
<tr>
<td>2-Linked LD-Hep</td>
<td>2.2</td>
<td>2.0</td>
<td>1.4</td>
</tr>
<tr>
<td>3-Linked DD-Hep</td>
<td>2.6</td>
<td>—</td>
<td>3.0</td>
</tr>
<tr>
<td>7-Linked DD-Hep</td>
<td>0.4</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>3-Linked LD-Hep</td>
<td>0.2</td>
<td>0.1</td>
<td>0.2</td>
</tr>
<tr>
<td>2,7-Linked DD-Hep</td>
<td>1.9</td>
<td>2.0</td>
<td>1.4</td>
</tr>
<tr>
<td>3,7-Linked LD-Hep</td>
<td>—</td>
<td>—</td>
<td>0.1</td>
</tr>
<tr>
<td>Terminal GlcNAc</td>
<td>1.4</td>
<td>0.1</td>
<td>0.3</td>
</tr>
<tr>
<td>4-Linked GlcNAc</td>
<td>2.3</td>
<td>—</td>
<td>0.5</td>
</tr>
<tr>
<td>3-Linked GlcNAc</td>
<td>0.6</td>
<td>0.1</td>
<td>0.9</td>
</tr>
<tr>
<td>3,4-Linked GlcNAc</td>
<td>2.0</td>
<td>—</td>
<td>0.2</td>
</tr>
</tbody>
</table>

The structural characterization of \textit{H. pylori} LPS mutant 26695::HP1105kan

Sugar analysis of the intact aqueous-phase LPS of \textit{H. pylori} 26695::HP1105kan revealed the presence of L-Fuc, D-Glc, D-Gal, GlcNAc, DD-Hep, and LD-Hep in the molar ratio of -0.4(2.0):3.5(3.2):1.6(4.8):2.7(7.1):3.8(3.5):1.0(1.0) and showed significant reduction in L-Fuc, D-Gal, and D-GlcNAc, sugar constituents of Le O-antigen, as compared with 26695 LPS (sugar ratios in parentheses). Purified aqueous-phase LPS was analyzed by SDS-PAGE, and immunoblotting revealed that it was negative for Le\(^x\) and Le\(^y\) antigens (Figure 6B). An enzyme-linked immunosorbent assay performed on aqueous-phase LPS confirmed that it was devoid of Le\(^a\) and Le\(^b\) antigens (data not shown). Methylation analysis carried out on the intact aqueous-phase LPS from 26695::HP1105kan

Fig. 6. Silver-stained sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) and immunoblot analysis of purified lipopolysaccharide (LPS). (A) Lanes 1, 3, and 5: 26695 LPS; lanes 2, 4, and 6: 26695 HP0159::kan LPS; lanes 1 and 2: LPS silver stain; lanes 3 and 4: western blot with anti-Lewis x (anti-Le\(^x\)); lanes 5 and 6: western blot with anti-Lewis y (anti-Le\(^y\)). (B) Lanes 1, 3, and 5: 26695 LPS; lanes 2, 4, and 6: 26695 HP1105::kan LPS; lanes 1 and 2: LPS silver stain; lanes 3 and 4: western blot with anti-Le\(^a\); lanes 5 and 6: western blot with anti-Le\(^b\).
confirmed that in accordance with SDS–PAGE and western blot analysis, it contained only negligible amounts of sugars associated with the presence of Le O-antigen, namely terminal Fuc, 2-linked Gal and 3,4-substituted GlcNAc, and no 3-linked Gal. The presence of 3,7-substituted LD-Hep, 2- and 3-substituted LD-Hep (Table IV) was consistent with the presence of the backbone core oligosaccharide fragment capped with [Hep, Fuc, GlcNAc] or [GlcNAc] as found in 26695::HP0159kan mutant LPS (Figure 8A) and other strains of H. pylori (Logan et al., 2000; Altman et al., 2003). In addition, CE-MS analysis in the positive ion mode suggested the presence of a series of triply charged ions, consistent with the consecutive addition of Hex residues, the longest glucan chain corresponding to ~14 residues and the most abundant glycoform containing 12 Glc residues, as confirmed in a separate tandem MS (MS/MS) experiment in which a triply charged ion at m/z 1319.5 was selected as a precursor (Figure 8). Its fragmentation pattern was consistent with the fragment ion Hex$_3$Hep$_3$Fuc at m/z 2474.8 and the previously observed backbone core oligosaccharide fragment GlcNAcHex$_2$Hep$_3$ (PE)KDO at m/z 1446.5. The presence of multiple Hex$_n$Hep fragments containing up to six Hex was also evident (Figure 8). In addition, consistent with FAB-MS results, fragment ion
at m/z 366 suggested that some core structures were also capped by terminal [LacNAc] residues. The LPS structures identified in 26695::HP1105 mutant strain are shown in Figure 7B.

Discussion

To identify LPS biosynthetic enzymes from *H. pylori*, we have employed a strategy utilizing both activity screening and insertional inactivation, which has led to the successful characterization of two key enzymes which were predicted to be functional homologs of RfaJ, an α(1,2)-glucosyltransferase (Heinrichs et al., 1999). Both enzymes play an integral role in the biosynthesis of the LPS molecule and, in the CAZY classification system, are members of the large and diverse family 8 group of glycosyltransferases. In this study, we demonstrate that each enzyme displays a unique activity which is not predicted from the functional assignment based only on sequence homology.

The first enzyme encoded by the HP1105 (JHP1032) gene is responsible for the addition of GlcNAc in β-1,3 linkage to Gal in the O-chain backbone, and the chemistry of reaction of this enzyme is inverting. Mutation of this gene leads to the production of truncated LPS molecules devoid of Lewis antigens but which retain a complete inner and outer core structure. Formation of a long glucan chain was observed in this strain, possibly compensating for the loss of O chain.

The sequence-based CAZY classification system is such that given families contain enzymes which display the same stereochemical outcome and therefore can be annotated as a putative retaining or inverting chemistry. In the case of GT-8 members, the functional characterization and crystal structure of LgtC and glycogenin have provided considerable data to indicate that members of this family would display a retaining chemistry and GT-A fold. An alignment of the six RfaJ homologs from *H. pylori* with LgtC is presented in Figure 9. The three-dimensional structure of LgtC has shown that there are two critical DXD motifs involved in the active site (Persson et al., 2001). The first motif at residues 103–105 in LgtC is involved in binding Mn^{2+} as well as the phosphates from the UDP-sugar donor. The first DXD motif involved in donor binding is conserved in all of the HP GT-8 family members. The second DXD motif in LgtC, residues 188–190, is involved in the catalytic mechanism, and a critical residue is Q189. This motif is present as either a DQD or EQD in four of the six HP homologs. The remaining two enzymes HP1578 and HP1105/JHP1032 do not have the first acidic residue, and HP1105/JHP1032 enzyme is missing the critical Q189 equivalent. It is tempting to speculate that this is an indicator of the change in reaction stereochemistry from retaining to inverting for a subgroup of GT-8 enzymes.

The enzyme encoded by the HP0159 (JHP0147) gene was shown to be responsible for the biosynthesis of the novel

![Alignment of rfaJ homologs from Helicobacter pylori.](https://academic.oup.com/glycob/article-abstract/15/7/721/689780)


Materials and methods

Bacterial strains

\textit{Helicobacter pylori} strains 26695 and J99 were obtained from R. Alm (Astra Zeneca, Boston, MA), SS1 from A. Lee (Sydney, Australia), O:3 isolate from J. Penner (Toronto, Ontario), M6 isolate from K. Eaton (Ann Arbor, MI), and PJ1 and PJ2 from W. Conlan (Ottawa, Ontario) and were
grown at 37°C on antibiotic-supplemented tryptic soy agar plates containing 7% horse blood in a microaerophilic environment for 48 h. For growth in liquid culture, antibiotic-supplemented brucella broth containing 5% fetal bovine serum (FBS) was inoculated with \textit{H. pylori} cells at a starting OD\textsubscript{600} of 0.15. Flasks were incubated for 24–36 h in a tri-gas incubator on a shaking platform. \textit{Escherichia coli} strain DH10B or AD202 was used as the host for plasmid cloning experiments and was grown on Luria broth (LB) plates supplemented with ampicillin (50 µg/mL) and kanamycin (20 µg/mL) where necessary.

Glycosyltransferase assays

Glycosyltransferase assays were performed essentially as described previously (Gilbert et al., 1997) by using either cell extracts of \textit{H. pylori} or \textit{E. coli} expressing recombinant protein. Reactions contained 0.5 mM FCHASE-aminophenyl-β-GlcNAc and 0.5 mM UDP-Gal or FCHASE-aminophenyl-α-Glc and UDP-Glc or 0.5 mM FCHASE-aminophenyl-β-Lac-NAc and 0.5 mM UDP-GlcNAc. Initial reactions screening for activity contained 10 mM MnCl\textsubscript{2}, 10 mM MgCl\textsubscript{2}, 1 mM dithiothreitol (DTT), and 50 mM N-[2-hydroxyethyl]piperazine-N'-[2-ethanesulfonic acid] (HEPES), pH 7.4. Optimized reactions for HP0159 were performed with 1 mM FCHASE-aminophenyl-α-Glc, 1 mM UDP-Glc, 10 mM MnCl\textsubscript{2}, and 50 mM 2-(-N-morpholino)ethanesulfonic acid (MES)-NaOH, pH 6.0. Optimized reactions for JH1032 were performed with 2 mM FCHASE-aminophenyl-β-Lac-NAc, 1 mM UDP-GlcNAc, 10 mM MgCl\textsubscript{2}, and 50 mM MES-NaOH, pH 6.5. For the determination of kinetic parameters for HP0159, the reactions contained a range of UDP-Glc from 50 to 2000 µM with a fixed FCHASE-aminophenyl-α-Glc concentration of 1 mM. The reciprocal reactions had the concentration of UDP-Glc fixed at 1 mM and the concentration of FCHASE-aminophenyl-α-Glc varied from 25 to 1500 µM. All reactions were incubated at 37°C for various lengths of times, where TLC or CE followed the reaction progress. The quantitation of the reactions was performed by integrating the CE trace peaks by using the PACE Station software (Beckman, CA). Kinetic parameters were calculated using Prism 3 software (GraphPad Software, San Diego, CA). TLC analysis was performed with aluminum-backed silica plates that were developed in ethyl acetate : methanol : water : acetic acid (7:2:1:0.1 or 4:2:1:0.1). Reaction products for NMR analysis were purified by preparative TLC, followed by desalting on SepPak C18 SPE cartridges. Elution was with 70% acetonitrile, followed by freeze drying to remove the solvent.

Methylation analysis of reaction products

Sugar composition analysis was performed by the alditol acetate method (Sawardeker et al., 1967), and methylation analysis was carried out by the NaOH/DM/CHO\_I procedure (Ciucanu and Kerek, 2004) on GlcNAc-Lac-FCHASE. The hydrolysis was performed in 4 M trifluoroacetic acid for 4 h at 100°C, followed by reduction with NaBD\textsubscript{4} in H\textsubscript{2}O overnight, and hydrolysis products then acetylated with acetic anhydride at 100°C for 2 h by using residual sodium acetate as catalyst. The methyl derivatives were characterized.
by GLC-MS by using a Hewlett-Packard chromatograph equipped with a 30 M DB-17 capillary column (180–230°C at 2.5°C/min). MS was carried out in the electron impact mode and recorded on a Varian Saturn II mass spectrometer.

J99 chromosomal DNA was nebulized with an IPI nebulizer (Medex, Toronto, Ontario) at 12 psi for 30 s, and then DNA was collected by ethanol precipitation. Following end repair, the sheared DNA was resolved on 1% agarose trisacetate/EDTA gel, and the region of the gel corresponding to fragments of size 1–2 kb was excised and purified by using a Qiagen agarose purification kit (Mississauga, Ontario). The purified genomic DNA was then ligated to SmaI digested pTrueblue (BioCan Scientific, Etobicoke, Ontario). Recombinant plasmid was electroporated into DH10B, and transformants were selected on LB plates supplemented with ampicillin (50 µg/mL). Colonies were picked into 0.5-mL brain heart infusion broth with 20% glycerol (30 pools containing 200 colonies/pool) and frozen at –80°C overnight. Cell pellets were resuspended in 600 µL of 50 mM HEPES, pH 7.4, and sonicated for 30 s with a microtip probe, and then the unclarified lysate was kept on ice before assay. β-1,4-Galactosyltransferase and β-1,3-N-acetylglucosaminyl transferase reactions were performed on each of the 30 pools by using 5 µL of sonicated cell extract in a final assay volume of 10 µL (see section Glycosyltransferase assays). Initial screening was performed by TLC analysis and then by CE to confirm products.

Recombinant expression of HP0159 and JHP1032
PCR primers were designed to amplify each ORF by using 26695 sequence data. Each product was cloned into pUC19 and plasmid DNA purified and sequenced. Briefly, each clone was disrupted by using reverse primers that were internal to each gene in a PCR, which resulted in the deletion of ~30 bp within each ORF. This was followed by ligation of a kanamycin cassette (Labigne-Roussel et al., 1988) to the gel-purified product of the PCR, generating plasmids pHPl105::kan, pHPl1416::kan, and pHPl0159::kan. The mutated allele was returned to H. pylori by natural transformation, according to the method of Haas et al. (1993), and double recombination confirmed by PCR analysis.

LPS isolation and delipidation
LPS was isolated by the hot phenol–water extraction procedure (Westphal and Jann, 1965) and, following the removal of insoluble material by low-speed centrifugation, purified by ultracentrifugation (105,000 × g, 4°C, 12 h).

Purified LPS (20 mg) was hydrolyzed in 0.1 M sodium acetate buffer, pH 4.2, for 2 h at 100°C. The solution was cooled and the precipitated lipid A removed by low-speed centrifugation. The supernatant solution was lyophilized and water-soluble components were fractionated by gel filtration on a Bio-Gel P-2 column (1.6 cm × 95 cm) equilibrated with pyridinium acetate (0.02 M, pH 5.4). Elution was performed with pyridinium acetate (0.02 M, pH 5.4). The fractions (1 mL) were monitored for neutral glycoses, and those giving positive reaction were combined and lyophilized.

Sugar composition and methylation analyses of LPS
Sugar composition analysis was performed by the alditol acetate method (Sawardeker et al., 1967). The hydrolysis was done in 4 M trifluoroacetic acid at 100°C for 4 h or 2 M trifluoroacetic acid at 100°C for 16 h, followed by reduction in water with sodium borohydride and subsequent acetylation with acetic anhydride/pyridine. Alditol acetate derivatives were analyzed by GLC-MS by using Hewlett-Packard chromatograph equipped with a 30 M DB-17 capillary column [210°C (30 min) to 240°C at 2°C/min], and MS spectra in the electron impact mode were recorded by using a Varian Saturn II mass spectrometer. Methylation linkage analysis was carried out by the NaOH/DMSO/CH3I procedure (Ciucanu and Kerek, 2004) and with the characterization of permethylated alditol acetate derivatives by GLC-MS in the electron impact mode (DB-17 column, isothermally at 190°C for 60 min).

Fast atom bombardment mass spectrometry
A fraction of the methylated sample was used for positive ion FAB-MS performed on a JEOL JMS-AX505H mass spectrometer with glycerol–thioglycerol (1:3) as the matrix. A 6-kV xenon beam was used to produce pseudomolecular ions that were then accelerated to 3 kV and their mass was analyzed. Product ion scan (B/E) and precursor ion scan (B2/E) were performed on metastable ions created in the first free field with a source pressure of 5 × 10⁻³ torr.
Electrospray mass spectroscopy

Samples were analyzed on a crystal Model 310 CE instrument (ATI Unicam, Boston, MA) coupled to a Q-Star quadrupole/time-of-flight mass spectrometer (Applied Biosystems/Sciex, Concord, Ontario) via a micro-ionspray interface. Sheath solution (isopropanol-methanol, 2:1) was delivered at a flow rate of 1 µL/min. An electrospray stainless steel needle (27-gauge) was butted against the low dead volume tee and enabled the delivery of the sheath solution to the end of the capillary column. The separations were obtained on ∼90-cm length bare-fused silica capillary by using 10 mM ammonium acetate in deionized water, pH 9.0, containing 5% methanol. A voltage of 25 kV was typically applied at the injection. The outlet of the capillary was tapered to −15 µm (internal diameter) by using a laser puller (Sutter Instruments, Novato, CA). Mass spectra were acquired with dwell times of 3.0 ms per step of 1 ms unit in full-mass scan mode. Fragment ions formed by collision activation of selected precursor ions with nitrogen in the radio frequency-only quadrupole collision cell were registered by mass and analyzed by time-of-flight mass analyzer.

SDS–PAGE and western blot

SDS–PAGE was performed according to the method of Laemmli (1970). Samples solubilized in sample buffer were stacked in 4.5% acrylamide and separated in 12.5% acrylamide. For western blot experiments, LPS was transferred to nitrocellulose membranes by the method of Towbin et al. (1979). Blots were incubated with mouse monoclonal antibodies specific for either anti-Leα or anti-Leβ, and binding was visualized with horseradish peroxidase-conjugated goat anti-mouse antibody.

NMR spectroscopy

For the FEX-glycosides, NMR experiments were performed at 600 MHz (1H) by using a 5-mm Z gradient triple resonance probe. NMR samples were prepared from 1 mg of compound dissolved in 600 µL of D2O and inserted in a 5-mm NMR tube. For the FCHASE-glycoside, 0.3 mg of sample was dissolved in 160 µL of D2O and inserted in a 3-mm NMR tube. For this sample, NMR experiments were performed at 600 MHz (1H) by using a 5-mm Z gradient triple resonance probe that is cryogenically cooled (Varian, Mississauga, Ontario). All NMR experiments were performed as previously described, by using standard techniques such as COSY, TOCSY, NOESY, selective TOCSY, HSQC, and HMBC (Uhrin and Brisson, 2000). For the proton chemical shift reference, the methyl resonance of internal acetone was set at 31.07 ppm at 2.225 ppm (1H). For the 13C chemical shift reference, the methyl resonance of internal acetone was set at 31.07 ppm relative to external dioxygen at 67.40 ppm.

Acknowledgments

The authors thank Marie-Claude Blanchard, Annie Aubry, Neety Panu, J. Carey, L. Willis, and Sonia Bardy for technical assistance and T. Devescari for help with photography.

Abbreviations

COSY, correlated spectroscopy; D-Gal, D-galactose; D-Glc, D-glucose; D-GlcNAc, N-acetyl-D-glucosamine; DD-He, D-glycero-D-manno-heptose; ES-MS, electrospray mass spectrometry; FAB-MS, fast atom bombardment mass spectrometry; FCHASE, 6-(5-fluorescein-carboxamido)-hexanoic acid succinimidyl ester; FEX, fluorescein-5-EX succinimidyl ester; Gal, galactose; GlcNAc, N-acetyl-glucosamine; GLC-MS, gas liquid chromatography mass spectrometry; HMBC, heteronuclear multiple bond coherence; HSQC, heteronuclear single quantum coherence; LacNAc, N-acetyllactosamine; L-D, Hep, L-glycero-D-manno-heptose; Leα, Lewis x; Leβ, Lewis y; PS, lipopolysaccharide; S-Fuc, L-fucose; GlcNAc, N-acetyl-D-glucosamine; Lac-α-2,3-sialyltransferase, its acceptor specificity.

References


