The lipopolysaccharide biosynthesis locus of Campylobacter jejuni 81116

Ben N. Fry,† Victoria Korolik,‡ Janna A. ten Brinke,† Maarten T. T. Pennings,‡ Robert Zalm,† Bart J. J. Teunis,† Peter J. Coloe‡ and Bernard A. M. van der Zeijst†‡

Author for correspondence: Bernard A. M. van der Zeijst. Tel: +31 30 274 41 64. Fax: +31 30 274 44 39. e-mail: ben.van.der.zeijst@rivm.nl

Most Campylobacter jejuni strains express lipo-oligosaccharides. Some strains also express lipopolysaccharides (LPS), with O-antigen-like carbohydrate repeats. C. jejuni 81116 expresses an LPS containing both lipo-oligosaccharides and O-antigen-like repeats, but nothing is known about the structure or sugar composition of these LPS species. A cosmid library of the genome of C. jejuni 81116 was constructed and probed with Campylobacter hyoilei genes involved in LPS synthesis. Five cosmids hybridized with the probe and two of these expressed C. jejuni 81116 LPS in Escherichia coli. By subcloning, a 16 kb DNA region was identified which contains the genetic information required to express C. jejuni LPS. DNA sequence analysis revealed 11 ORFs homologous to genes involved in LPS synthesis of other bacteria. They consisted of three homologues of sugar biosynthesis genes, two homologues of transport genes and six homologues of sugar transferases.

Keywords: Campylobacter jejuni, rfb gene cluster, lipo-oligosaccharide (LOS), lipopolysaccharide (LPS), O-antigen

INTRODUCTION

Campylobacter jejuni is an agent causing human enterocolitis and is the most common cause of bacterial diarrhoea in many countries (Tauxe, 1992). Symptoms most frequently seen are acute abdominal pain and inflammatory diarrhoea, often with fever (Butzler & Skirrow, 1979). C. jejuni is also a commensal in the intestine of birds (Blaser & Reller, 1981; Stern et al., 1988), which probably explains why it grows well at 42 °C, the body temperature of birds.

Lipopolysaccharides (LPS) are an abundant surface component of the outer membrane of Gram-negative bacteria. The LPS molecule consists of three distinct regions. Anchored in the outer membrane is the lipid A moiety, which is the endotoxic part of the LPS molecule.

Attached to the lipid A is the core, a branched-chain oligosaccharide linked to a ketodeoxy octulosonic acid (KDO) molecule, which has an inner and an outer part. Specific for the inner core is the presence of KDO and heptose residues, which are not found in the outer core. Extending from the cell surface is the O-antigen, a repeat of 10–30 oligosaccharides composed of 1–5 sugar residues (Schnaitman & Klena, 1993). Bacterial O-antigen protects the organism against complement and other serum components (Hackett et al., 1987; Joiner et al., 1986) and is involved in resistance to microbicidal human polymorphonuclear leukocytes (Stinavage et al., 1989).

A variant of the LPS molecule, seen in bacteria like Neisseria spp., Haemophilus spp. and Bordetella spp., is the lipo-oligosaccharide (LOS) molecule (Inzana et al., 1985; Jennings et al., 1980; Ray et al., 1991), which lacks the O-antigen polymer. C. jejuni strains synthesize LOS molecules. In addition, some strains also possess polysaccharide repeats (Preston & Penner, 1987). In this paper, the term LPS will be used to indicate both saccharide structures of Campylobacter (LOS or LPS containing additional polysaccharide repeats). The LPS molecules of Campylobacter have been shown to have endotoxic properties (Branquinho et al., 1983; Naess &
B. N. FRY and OTHERS

Fig. 1. Known structures of core molecules and O-antigen-like polysaccharides from Campylobacter jejuni strains. O:1 (Aspinall et al., 1993c); O:2 (Aspinall et al., 1993b); O:3 (Aspinall et al., 1995a); O:4 (Aspinall et al., 1993c); O:10 (Salloway et al., 1996); O:19 (Aspinall et al., 1994a, b); O:23 and O:36 (Aspinall et al., 1992a, 1993c). The serotypes are shown. ‘n’ indicates that the oligosaccharide is repeated, forming an O-antigen-like polysaccharide. ‘P’ denotes a phosphate group.

Hofstad, 1984). Furthermore they have been reported to be important as adhesins (McSweegan & Walker, 1986) and they may play a role in antigenic variation, as the bacteria have the ability to shift the LPS antigenic composition (Mills et al., 1992).

The sugar composition and structure of the core oligosaccharide from several C. jejuni strains, belonging to eight serotypes, have been analysed (Fig. 1; Aspinall et al., 1993b, c, 1994b, 1995a; Salloway et al., 1996). The presence of N-acetylmuramic acid (NeuNAc) is surprising, as it is not usually found in prokaryotes. These NeuNAc residues resemble gangliosides when attached by β-2,3 linkages to α-Gal (Aspinall et al., 1992b, 1993c). This molecular mimicry is thought to play a role in the neuropathological autoimmune diseases Guillain–Barré syndrome and Miller–Fisher syndrome (Salloway et al., 1996; Schwerer et al., 1995). In Neisseria and Haemophilus, sialylation of LPS plays a role in pathogenicity by enhancing serum resistance (Demarco de Hormaeche et al., 1991; Moxon & Maskell, 1992). The role of sialylation of the Campylobacter LPS in pathogenicity has not yet been determined.

For many bacteria, the LPS core structure is highly conserved within a genus (Jansson et al., 1981). The core oligosaccharide structure of C. jejuni however is highly variable, although the sugar composition is not very different between strains (Fig. 1). Limited variability has also been observed for the core oligosaccharide from Neisseria and Haemophilus (Mandrell & Apicella, 1993; Pavliak et al., 1993), although not to the extent seen in C. jejuni. This variability in the C. jejuni LPS core explains the ability of the Penner serotyping system (Penner, 1988; Penner & Hennessy, 1980) to discriminate between strains with LOS-type LPS.

The O-antigen repeats of C. jejuni serotypes O:3, O:19, O:23 and O:36 have been analysed so far (Aspinall et al., 1992a, 1994a, 1995a). They consist of only two or three sugar species (Fig. 1). The position of the core oligosaccharide to which the O-antigen is attached has not been determined; the inner core heprose-KDO region has been suggested (Aspinall et al., 1994a), but C. jejuni serotype O:3 synthesizes a polysaccharide, consisting of a disaccharide repeat of L-glycero-D-idopyranosyl-1 → 4Gal, which is not covalently linked to the LPS core (Aspinall et al., 1995a).

The metabolic pathways and enzymes required to synthesize the LPS molecules have not yet been characterized. Rapid progress in the study of LPS synthesis in other bacteria has been made by a genetic approach in combination with knowledge of the structure of LPS. Understanding LPS genetics for C. jejuni would allow more insight into the role of LPS in infection and colonization, through the ability to create specific mutants. Moreover, it would enable a better understanding of the mechanism of LPS synthesis in C. jejuni.

C. jejuni strain 81116 was originally isolated from a human water-borne outbreak of gastroenteritis (Palmer et al., 1983). In this paper we report the cloning of the LPS gene cluster, which is located on a 16 kb chromosomal DNA fragment, from this pathogenic strain. This region, when expressed in Escherichia coli, gives rise to an LPS which reacts with the C. jejuni Penner antiserum specific for serotype 6. The nucleotide sequence of this region was also determined.

METHODS

Bacterial strains and plasmids. The bacterial strains and plasmids used in this study are listed in Table 1. C. jejuni strains were grown under microaerophilic conditions on Skirrow agar medium (Skirrow, 1977) at 42 °C for 24 h. E. coli was grown in LB broth or agar for 16 h at 37 °C. Antibiotic concentrations used were as follows: kanamycin (Sigma), 30 μg ml⁻¹; tetracycline (Sigma), 20 μg ml⁻¹; ampicillin (Centrafarm), 100 μg ml⁻¹.

DNA manipulation. DNA isolations, restriction enzyme digestions and DNA ligations were performed as described by Ausubel et al. (1995). Restriction enzymes, alkaline phosphatase and a nested deletion kit were obtained from Pharmacia and used according to the manufacturer’s instructions. Plasmid DNA purification for sequencing reactions was performed by using the Flexiprep plasmid kit (Pharmacia) or the Qiagen plasmid kit (Qiagen). DNA fragments for subcloning and probes were isolated from agarose gels using the GeneClean II kit (Bio101).

Construction of the C. jejuni 81116 genomic library. High-molecular-mass genomic DNA of C. jejuni 81116 was partially cleaved using Sau3AI and fragments were size-fractionated by sucrose-gradient centrifugation (Sambrook et al., 1982). Fragments of 25–35 kb were selected and cloned into the unique BglII site in the kanamycin-resistance gene of a 21 kb alkaline-phosphatase-treated cosmid vector pLA2917 (Allen & Hanson, 1985). The ligated mixture (a ratio of two vector molecules to one chromosomal fragment) was then packaged (0.5 μg ligation mix to 50 μl packaging mix) with Packagene λ particle packaging mix (Promega) and the mixture was
transduced into the host *E. coli* HB101 as recommended by the manufacturers. Of the resulting transformants, 600 were taken to assemble the genomic library, ensuring that the complete *C. jejuni* genome of 1600 kb was present.

**Library screening by Southern blot analysis.** The library was plated onto LB plates and grown overnight at 37 °C. The bacteria were transferred to nylon membranes (Hybond-N; Amersham) and lysed on Whatman 3 MM paper saturated with 2× SSPE, 5% SDS. The DNA was fixed by microwaving for 2–5 min at 650 W. The filters were prehybridized for 2 h in hybridization mix (6× SSPE, 5% Denhardt’s, 0.5% SDS) with 0.1 μg herring sperm DNA ml⁻¹ at 60 °C. The insert from pBT9105 (pBR322 with an 11.8 kb EcoRV fragment from *Campylobacter bryoidei* RMIT-32A containing several genes showing homology with genes involved in LPS synthesis, including a *rfbB*-like gene; Korolik et al., 1997) was labelled with ³²P using a random-priming labelling kit (Amersham). After boiling the probe for 3 min, it was added to the filters in the hybridization mix and hybridized overnight at 60 °C. The filters were washed twice with 2× SSPE, 0.1% SDS for 15 min at 60 °C and exposed to Fuji RX film overnight.

**Subcloning of the C. jejuni LPS coding region.** Recombinant cosmid pBT9502 DNA was partially cleaved with *BglII*. The resulting fragments were subcloned into *Bam*HI-cleaved, alkaline-phosphatase-treated plasmid vector pBR322 (Bolivar et al., 1977) and transformed into *E. coli* HB101.

**DNA sequence analysis.** The sequence of the cloned DNA was determined by the dideoxy chain-termination method (Sanger et al., 1977) with an Automated laser fluorescent DNA sequencer (Pharmacia), the autoread sequencing kit using T7 DNA polymerase (Pharmacia) and fluorescein-labelled nucleotide primers (Pharmacia). PC/GENE 6.70 (Korn & Queen, 1984) was used to analyse nucleotide and amino acid sequences, which were compared to databases available at GenomeNet using the BLAST program (Altschul et al., 1990). The Macaw program (Lawrence et al., 1993) was used for multiple sequence alignment.

**LPS isolation, PAGE, silver staining and immunoblotting.** Cell envelopes were isolated following the procedure of Lugtenberg et al. (1975) and treated with proteinase K for 1 h at 56 °C. The isolated LPS was resolved by Tricine-SDS-PAGE (Lesse et al., 1987) and analysed by silver staining (Tsai et al., 1982) and Western blotting (Sambrook et al., 1982). Penner 6 antisera was used in immunoblotting diluted 1:1000 in PBS containing 5% skimmed milk and 0.3% Tween 20. Goat anti-rabbit alkaline-phosphatase-conjugated immunoglobulins (Promega) were used as the second antibody diluted 1:5000 in PBS containing 5% skimmed milk and 0.3% Tween 20. The bound phosphatase was visualized with NBT/BCIP (nitro blue tetrazolium/5-bromo-4-chloro-3-indolyl phosphate). The Multi-mark multi-coloured standard (Novel Experimental Technology) was used as a molecular-mass marker.

## RESULTS

**Isolation of the DNA region involved in LPS synthesis**

A genomic cosmid library of *C. jejuni* 81116 (Newell et al., 1985; Palmer et al., 1983) was constructed and screened by Southern hybridization with the *rfbB* gene.
B. N. FRY and OTHERS

Fig. 2. Restriction map of the genomic region of C. jejuni 81116 encoding the synthesis of O-antigen. The inserts of cosmids pBT9502, pBT9505 and pBT9506 are shown as bars. The insert of plasmid pBTLPS is also shown. Black bars indicate clones that express O-antigen in E. coli. C and B are abbreviations for Clai and BgIII respectively.

Fig. 3. LPS profiles in Tricine-SDS-PAGE. (a) Silver-stained samples. (b) Western blots reacted with Penner 6 antiserum. Lanes: 1, E. coli 0111; 2, C. jejuni 81116; 3, E. coli HB101(pBT9502); 4, E. coli HB101(pLA2917); 5, E. coli HB101(pBTLPS).

from C. hyoilei (Korolik et al., 1997). This gene is homologous to genes involved in transferring the first sugar of the O-antigen subunit to the lipid precursor undecaprenol phosphate. Five hybridizing clones were isolated. The cosmids contained in these strains were designated pBT9502–pBT9506 (Table 1). Restriction analysis showed that pBT9502, pBT9505 and pBT9506 contained overlapping inserts (Fig. 2). The three clones carrying these overlapping cosmids were tested for expression of C. jejuni LPS antigens by colony blotting. Both E. coli/pBT9502 and E. coli/pBT9506 reacted with Penner 6 antiserum in a Western blot.

LPS analysis

The silver-stained C. jejuni 81116 LPS showed one band of 6 kDa after separation by Tricine-SDS-PAGE. This size is in accordance with LPS isolated from other Campylobacter strains. Using the Tricine-SDS-PAGE system allows the use of protein markers to estimate the molecular mass of LPS components (Lesse et al., 1990), as opposed to the standard glycine-buffer system in which the LPS components do not migrate as a function of their molecular masses. Moreover, aggregates were observed for C. jejuni 81116 LPS when the standard glycine-SDS-PAGE method of Laemmli (1970) was used (data not shown); this was never seen using the Tricine-SDS-PAGE method.

The E. coli strains carrying recombinant cosmids showed different LPS profiles when stained with silver. In the control lane, E. coli/pLA2917, three bands around 5–5 kDa and two bands around 4 kDa were visible. E. coli/pBT9505 LPS demonstrated the same pattern (data not shown). E. coli/pBT9502 and E. coli/pBT9506 LPS showed identical patterns with additional bands of 4–5, 6–5, 7–5 and 8 kDa compared to E. coli/pLA2917. The LPS from E. coli/pBT9502 is shown in Fig. 3. The changed LPS pattern shows that the cloned DNA region is involved in LPS synthesis.

Western blotting analysis using Penner 6 antiserum demonstrated that C. jejuni 81116 LPS is of the smooth type, with a high-molecular-mass O-antigen-specific ladder pattern between 6 and 40 kDa. LPS isolated from E. coli/pBT9502 and E. coli/pBT9506 (not shown) gave identical profiles, i.e. an O-antigen ladder pattern with immunoreactive bands between 10 and 45 kDa. This complex appeared to have a higher molecular mass than that of C. jejuni 81116 LPS. This may be explained by the fact that the lipid A-core complex of E. coli is used as a carrier for the C. jejuni O-antigen subunits. As expected, the C. jejuni O-antigen expressed in E. coli does not stain with silver. LPS prepared from E. coli/pBT9505 did not react with Penner 6 antiserum (data not shown).

Further definition of the LPS-expressing DNA region

The smallest insert still able to encode the complete ladder structure reacting with Penner 6 antiserum was 16 kb; it was subcloned into the pBR322 plasmid vector. The plasmid containing this insert was named pBTLPS (Figs 2 and 3). To exclude possible complementation of the E. coli HB101 LPS genes by the C. jejuni genes, pBT9502 and pBTLPS were also expressed in E. coli Sp874, a strain lacking all genes involved in O-antigen synthesis. LPS isolated from these Sp874 strains showed the same O-antigen ladder pattern reacting with Penner 6 antiserum (data not shown).

DNA sequence analysis

Fragments of the insert of plasmid pBTLPS, generated by the restriction endonucleases BgIII, Clai, HindIII and Sau3AI, were subcloned into sequencing vectors. Nested deletions and subcloning of these fragments allowed the
sequencing of the complete 16 kb region. To demonstrate continuity of adjacent but non-overlapping fragments, specifically primed sequences were generated.

Sequence analysis of the insert from pBTLPS showed that it contains 13 complete and two incomplete ORFs (Fig. 4; Table 2). Twelve ORFs showed homology to genes involved in LPS synthesis (Table 2) and were named galE and wlab-wlam (wla genes), according to the newly proposed bacterial polysaccharide gene nomenclature (Reeves et al., 1996). The remaining complete ORF was that of the chemotaxis gene cheY from C. jejuni, which is not thought to be involved in LPS synthesis. The incomplete ORF at the 5' terminus was similar to waaC genes and was therefore named waaC. At the 3' end, an incomplete ORF without homology was found. waaC is transcribed in the opposite direction to the other genes (Fig. 4). Not all the genes start with ATG codons preceded by typical Shine–Dalgarno sequences. wlab and cheY have the putative start codon GUG (Table 2); to our knowledge, this is the first time this start codon has been suggested for a Campylobacter gene. As only a few Campylobacter genes have been identified, not much is known about the promoter consensus sequences. Using the E. coli σ70 promoter consensus sequence, weak putative promoter sites were recognized along the sequence, which were located upstream of the waaC, the galE and the cheY genes. galE and wlab-wlam are very closely spaced, with most of the designated ORFs showing overlap (Table 2). The largest spacing between two ORFs (97 bp) was found between genes wlab and wlaK. No putative promoter sequences were found within this spacing. All these genes have the same transcriptional direction and therefore it is likely they are transcribed as a single operon. No putative rho-independent terminator-like hairpin loops were found. All ORFs possessed a G + C content between 28 and 34 mol% (Table 2), which is consistent with the low G + C content of Campylobacter spp. of 30 mol%.

The characteristics of the genes within the analysed region will be discussed in the order in which they are present, starting with waaC. The BLAST algorithm (Altschul et al., 1990) was used to find proteins homologous to the deduced gene products (Table 3).

**waaC.** The deduced partial protein sequence is homologous to the N-terminal regions of LPS heptosyltransferases -I, which transfer a heptose residue to the KDO molecule, initiating the synthesis of the inner core (Ray & Raetz, 1987). At least four other genes adjacent to waaC are similar to genes involved in inner core synthesis (data not shown). Therefore, waaC may be the first of a gene cluster involved in inner core synthesis.
Table 3. Deduced proteins present in the C. jejuni wla region and their homologues

<table>
<thead>
<tr>
<th>C. jejuni protein</th>
<th>Similar protein</th>
<th>Organism</th>
<th>Function</th>
<th>Identity/similarity (%)*</th>
<th>Accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>WaaC†</td>
<td>RfA</td>
<td>E. coli</td>
<td>Heptosyltransferase I</td>
<td>28.6/51.2</td>
<td>P24173</td>
</tr>
<tr>
<td></td>
<td>RfA</td>
<td>N. gonorrhoeae</td>
<td>Heptosyltransferase I</td>
<td>28.6/51.2</td>
<td>S60783</td>
</tr>
<tr>
<td></td>
<td>RfA</td>
<td>B. pertussis</td>
<td>Heptosyltransferase I</td>
<td>28.6/48.8</td>
<td>X90711</td>
</tr>
<tr>
<td></td>
<td>RfA</td>
<td>S. typhimurium</td>
<td>Heptosyltransferase I</td>
<td>26.2/52.4</td>
<td>P26469</td>
</tr>
<tr>
<td>GalE</td>
<td>GalE</td>
<td>H. influenzae</td>
<td>UDPglucose 4-epimerase</td>
<td>37.5/52.4</td>
<td>A64063</td>
</tr>
<tr>
<td>GalE</td>
<td>GalE</td>
<td>N. meningitidis</td>
<td>UDPglucose 4-epimerase</td>
<td>36.3/53.1</td>
<td>S39638</td>
</tr>
<tr>
<td></td>
<td>GalE</td>
<td>E. coli</td>
<td>UDPglucose 4-epimerase</td>
<td>36.0/52.2</td>
<td>P09147</td>
</tr>
<tr>
<td>WlaB</td>
<td>HyaA</td>
<td>A. baena sp.</td>
<td>Polysaccharide involvement</td>
<td>30.1/48.7</td>
<td>P22638</td>
</tr>
<tr>
<td></td>
<td>HlyB</td>
<td>E. coli</td>
<td>Export haemolysin</td>
<td>30.1/46.2</td>
<td>P08716</td>
</tr>
<tr>
<td></td>
<td>HlyB</td>
<td>A. pleuropneumoniae</td>
<td>Export haemolysin</td>
<td>27.5/54.4</td>
<td>P26760</td>
</tr>
<tr>
<td></td>
<td>M8bA</td>
<td>E. coli</td>
<td>Export†</td>
<td>24.3/43.3</td>
<td>P27299</td>
</tr>
<tr>
<td></td>
<td>ValA</td>
<td>F. novicida</td>
<td>Secretion LPS</td>
<td>23.4/41.5</td>
<td>L17003</td>
</tr>
<tr>
<td>WlaC</td>
<td>TrsD</td>
<td>Y. enterocolitica</td>
<td>Glycosyltransferase†</td>
<td>28.8/47.0</td>
<td>S51263</td>
</tr>
<tr>
<td></td>
<td>AmsD</td>
<td>E. amylovora</td>
<td>Glycosyltransferase†</td>
<td>27.4/44.8</td>
<td>S52144</td>
</tr>
<tr>
<td></td>
<td>EpsG</td>
<td>S. thermophilus</td>
<td>Glycosyltransferase*</td>
<td>26.8/40.8</td>
<td>U40830</td>
</tr>
<tr>
<td>WlaD</td>
<td>Orf87</td>
<td>Y. pseudotuberculosis</td>
<td>Aboequosyltransferase*</td>
<td>27.2/38.9</td>
<td>L01777</td>
</tr>
<tr>
<td></td>
<td>TrsB</td>
<td>Y. enterocolitica</td>
<td>Glycosyltransferase†</td>
<td>23.3/42.4</td>
<td>S51261</td>
</tr>
<tr>
<td></td>
<td>EpsI</td>
<td>S. thermophilus</td>
<td>Glycosyltransferase†</td>
<td>230.3/38.8</td>
<td>U40830</td>
</tr>
<tr>
<td></td>
<td>SpsA</td>
<td>B. subtilis</td>
<td>Polysaccharide synthesis</td>
<td>22.2/41.0</td>
<td>P39621</td>
</tr>
<tr>
<td>WlaE</td>
<td>AmsD</td>
<td>E. amylovora</td>
<td>Glycosyltransferase†</td>
<td>25.1/41.3</td>
<td>S52144</td>
</tr>
<tr>
<td></td>
<td>TrsD</td>
<td>Y. enterocolitica</td>
<td>Glycosyltransferase†</td>
<td>24.9/40.5</td>
<td>S51263</td>
</tr>
<tr>
<td></td>
<td>TrsE</td>
<td>Y. enterocolitica</td>
<td>Glycosyltransferase†</td>
<td>25.6/40.1</td>
<td>S51264</td>
</tr>
<tr>
<td></td>
<td>CapH</td>
<td>S. aureus</td>
<td>Glycosyltransferase†</td>
<td>23.1/37.8</td>
<td>U10927</td>
</tr>
<tr>
<td>WlaF</td>
<td>STT-3</td>
<td>C. elegans</td>
<td>Oligosaccharyltransferase</td>
<td>19.5/35.9</td>
<td>P46975</td>
</tr>
<tr>
<td></td>
<td>Wzy (RfC)</td>
<td>E. coli</td>
<td>O-antigen polymerase</td>
<td>18.3/39.7</td>
<td>P37748</td>
</tr>
<tr>
<td></td>
<td>STT3</td>
<td>S. cerevisiae</td>
<td>Oligosaccharyltransferase</td>
<td>16.3/34.5</td>
<td>JC345</td>
</tr>
<tr>
<td>WlaG</td>
<td>CpsF</td>
<td>P. mirabilis</td>
<td>Glycosyltransferase†</td>
<td>30.8/48.1</td>
<td>L36873</td>
</tr>
<tr>
<td></td>
<td>RfbF</td>
<td>S. marcescens</td>
<td>Galactantransferase†</td>
<td>27.1/41.2</td>
<td>L34167</td>
</tr>
<tr>
<td></td>
<td>RfbP</td>
<td>S. dysenteriae</td>
<td>Galactosyltransferase</td>
<td>26.6/44.7</td>
<td>S27671</td>
</tr>
<tr>
<td></td>
<td>RfbE</td>
<td>K. pneumoniae</td>
<td>Galactantransferase</td>
<td>26.1/41.5</td>
<td>S31762</td>
</tr>
<tr>
<td>WlaH</td>
<td>WbaP (RfP)</td>
<td>S. enterica</td>
<td>Galactosyltransferase first step</td>
<td>39.5/55.5</td>
<td>P26406</td>
</tr>
<tr>
<td></td>
<td>WbaP (RfP)</td>
<td>H. influenzae</td>
<td>Galactosyltransferase first step</td>
<td>38.5/53.0</td>
<td>B64099</td>
</tr>
<tr>
<td></td>
<td>EpsE</td>
<td>S. thermophilus</td>
<td>Galactosyltransferase first step</td>
<td>36.0/33.0</td>
<td>U40830</td>
</tr>
<tr>
<td></td>
<td>ExoY</td>
<td>R. sphaeroides</td>
<td>Galactosyltransferase first step†</td>
<td>36.0/52.0</td>
<td>X16704</td>
</tr>
<tr>
<td>WlaI</td>
<td>NeuD</td>
<td>E. coli</td>
<td>Polysialic acid capsule synthesis</td>
<td>28.1/45.8</td>
<td>U05248</td>
</tr>
<tr>
<td></td>
<td>LpxD</td>
<td>S. enterica</td>
<td>Acetyltransferase</td>
<td>26.6/44.3</td>
<td>P18482</td>
</tr>
<tr>
<td></td>
<td>LpxD</td>
<td>Y. enterocolitica</td>
<td>Acetyltransferase</td>
<td>26.1/42.0</td>
<td>P32203</td>
</tr>
<tr>
<td></td>
<td>LpxD</td>
<td>E. coli</td>
<td>Acetyltransferase</td>
<td>25.1/43.8</td>
<td>P21645</td>
</tr>
<tr>
<td>WlaK</td>
<td>RfE</td>
<td>V. cholerae</td>
<td>Perosamine synthetase†</td>
<td>32.2/51.2</td>
<td>S28471</td>
</tr>
<tr>
<td></td>
<td>DegT</td>
<td>B. steatorrhombus</td>
<td>Transamination</td>
<td>29.0/44.4</td>
<td>P15263</td>
</tr>
<tr>
<td></td>
<td>FlaA2</td>
<td>C. crescentus</td>
<td>Monosaccharide synthetase</td>
<td>25.9/44.0</td>
<td>U27301</td>
</tr>
<tr>
<td></td>
<td>SpsC</td>
<td>B. subtilis</td>
<td>Polysaccharide synthetase</td>
<td>25.4/33.0</td>
<td>P39623</td>
</tr>
<tr>
<td>WlaL</td>
<td>CapD</td>
<td>S. aureus</td>
<td>Capsule synthesis</td>
<td>38.5/55.6</td>
<td>P39853</td>
</tr>
<tr>
<td></td>
<td>TrsG</td>
<td>Y. enterocolitica</td>
<td>Acetylgalactosamine synthetase†</td>
<td>37.5/51.7</td>
<td>S51266</td>
</tr>
<tr>
<td></td>
<td>WbaP</td>
<td>S. enterica</td>
<td>LPS B-band synthesis</td>
<td>34.4/53.2</td>
<td>U44089</td>
</tr>
<tr>
<td>WlaM</td>
<td>AecB</td>
<td>V. cholerae</td>
<td>Accessory colonization factor</td>
<td>26.9/48.2</td>
<td>L25660</td>
</tr>
<tr>
<td>CheY</td>
<td>CheY</td>
<td>C. jejuni</td>
<td>Chemotaxis regulator</td>
<td>100</td>
<td>U75208</td>
</tr>
<tr>
<td></td>
<td>CheY</td>
<td>P. aeruginosa</td>
<td>Chemotaxis regulator</td>
<td>48.4/65.3</td>
<td>S20545</td>
</tr>
<tr>
<td></td>
<td>CheY</td>
<td>E. coli</td>
<td>Chemotaxis regulator</td>
<td>453/57.8</td>
<td>P06143</td>
</tr>
<tr>
<td></td>
<td>CheY</td>
<td>S. typhimurium</td>
<td>Chemotaxis regulator</td>
<td>453/57.8</td>
<td>P06657</td>
</tr>
</tbody>
</table>

* The program PALIGN from PCGene (Korn & Queen, 1984) was used to compare the deduced gene products. Open gap cost: 2. Percentage similarities represent the sum of identical amino acids (%) plus conserved substitutions (%).
† Incomplete.
‡ Putative function.

**galE.** The protein encoded by galE is homologous to UDPgalactose 4-epimerase (GalE), an enzyme catalysing the interconversion between UDPGlc and UDPGal. Gal is a component of the LPS core structure of *E. coli* and *S. typhimurium*. Defects in galE result in a rough-LPS phenotype (Rick, 1987). In *Neisseria meningitidis*,
**Campylobacter jejuni** lipopolysaccharide genes

**WlaB-Cj**

KYLFKNLNLHKEKRIAPTEGKKAEDKGTWVDLILGLLKKPEISQHEKQGRQELNASNAYQKGYPEHONYLFNSDIAKQ

**HetA-An**

NLYLHNHITRKKQATLLVWNLPSADWLLADLPRYPDTEQGQVLIDLVQYFBELARKMAVSTDTFTMTTSRIFA

**ValA-Fn**

HKVLGOVSVDPLDKTVAFVQTLSSRLWTIGMRTPYQKIIAIQVHDWTRLILELHLSILTVSWSVLFHYDTDVTW

**MsbB-Ec**

VPALRNLQKRLAHTKIALRRPVYSLRQYTVNSHLRQIYALVARLVQGWYVWNLNHTDVAQP

**HlyB-Ec**

PVILLNLSLQEGGTVIGLGVQTVKSKLQRTQYIPENQVSITSCNDLIALADPNWRQGVVULCVNLNNRISIDQ

**HlyB-Ac**

PVILLNLSLQEGGTVIGLGVQTVKSKLQRTQYIPENQVSITSCNDLIALADPNWRQGVVULCVNLNNRISIDQ

**Fig. 5.** Alignment of the C-terminal part of WlaB with homologous proteins. Identical amino acids within all six sequences are highlighted. The ATP-binding site motif and the ABC-transporter family signature (underlined) are presented below the alignment. When multiple amino acids are possible in one of the positions of the ABC-transporter family signature, numbers are inserted: 1 (L, I, V, M, F or Y); 2 (S, A or G); 3 (R, K or A); 4 (L, I, V, M, Y or A); 5 (L, I, V, M or F); 6 (5, A or G).

_C. jejuni_, a mutated _galE_ gene resulted in truncation of the LOS chain by deletion of the outer core (Hammer-schmidt _et al._, 1994). The N-terminal region of the predicted GalE protein possesses the consensus NAD-binding domain (GxxGxxG; Macpherson _et al._, 1994) commonly present in epimerases. All known Campylobacter LPS structures contain both Gal and Glc; therefore it is likely that GalE functions as a UDP-galactose 4-epimerase.

**wlaB.** The protein product of _wlaB_ contains an ATP-binding site and the signature for ABC transporters (Higgins _et al._, 1990; Walker _et al._, 1982), suggesting that this protein is involved in transport across the cytoplasmic membrane. The protein contains six potential transmembrane segments, which are also seen in other transmembrane transport molecules (Blight & Holland, 1990). Homology to other transport proteins is mainly located in the C-terminal part, where the ATP-binding site and the ABC transporters family motif are located (Fig. 5). WlaB is also homologous to mammalian multidrug resistance proteins (P-proteins), which bind to and export a variety of drugs, thereby conferring resistance (Endicott & Ling, 1989). Comparison of the hydrophobicity plots of these proteins indicated that the six transmembrane segments were located at identical positions (data not shown). WlaB probably has a signal peptide which is cleaved between amino acids 45 and 46. The similarities described above suggest that the WlaB protein from _C. jejuni_ is involved in the transport of LPS molecules across the cytoplasmic membrane.

**wlaC, wlaD and wlaE.** The protein products of these three genes all possess one or two putative transmembrane segments. All three deduced proteins are similar to putative glycosyltransferases. TrsD and TrsE from _Y. enterocolitica_, showing homology to WlaC and WlaE, are putative galactosyl- or N-acetylgalactosaminyl-transferases involved in outer core synthesis (Skurnik _et al._, 1995). WlaC and WlaE are closely related, with 30.2% identical amino acids and 43.6% similar amino acids. WlaC and WlaE may link Gal or GalNAc to the growing polysaccharide chain.

**wlaF.** _wlaF_ is the largest gene found in the cluster, encoding a predicted protein of 822 kDa (Table 2). This protein is similar to the transmembrane oligosaccharyl-transferases (STT-3) from eukaryotes and the O-antigen polymerase from _E. coli_. The STT3 protein from yeast is possibly involved in the transfer of an oligosaccharide from a lipid precursor to a protein (Yoshida _et al._, 1995). The amino acid sequence of WlaF revealed a putative signal peptide, which is most probably cleaved between amino acids 36 and 37, as also seen for the STT-3 protein. The N-terminal part of WlaF contains 12 putative transmembrane segments. The C-terminal part is hydrophilic. These latter two characteristics are also seen in the STT-3 protein. The Rfc polymerase is often also highly hydrophobic, containing 12 putative trans-
Fig. 6. Hydrophobicity plots of WlaH and homologous proteins (Table 2). The plots were made using SOAP from PCGENE (Korn and Queen, 1984) using an interval of 15 amino acids. Se, S. enterica; Hi, Haemophilus influenzae.

membrane segments. We speculate that the C. jejuni WlaF protein is involved in the transfer of oligosaccharides, possibly from a lipid carrier onto the lipid A-core complex. However, it may also have flipase activity. Wzx proteins are also hydrophobic and their genes are often found in LPS gene clusters (Reeves, 1994).

wlaG. Throughout the whole gene, wlaG is homologous to genes encoding galactosyltransferases. The rfbF gene products of Klebsiella pneumoniae and Serratia marcescens, transfer the disaccharide Gal(1→3)Gal, representing a single repeating unit of β-galactan I, to the GlcNAC-containing lipid intermediate (Clarke et al., 1995; Szabo et al., 1995). The Rpb protein of Shigella dysenteriae transfers a single Gal residue to a GlcNAC-containing lipid (Gohmann et al., 1994). Unlike the Rpb protein of S. dysenteriae and the CpsF protein of Proteus mirabilis (Gygi et al., 1995), the predicted WlaG protein of C. jejuni is probably not membrane bound, as it has no significant hydrophobic regions. However, the RfbF protein of Klebsiella pneumoniae is also not membrane bound. As at least two galactose residues are present in the LPS molecules of all C. jejuni strains analysed, we suggest that the WlaG protein functions as a galactosyltransferase.

wlaH. The wlaH gene is homologous to genes involved in transferring the first sugar of the subunits of O-antigen to the lipid precursor undecaprenol phosphate. The present nomenclature for this type of gene is wbaP. This first sugar differs between bacterial species and strains, but in most cases Gal is the first sugar with which O-antigen synthesis is initiated. The proteins showing highest similarity with WlaH transfer a Gal residue. Like other WbaP proteins, WlaH contains a hydrophobic domain required for the interaction with the undecaprenol phosphate lipid carrier (Fig. 6). WbaP-like proteins have been shown to be bifunctional, with the above-mentioned transferase activity assigned to the C-terminal domain. Recently it has been proposed that the N-terminal domain of the WbaP protein of Salmonella enterica is involved in the release of undecaprenol-pyrophosphate-linked galactose from WbaP (Wang et al., 1996) and does not have a flipase function as earlier reported (Wang & Reeves, 1994). From the protein alignment it can be concluded that WlaH from C. jejuni probably only possesses the transferase activity, as the N-terminal domain is missing, compared to the WbaP proteins of S. enterica and Haemophilus influenzae (Fig. 6).

wlaI. This ORF, encoding a protein of 203 amino acids, is most similar to the 207 amino acid NeuD protein from E. coli, an enzyme involved in the synthesis of the K1 polysaccharide (Annunziato et al., 1995), an α-2,8-linked linear polymer of about 200 NeuNAC residues. The NeuD protein is suggested to be involved in NeuNAC transfer (Annunziato et al., 1995). The C-terminal part of the WlaI protein is also similar, although to a lesser extent, to the 340 amino acid UDP-N-acetylgalcosamine acyltransferases, LpxD, of E. coli, S. enterica and Y. enterocolitica, which play a role in lipid A synthesis (Dicker & Seetharam, 1991; Hirvas et al., 1990; Vuorio et al., 1994). LpxD-like genes have been found in several gene clusters involved in LPS O-antigen synthesis. The BpB protein from Bordetella pertussis (Allen & Maskell, 1996) and the WbpD protein from Pseudomonas aeruginosa (Burrows et al., 1996) are both acetyltransferases. The WlaI deduced protein sequence also shows homology to this group of transferases.

A hexapeptide motif, also termed the isoleucine patch (Dicker & Seetharam, 1992), is present in some transferases (Vaara, 1992). Each motif starts with an isoleucine, leucine or valine residue and often contains a glycine in the second position. In the WlaI protein of C. jejuni, as well as in all the homologous proteins described above, this repeating motif is present (Fig. 7). We propose a function for the WlaI protein as a NeuNAC transferase, since the NeuD protein from E. coli shows the highest similarity. The NeuD protein, like the WlaI protein, is small, containing only 207 amino acids,
whereas all known LpxD proteins contain around 340 amino acids. The presence of NeuNAc residues in seven of the eight analysed C. jejuni LPS molecules indicates the need of such an enzyme activity.

\textbf{wlaK}. The predicted protein designated \textit{WlaK} shows the highest similarity with the \textit{RfbE} protein from \textit{Vibrio cholerae}, which encodes a putative perosamine synthetase (Manning et al., 1995). Homology with \textit{DegT} from \textit{Bacillus stearothermophilus} is also found. This protein is required in the pathway to synthesize 2,6-, 3,6- and 4,6-dideoxyhexoses and is thought to be an enzyme for transaminations leading to amino sugars (Thorson et al., 1993). The \textit{SpsC} protein from \textit{Bacillus subtilis} (Glaser et al., 1993), also belonging to the \textit{DegT} family, and involved in spore coat polysaccharide biosynthesis, is also homologous to the \textit{C. jejuni} \textit{WlaK} protein. The \textit{flaA2} gene product from \textit{Caulobacter crescentus}, involved in flagellin synthesis (B. Ely, personal communication) is also similar to \textit{WlaK}. It is thought that polysaccharide may be involved in \textit{C. crescentus} flagellum biogenesis (B. Ely, personal communication). The \textit{WlaK} protein of \textit{C. jejuni} is probably also involved in the synthesis of an amino sugar, possibly GalNAc or NeuNAc, which are both common components of the \textit{Campylobacter} LPS molecule. \textit{N. gonorrhoeae}, however, uses exogenous NeuNAc to sialylate its LOS by an outer-membrane-bound NeuNAc transferase which has been detected in the outer membrane (Mandrell et al., 1993). If exogenous NeuNAc is also used by \textit{C. jejuni}, no enzymes for the synthesis of NeuNAc would be needed.

\textbf{wlaL}. The \textit{WlaL} protein is homologous throughout its entire length to the \textit{CapD}, \textit{WbpM} and \textit{TrsG} proteins from \textit{Staphylococcus aureus}, \textit{P. aeruginosa} and \textit{Y. enterocolitica}, respectively (Burrows et al., 1996; Lin et al., 1994; Skurnik et al., 1995). All proteins contain five hydrophobic domains in the N-terminal part that are predicted to be transmembrane segments. The central part contains an NAD-binding site and is homologous to UDPglucose 4-epimerases. The function of the \textit{CapD} protein is not known. The \textit{TrsG} protein is thought to be involved in the biosynthesis of GalNAc or FucNAc (acetylfucosamine), the \textit{WbpM} protein is possibly a dehydrogenase or epimerase needed for the biosynthesis of a 2-acetamido-2,6-dideoxy-\alpha-galactose residue. A putative signal peptide has been found in \textit{WlaL}, which is probably cleaved at position 29. We suggest the \textit{WlaL} protein is involved in the synthesis of an amino sugar, probably GalNAc or NeuNAc.

\textbf{wlaM}. The \textit{wlaM} gene product shows only low similarity to an accessory colonization factor, \textit{AcfB} (Everiss et al., 1994). This protein contains the methyl-accepting chemotaxis motif as found in chemotaxis genes (Alley et al., 1992; Dahl et al., 1989). However the similarity with \textit{WlaM} was confined to the N-terminal part not containing this motif. Like \textit{WlaF}, \textit{WlaM} is a hydrophobic protein containing six putative transmembrane regions. Therefore \textit{WlaM} could also function as a flippase or a polymerase.

\textbf{cheY}. This gene is identical to the \textit{cheY} gene of \textit{C. jejuni} (Ketley, 1997; Yao et al., 1997), a gene involved in chemotaxis. It is unlikely that this gene is an LPS synthesis gene. Furthermore, unlike \textit{galE} and \textit{wla}, the \textit{cheY} gene has its own promoter consensus sequence.

\textbf{orf1}. No homology was found for this incomplete ORF or its putative protein product.

\textbf{DISCUSSION}

We have shown that a 16 kb chromosomal DNA region of \textit{Campylobacter jejuni} strain 81116 is able to express LPS in \textit{E. coli}. These heat-stable antigens are known to react with Penner 6 antiserum. The sequence analysis of this region is presented.

\textbf{The wla cluster}

The sequence data show that the 16 kb region contains 11 genes, of which the deduced protein sequences show homology to proteins involved in LPS synthesis in other bacteria. On the basis of the observed similarity of the deduced \textit{Wla} proteins to proteins involved in LOS synthesis, and especially the outer-core synthesis of \textit{Y. enterocolitica} and \textit{Bordetella pertussis} (Allen & Maskell, 1996; Skurnik et al., 1995), it is likely that the \textit{wla} gene cluster of \textit{C. jejuni} is involved in the synthesis of the O-antigen-like polysaccharide, as well as the outer core, but not of the inner core. This is in agreement with the finding of the \textit{waatC} gene, situated at the beginning of a separate gene cluster of at least five genes involved in inner core synthesis (data not shown).

The presence of the \textit{wlaH} gene, which probably encodes a galactosyltransferase, linking a Gal residue to the lipid carrier, and the observation that almost all analysed outer core molecules from \textit{C. jejuni} start with a Gal residue, suggests that the outer core molecule is synthesized on a lipid carrier. This has also been suggested for the outer core of \textit{Y. enterocolitica}, where also a Gal-lipid carrier transferase gene was found to be present in the outer-core-synthesis gene cluster (Skurnik et al., 1995). Both the O-antigen-like polysaccharide and the outer core may be synthesized on lipid carriers, in which case the \textit{WlaH} protein would catalyse the first step.

\textbf{Sugar biosynthesis}

Three genes involved in the biosynthesis of sugars are present in the \textit{wla} cluster. The presence of \textit{galE}, converting Glc to Gal, is not surprising, as all known outer core molecules of \textit{C. jejuni} contain Gal residues. The \textit{wlaK} and \textit{wlaL} genes are probably involved in the synthesis of amino sugars. GalNAc and NeuNAc are amino sugars found in core molecules of several \textit{C. jejuni} strains.

\textbf{Transport of sugar structures}

Three transport steps have been observed during LPS synthesis of \textit{E. coli} and \textit{Salmonella typhimurium}: flipping of an oligosaccharide subunit attached to a lipid carrier across the inner membrane; transport of saccharide structures from a lipid carrier to the
lipid A–core complex; and transport of the complete LPS molecule from the inner membrane to the outer membrane.

We assume that the outer-core molecule of C. jejuni is synthesized at the cytoplasmic side of the inner membrane, like an LPS O-antigen subunit. After completion it has to be ‘flipped’ to the periplasmic side of the inner membrane. This could happen before or after it has been ligated to the lipid A–inner core complex. Possibly WlaM is involved in this flippase process and WlaF synthesized at the cytoplasmic side of the inner membrane. This could happen before or after it has been ligated to the lipid A–inner core complex. Possibly WlaM is involved in this flippase process and WlaF could be used to link the outer core to the lipid A–inner core complex, like an LPS O-antigen subunit. After completion this could happen before or after it has been ligated to the lipid A–inner core complex. Possibly WlaM is involved in this flippase process and WlaF could be used to link the outer core to the lipid A–inner core complex, or vice versa. The complete LOS molecule is then translocated across the inner membrane to the outer membrane by WlaB.

Sugar transferases

The proposed sugar transferases within the wla gene cluster consist of Gal or GalNAc transferases WlaC and WlaE, an unknown transferase WlaD, a Gal transferase WlaG, a Gal-to-lipid-carrier transferase WlaH and a NeuNAc transferase WlaI. If, as suggested above, the outer core is synthesized on a lipid carrier, WlaH probably initiates the synthesis of the outer core by transferring a Gal residue onto the lipid carrier.

All but one of the known core molecules of C. jejuni contain a GalNAc residue bound to a backbone Gal residue. The proposed function of WlaC and WlaE as Gal transferases would predict the presence of a GalNAc residue in the outer core of strain 81116. Most Gal residues in the outer core molecules of C. jejuni are branched with a NeuNAc residue, for which WlaI is the proposed transferase.

O-antigen-like polysaccharide structure and synthesis

Three observations suggest that the O-antigen-like repeat of C. jejuni is not linked to the lipid A–inner core complex and therefore is not a proper O-antigen. (1) Significant amounts of lipid A–inner core residues were not found in any of the analysed O-antigen-like repeats (Aspinall et al., 1993b, c, 1994b, 1995a; Salloway et al., 1996). (2) The O-antigen-like repeat unit is not stained by the silver-staining method, whereas the lipid A–core complex is. Thus, if the repeat unit were linked to the lipid A–core complex it would surely be seen using this technique. (3) Expression of the wla cluster in E. coli shows a lipid A–molecule larger than that of Campylobacter, which reacts with the serospecific antiserum of C. jejuni 81116. This suggests that the lipid A–core complex of E. coli is used, to which a specific C. jejuni sugar structure is linked. The molecular mass of the expressed O-antigen-like repeat, however, is not increased. The O-antigen-like repeat therefore resembles capsular polysaccharide or enterobacterial common antigen as found in E. coli and not LPS O-antigen.

Three variable O-antigen-like polysaccharides from different Campylobacter strains have been examined. Two consist of a disaccharide and one of a trisaccharide repeat. In two of these repeat units, a heptose derivative is present, as also found in Campylobacter coli O:30 (Aspinall et al., 1993a) and in strains of Campylobacter lari (Aspinall et al., 1993b, 1995c). In all three analysed repeat units, an amino sugar is present; in two of these it is a GlcNAc residue.

We predict that the polysaccharide synthesis follows the same route as the start of the outer-core assembly, namely on a lipid carrier. The first sugar, a Gal residue, will be linked by WlaH. After this, the synthesis of the outer core and the O-antigen-like polysaccharide would diverge.

When comparing the known O-antigen-like polysaccharide structures with the core structures of C. jejuni, a remarkable difference is seen. The Gal residues in the repeat units are not sialylated, whereas all Gal residues found in the backbone of the outer core molecule are. It could therefore be reasoned that the outer core oligosaccharide is sialylated in a compartment where the O-antigen-like polysaccharide is not present. Sialylation may occur on the periplasmic side of the inner membrane or in the outer membrane, where the O-antigen-like polysaccharide never goes.

The characterization of this first wla gene cluster of C. jejuni marks a starting point for further studies and applications, possibly in serotyping, diagnostics and vaccine development. Also the study of LPS-induced diseases can be advanced.

Insertional-inactivation experiments in C. jejuni and in the cloned wla region in E. coli are needed to confirm the functions assigned to the various proteins. In addition, attempts to complement define S. typhimurium LPS mutants are in progress. The LOS structure of C. jejuni 81116 is currently being analysed by Dr A. P. Moran and coworkers in Galway, Ireland.

ACKNOWLEDGEMENTS

We are grateful to Dr D. G. Newell for her gift of the Penner 6 antiserum and to Dr J. M. Ketley for his help with designing Fig. 1. We thank Dr P. R. Reeves for his help with the gene nomenclature. We thank A. J. A. M. van Asten, L. Heijmen-van Dijk and K. A. Zwaagstra for technical assistance with the automated sequencing.

REFERENCES


Received 10 October 1997; revised 13 March 1998; accepted 2 April 1998.