Genetic basis for expression of the major globotetraose containing lipopolysaccharide from Haemophilus influenzae strain Rd (RM118)

Hood, Derek W.; Cox, Andrew D.; Wakarchuk, Warren W.; Schur, Melissa; Schweda, Elke K. H.; Walsh, Shannon L.; Deadman, Mary E.; Martin, Adele; Moxon, E. Richard; Richards, James C.

Publisher's version / Version de l'éditeur:
https://doi.org/10.1093/glycob/11.11.957
Glycobiology, 11, 11, pp. 957-967, 2001-11-01

Glycobiology vol. 11 no. 11 pp. 957–967, 2001

Genetic basis for expression of the major globotetraose-containing lipopolysaccharide from *H. influenzae* strain Rd (RM118)

Derek W. Hood[1,2], Andrew D. Cox[3], Warren W. Wakarchuk[3], Melissa Schur[3], Elke K.H. Schweda[4], Shannon L. Walsh[2], Mary E. Deadman[2], Adele Martin[3], E. Richard Moxon[2], and James C. Richards[3]

[2]Molecular Infectious Diseases Group, University of Oxford Department of Paediatrics, Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, Headington, Oxford, OX3 9DS, UK; [3]Institute for Biological Sciences, National Research Council of Canada, Ottawa, Ontario, K1A 0R6 Canada; and [4]Clinical Research Centre, Karolinska Institutet and University College of South Stockholm, NOVUM, S-141 86 Huddinge, Sweden.

Received on April 19, 2001; revised on July 9, 2001; accepted on July 11, 2001

A genetic basis for the biosynthetic assembly of the globotetraose containing lipopolysaccharide (LPS) of *Haemophilus influenzae* strain RM118 (Rd) was determined by structural analysis of LPS derived from mutant strains. We have previously shown that the parent strain RM118 elaborates a population of LPS molecules made up of a series of related glycoforms differing in the degree of oligosaccharide chain extension from the distal heptose residue of a conserved phosphorylated inner-core element, L-α-D-Hep*p*-(1→2)-L-α-D-Hep*p*-(1→3)-[β-D-Glc*p*-(1→4)-]-L-α-D-Hep*p*-(1→5)-α-Kdo. The fully extended LPS glycoform expresses the globotetraose structure, β-D-Gal*p*NAc-(1→3)-α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→4)-β-D-Glc*p*. A fingerprinting strategy was employed to establish the structure of LPS from strains mutated in putative glycosyltransferase genes compared to the parent strain. This involved glycose and linkage analysis on intact LPS samples and analysis of *O*-deacylated LPS samples by electrospray ionization mass spectrometry and 1D ^{1}H-nuclear magnetic resonance spectroscopy. Four genes, *lpsA*, *lic2A*, *lgtC*, and *lgtD*, were required for sequential addition of the glycoses to the terminal inner-core heptose to give the globotetraose structure. *lgtC* and *lgtD* were shown to encode glycosyltransferases by enzymatic assays with synthetic acceptor molecules. This is the first genetic blueprint determined for *H. influenzae* LPS oligosaccharide biosynthesis, identifying genes involved in the addition of each glycose residue.

Key words: globotetraose/*Haemophilus influenzae*/lipopolysaccharide

Introduction

Haemophilus influenzae is a bacterium that routinely colonizes the upper respiratory tract of humans. It is also the cause of both upper and lower respiratory tract infections that result from contiguous spread. Occasionally, *H. influenzae* can cause systemic and life-threatening bacteraemic diseases, such as septicaemia and meningitis. Lipopolysaccharide (LPS) of *H. influenzae* functions as a virulence determinant (Preston *et al.*, 1996a). LPS results in major cytotoxic injury to host cells, is a target for host immune responses, and can influence each stage of the pathogenesis of *H. influenzae* infection (Moxon and Maskell, 1992). Changes in LPS expression can alter the virulence of this pathogen (Kimura and Hansen, 1986; Cope *et al.*, 1990; Weiser *et al.*, 1990b; Hood *et al.*, 1996a). A feature of *H. influenzae* LPS is that surface-exposed epitopes of the oligosaccharide are subject to high frequency on-off switching of expression (phase variation) (Weiser *et al.*, 1990a; High *et al.*, 1993; Jarosik and Hansen, 1994). This heterogeneity may be an advantage to the bacteria, allowing them to better confront different host compartments and microenvironments and to survive the host immune response (Weiser and Pan, 1998).

Determination of structure is crucial to understanding the biology of *H. influenzae* LPS and its role in bacterial virulence. *H. influenzae* LPS comprises a variable oligosaccharide moiety and a membrane-anchoring lipid A component (Zamze and Moxon, 1987). LPS from a number of different strains have been shown to contain a common L-*glycero*-D-*manno*-heptose-containing inner-core trisaccharide unit attached to the lipid A moiety via a phosphorylated 2-keto-3-deoxyoctulosonic acid (Kdo) residue (Phillips *et al.*, 1992, 1993, 1996; Gibson *et al.*, 1993; Schweda *et al.*, 1993, 1995; Masoud *et al.*, 1997; Risberg *et al.*, 1997, 1999a,b; Rahman *et al.*, 1999). Each of the Hep residues can provide a point for the addition of Hex residues, which in turn can lead to oligosaccharide chain extensions. The degree of substitution and chain extension from the triheptose unit varies within and between strains (Masoud *et al.*, 1997; Risberg *et al.*, 1999b). In addition, phosphate-containing substituents that include free phosphate groups (*P*), phosphoethanolamine (*P*Etn), pyrophosphoethanolamine (*PP*Etn), and phosphocholine (*P*Cho) also contribute to the structural variability of these molecules. Recently we reported the structure of a globotetraose (β-D-Gal*p*NAc-(1→3)-α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→4)-β-D-Glc*p*) containing LPS from *H. influenzae* strain RM118 (Risberg *et al.*, 1999b), the strain (Rd) for which the complete genome sequence has been determined (Fleischmann *et al.*, 1995). For strain RM118, three major populations of LPS glycoforms were identified, all containing a *P*Cho→6)-β-D-Glc*p* group off the Hep attached to the Kdo unit, but differing in the length of the oligosaccharide chains off the third Hep of the inner-core element. LPS glycoforms expressing a fully assembled globotetraose side chain and sequentially truncated glycoforms containing globoside (α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→4)-β-D-Glc*p*) and

[1]To whom correspondence should be addressed

lactose (β-D-Gal*p*-(1→4)-β-D-Glc*p*) were characterized (Risberg *et al.*, 1999b).

The availability of the complete genome sequence of *H. influenzae* strain Rd facilitated a comprehensive study of LPS biosynthetic loci in the type b strains RM153 and RM7004. Many predicted gene functions were correlated with particular steps in the synthesis of the LPS in strain RM153 (Hood *et al.*, 1996a). The LPS from strain RM118 has a significantly different structure to that of RM153, the pattern and degree of substitution of oligosaccharide chain extensions being entirely different. Thus it is not possible to assign the genetic basis for biosynthetic functions for RM118 LPS from the information currently available. In particular, the genetic basis for expression of the globotetraose structure and Hex addition to the first Hep has not previously been reported.

In this study we employ a structural fingerprinting strategy to determine and compare the structures of LPS obtained from a series of defined mutants in LPS biosynthetic genes in *H. influenzae* strain RM118. We identify the glycosyltransferases involved in the assembly of the globotetraose side chain and in the biosynthesis of the inner-core region of the LPS molecule. The transferase functions of gene products involved in sequential addition of α-1,4-linked Gal*p* (LgtC) and β-1,3-linked Gal*p*NAc (LgtD) to give the globoside and globotetraose structures, respectively, were unambiguously determined by enzymatic assays with synthetic acceptors.

Results

Construction and screening of mutant strains

A set of mutants was made (Table I) to investigate in detail the genetic basis of the biosynthesis of the oligosaccharide portion of LPS from *H. influenzae* strain RM118. The DNA constructs used to mutate the majority of these genes had been previously reported (Hood *et al.*, 1996a). *lic1* and *lic2A* are phase-variable LPS biosynthetic loci described previously (Weiser *et al.*, 1989; High *et al.*, 1993). In the course of this study there was no obvious candidate gene responsible for adding the Glc to the first Hep. Searching the Rd genome sequence database with the LgtF sequence from *Neisseria* gave a match (31% identify over 247 amino acids) to reading frame HI0653. *lgtF* was amplified by polymerase chain reaction (PCR) from chromosomal DNA of strain RM118, cloned then inactivated and used to transform *H. influenzae*. Genes responsible for the synthesis of Kdo (*kdsA*, *kdsB*) and the Kdo transferase (*kdtA*) have been identified from the genome sequence (Fleischmann *et al.*, 1995). Various attempts to construct strains mutated in the *kdtA* gene failed, similar to findings with type b strains (Hood *et al.*, 1996a). This is assumed to be due to nonviability of this mutant. LPS isolated from RM118 and the isogenic mutant strains was analysed by tricine–sodium dodecyl sulfate–polyacrylamde gel electrophoresis (T–SDS–PAGE) (data not shown). Strains mutated in genes most likely encoding glycosyltransferases for RM118 oligosaccharide synthesis and that showed an altered pattern of LPS bands when compared to wild type (Figure 1), were selected for detailed structural analysis of their LPS as described below. A mutant in which the *lic1* locus is inactivated was also investigated.

Table I. LPS-related genes investigated in strain RM118 in this study

Gene	HI number	Reference
kdtA	**0652**	Hood *et al.*, 1996a
lgtC	**0259**	Hood *et al.*, 1996a
lgtD	**1578**	Hood *et al.*, 1996a
lpsA	**0765**	Hood *et al.*, 1996a
orfZ	**0260.1**	Hood *et al.*, 1996a
opsX	**0261**	Hood *et al.*, 1996a
orfH	**0523**	Hood *et al.*, 1996a
rfaF	**1105**	Hood *et al.*, 1996a
lic2A	**0550**	High *et al.*, 1993
lic1	**1537–1540**	Weiser *et al.*, 1990a
lgtF	**0653**	This study
cld	0866	Hood *et al.*, 1996a
galU	0812	Hood *et al.*, 1996a
kfiC	0868	Hood *et al.*, 1996a
lsg1	0867	Hood *et al.*, 1996a
orfM	0260	Hood *et al.*, 1996a
orfE	0869	Hood *et al.*, 1996a
orfO	0870	Hood *et al.*, 1996a
orfY	0871	Hood *et al.*, 1996a
pgmB	0740	Hood *et al.*, 1996a
pgmC	1337	Hood *et al.*, 1996a
rfe	1716	Hood *et al.*, 1996a
rfbP	0872	Hood *et al.*, 1996a
rfbB	0873	Hood *et al.*, 1996a

HI numbers are the ORF designations given by the Institute for Genomic Research for the *H. influenzae* genome sequence database. The genes given in boldface type are those selected for detailed comparative analysis of the expressed LPS glycoforms.

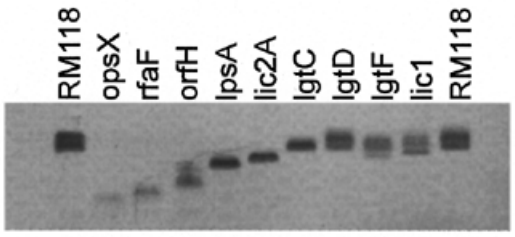


Fig. 1. The electrophoretic gel-migration patterns after T–SDS–PAGE of LPS purified from RM118 wild type and strains mutated in putative glycosyltransferase genes. RM118 corresponds to the wild-type strain, and the isogenic mutants are listed by the relevant LPS gene.

Structural characterization of LPS

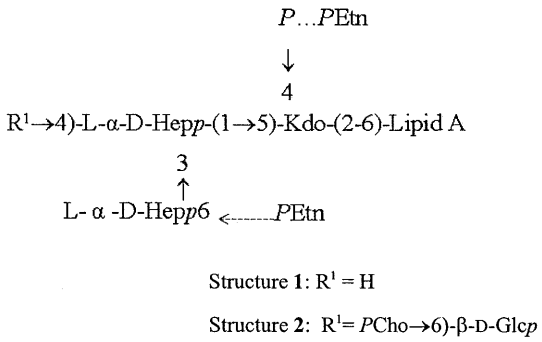
Analysis of LPS from strain RM118 by T–SDS–PAGE (Figure 1) showed a heterogeneous pattern of bands corresponding in electrophoretic mobility to populations of low-molecular-mass LPS composed of lipid A and oligosaccharide components differing in the number of attached sugar residues. We have previously shown that strain RM118 grown under similar

conditions expresses populations of LPS containing three to five glycose residues attached to a common inner-core element (Risberg *et al.*, 1999b). LPS from strains with mutations in *lic1*, *lgtF*, and *lgtD* showed similar complex banding patterns, and those from strains with mutations in *lgtC*, *lic2A*, *lpsA*, *orfH*, *rfaF*, and *opsX* gave less complex patterns comprising bands having consecutively faster mobilities consistent with successive sugar deletions. Identification of the nature of sugar deletions in the LPS samples from mutant strains, grown in liquid culture, was achieved by comparative structural analysis. Structural fingerprinting was done using electrospray ionization mass spectrometry (ESI-MS) and 1D ^{1}H–nuclear magnetic resonance (NMR) analysis of *O*-deacylated LPS (LPS-OH) samples. In addition, glycose and linkage analyses were carried out on intact LPS samples. Such analyses established the key structural features of the altered LPS glycoforms. The ESI-MS data obtained in the negative ion mode is presented in Table II. The LPS-OH samples from the mutant strains gave data consistent with the presence of oligosaccharides linked *via* Kdo-4-phosphate to a common *O*-deacylated lipid A moiety (lipid A-OH) differing in the number of Hep, Hex, and phosphate-containing substituents (Phillips *et al.*, 1992, 1996; Schweda *et al.*, 1993, 1995; Masoud *et al.*, 1997; Risberg *et al.*, 1997, 1999b). *H. influenzae* lipid A-OH is known (Masoud *et al.*, 1997; Helander *et al.*, 1988) to be composed of bisphosphorylated β-1,6-linked glucosamine disaccharide substituted by 3-hydroxytetradecanoamide groups at C-2 and C-2′.

opsX mutant. Inactivation of *opsX* gave rise to deep-rough LPS, which was devoid of Hep or hexose residues, containing only a phosphorylated Kdo attached to lipid A (Table II). We have previously shown that mutation in the *opsX* gene of *H. influenzae* type b strain RM153 results in truncation of the LPS between HepI and Kdo (Hood *et al.*, 1996a). Tandem mass spectrometry (MS-MS) analysis of the LPS-OH sample by low-energy collisional activation of the doubly charged molecular ions (*m/z* 625.5 and 616.5) afforded in both cases a major fragment ion at *m/z* 952 (lipid A-OH) arising from cleavage of the Kdo-β-D-glucosamine bond (data not shown). The mass of this fragment ion is consistent with that expected for *H. influenzae* lipid A-OH (Helander *et al.*, 1988). It seems likely that the anhydro-Kdo-P derivative (corresponding to *m/z* 616.5) is due to the Kdo molecule in this mutant only, not being substituted by the HepI residue at the 5-position. It is evident that RM118 and RM153 *opsX* mutants express LPS similar to that from the previously characterized Rd*isn* (I69) strain (Helander *et al.*, 1988; Preston *et al.*, 1996b). The I69 LPS phenotype arises from a mutation in the heptose biosynthetic gene *gmhA* (Brook and Valvano, 1996), rendering the mutant strain incapable of adding heptose to LPS.

rfaF mutant. ^{1}H-NMR analysis of LPS-OH from RM118*rfaF* showed, in addition to the expected ^{1}H resonance from the α-linked glucosamine residue of lipid A, an anomeric proton resonance (~5.19 ppm) in the low-field region from a single heptose unit. Sugar analysis confirmed the Hep residue to be L-*glycero*-D-*manno* heptose. Correspondingly, the ESI-MS spectrum was dominated by a single abundant doubly charged ion at *m/z* 721.6 consistent with the structure Hep_1-Kdo-lipid A-OH (Table II).

orfH mutant. Strain RM118*orfH* gave a mixture of LPS glycoforms, each containing two Hep residues, as evidenced from the ESI-MS data (Table II). In addition to the major population of glycoforms containing an additional Hep residue, that is, Hep_2•*P*Etn$_{0\text{-}2}$•Kdo-lipid A-OH, compared to RM118*rfaF* LPS, were species containing a Hex-*P*Cho unit. Sugar analysis indicated the presence of D-glucose and the *P*Cho methyl protons gave an intense signal in the ^{1}H-NMR at 3.24 ppm.. LPS from this strain reacted with TEPC-15, a *P*Cho specific monoclonal antibody (MAb) (Weiser *et al.*, 1997). Linkage analysis revealed the presence of terminal Hep, 3-substituted Hep and 3,4,-disubsituted Hep residues (Table III). This data is consistent with RM118*orfH* expressing the two major LPS glycoform structures 1 and 2 (see Scheme 1, *P*Etn shows partial substitution). The occurrence of two bands for the LPS of RM118*orfH* when analyszed by T–SDS–PAGE (Figure 1) is consistent with this conclusion.



Scheme 1. Structures 1 and 2 of the LPS glycoform.

lpsA mutant. ESI-MS analysis of LPS-OH from RM118*lpsA* indicated it to contain glycoforms having an additional Hep residue when compared to RM118*orfH* (Figure 2, Table II), the *P*Cho containing Hex1 glycoform being the major LPS species. Linkage analysis was consistent with sequential addition of Hep to the terminal Hep in structure 2 (Table III). Correspondingly, the ^{1}H-NMR spectrum of this LPS-OH showed the characteristic pattern in the low-field region (5.0–6.0 ppm) for the LPS tri-Hep inner-core element (HepII, 5.76 ppm; HepI/HepIII, 5.16/5.15 ppm) of *H. influenzae* (Figure 3) (Risberg *et al.*, 1999b). This data is consistent with the RM118*lpsA* derived LPS having the structure 3 (Figure 2, Table IV).

lic2A mutant. ESI-MS analysis of the LPS-OH samples from strain RM118*lic2A* revealed the presence of Hex2 glycoforms as the major LPS species (Table II). Analysis of the RM118*lic2A* LPS indicated the presence of D-glucose as the only neutral hexose, linkage analysis indicating it to be a terminal residue (Table III). A significant proportion of 2-linked Hep residues was also revealed by linkage analysis. It is noteworthy, that 2-subsituted Hep residues were not detected in the LPS sample from the *lpsA* mutant due to substitution of that residue by *P*Etn groups (cf. structure 3), which are not readily cleaved under the hydrolysis conditions employed in the linkage analysis procedure. In accord with these findings, it can be concluded that LPS from the *lic2A* mutant differs from that of the *lpsA* mutant in that it carries a glucose residue at the

Table II. Negative ion ESI-MS data and proposed compositions of LPS-OH from *H. influenzae* strain RM118 mutants.

	Observed ions (*m/z*)		Molecular mass (Da)			
Strain	$(M-2H)^{2-}$	$(M-3H)^{3-}$	Observed	Calculated	Relative intensity[b]	Proposed composition
opsX[a]	616.5	—	1235.0	1235.2	80.0	Kdo-*P* (-H_2O)c, Lipid A-OH
	625.5	—	1253.0	1253.2	20.0	Kdo-*P*, Lipid A-OH
rfaF	721.6	—	1445.2	1445.3	100.0	Hep, Kdo-*P*, Lipid A-OH
orfH	817.5		1637.5	1637.5	10.6	2Hep, Kdo-*P*, Lipid A-OH
	879.2	—	1760.4	1760.6	42.4	2Hep, *P*Etn, Kdo-*P*, Lipid A-OH
	940.5	—	1883.4	1883.6	15.2	2Hep, 2*P*Etn, Kdo-*P*, Lipid A-OH
	1042.3	—	2086.6	2087.8	21.2	*P*Cho, Hex, 2Hep, *P*Etn, Kdo-*P*, Lipid A-OH
	1104.8	—	2211.6	2210.8	10.6	*P*Cho, Hex, 2Hep, 2*P*Etn, Kdo-*P*, Lipid A-OH
lpsA	1056.3	—	2114.6	2114.9	10.3	Hex, 3Hep, *P*Etn, Kdo-*P*, Lipid A-OH
	1139.0	759.0	2280.0	2279.9	69.0	*P*Cho, Hex, 3Hep, *P*Etn, Kdo-*P*, Lipid A-OH
	1200.4	800.0	2402.9	2403.0	20.7	*P*Cho, Hex, 3Hep, 2*P*Etn, Kdo-*P*, Lipid A-OH
lic2A	1137.5	758.3	2277.5	2277.0	18.0	2Hex, 3Hep, *P*Etn, Kdo-*P*, Lipid A-OH
	1220.0	813.3	2442.5	2442.1	51.0	*P*Cho, 2Hex, 3Hep, *P*Etn, Kdo-*P*, Lipid A-O H
	1281.9	854.3	2565.8	2565.1	31.0	*P*Cho, 2Hex, 3Hep, 2*P*Etn, Kdo-*P*, Lipid A-OH
lgtC	1301.1	867.1	2603.8	2604.2	80.0	*P*Cho, 3Hex, 3Hep, *P*Etn, Kdo-*P*, Lipid A-OH
	1362.7	908.2	2727.5	2727.3	20.0	*P*Cho, 3Hex, 3Hep, 2*P*Etn, Kdo-*P*, Lipid A-OH
lgtD	1300.6	867.1	2603.9	2604.2	62.0	*P*Cho, 3Hex, 3Hep, *P*Etn, Kdo-*P*, Lipid A-OH
	1362.0		2726.0	2727.3	5.0	*P*Cho, 3Hex, 3Hep, 2*P*Etn, Kdo-*P*, Lipid A-OH
	1381.6	921.2	2765.5	2766.4	28.0	*P*Cho, 4Hex, 3Hep, *P*Etn, Kdo-*P*, Lipid A-OH
	1443.3	962.1	2888.9	2889.4	5.0	*P*Cho, 4Hex, 3Hep, 2*P*Etn, Kdo-*P*, LipidA-OH
lgtF	1137.4	757.8	2276.6	2277.0	16.0	2Hex, 3Hep, *P*Etn, Kdo-*P*, Lipid A-OH
	—	798.9	2399.7	2400.0	5.0	2Hex, 3Hep, 2*P*Etn, Kdo-*P*, Lipid A-OH
	1218.6	812.2	2439.4	2439.2	18.0	3Hex, 3Hep, *P*Etn, Kdo-*P*, Lipid A-OH
	1280.0	853.2	2562.3	2562.2	7.0	3Hex, 3Hep, 2*P*Etn, Kdo-*P*, Lipid A-OH
	1320.3	879.9	2642.8	2642.4	38.0	HexNAc, 3Hex, 3Hep, PEtn, Kdo-*P*, Lipid A-OH
	—	920.8	2765.4	2765.4	16.0	HexNAc, 3Hex, 3Hep, 2*P*Etn, Kdo-*P*, Lipid A-OH
lic1	1056.2	—	2114.4	2114.9	12.8	Hex, 3Hep, *P*Etn, Kdo-*P*, Lipid A-OH
	1117.4	744.9	2237.3	2237.9	7.2	Hex, 3Hep, 2*P*Etn, Kdo-*P*, Lipid A-OH
	1218.3	812.1	2438.9	2439.2	20.0	3Hex, 3Hep, *P*Etn, Kdo-*P*, Lipid A-OH
	1279.7	852.6	2561.1	2562.2	12.9	3Hex, 3Hep, 2*P*Etn, Kdo-*P*, Lipid A-OH
	1299.2	865.7	2600.3	2601.3	9.0	4Hex, 3Hep, *P*Etn, Kdo-*P*, Lipid A-OH
	1360.6	906.5	2722.9	2724.4	10.0	4Hex, 3Hep, 2*P*Etn, Kdo-*P*, Lipid A-OH
	1401.1	933.7	2804.6	2804.5	12.9	HexNAc, 4Hex, 3Hep, *P*Etn, Kdo-*P*, Lipid A-OH
	1462.5	974.6	2927.1	2927.6	15.1	HexNAc, 4Hex, 3Hep, 2*P*Etn, Kdo-*P*, Lipid A-OH

Average mass units were used for calculation of molecular weight based on proposed composition as follows: Hex, 162.15; HexNAc, 203.19; Hep, 192.17; Kdo-*P*, 300.16; *P*Etn, 123.05; *P*Cho, 165.05. The average molecular mass of *O*-deacylated lipid A (lipid A-OH) is 953.03.
[a]Data acquired by CE-ESI- MS on a crystal Model 310 CE instrument interfaced to an API 3000 triple quadrupole mass spectrometer (Perkin-Elmer/Sciex) fitted with a bare fused silica capillary column and using 30 mM morpholine-acetate (pH 9.0) containing 5% methanol as the separation buffer.
[b]Measured from the respective molecular ions in the reconstructed spectrum.
[c]The major ion observed corresponded to the molecular ion –18 (loss of H_2O).

2-position of HepIII as shown in structure 4 (Table IV). The presence of an additional ^{1}H-NMR signal at 4.65 ppm indicated the terminal D-Glc*p* to have the β-configuration, the upfield shifted value of the resonance for HepII (5.58 ppm) compared to that of the unsubstituted analogue (5.76 ppm), being indicative of the 1,2-linkage to HepIII (structure 4) (Masoud *et al.*, 1997; Schweda *et al.*, 1993, 1995).

lgtC mutant. For the RM118*lgtC* mutant, ESI-MS analysis of the LPS-OH sample revealed the presence of Hex3 glycoforms.

Table III. Linkage analysis data for LPS derived from *H. influenzae* RM118 mutated in LPS biosynthesis genes

		Relative detector response for mutant strain						
Methylated sugar[a]	T_{gm}[b]	*orfH*	*lpsA*	*lic2A*	*lgtC*	*lgtD*	*lgtF*	Linkage assignment
2,3,4,6-Me_4-Glc	1.00	-tr[c]	6.7	22.5	tr	tr		D-Glc*p*-(1→
2,3,4,6-Me_4-Gal	1.05	-tr			34.3	22.9	10.3	D-Gal*p*-(1→
2,3,6-Me_3-Glc	1.18	-tr		6.3	33.0	24.4	18.8	→4)-D-Glc*p*-(1→
2,3,6-Me_3-Gal	1.17					12.4	16.6	→4)-D-Gal*p*-(1→
2,4,6-Me_3-Gal[d]	1.20	5.0	12.4	5.4	tr	4.0	10.3	→3)-D-Gal*p*-(1→
2,3,4,6,7-Me_5-Hep	1.30	27.4	35.4					L,D-Hep*p*(1→
3,4,6,7-Me_4-Hep	1.44			53.3	15.8	16.0	16.3	→2)-L,D-Hep*p*(1→
2,4,6,7-Me_4-Hep	1.47	37.0					19.7	→3)-L,D-Hep*p*(1→
2,6,7-Me_3-Hep	1.52	23.7	45.4	12.5	10.6	17.1		→3,4)-L,D-Hep*p*-(1→
2,3,4,6-Me_4-GalN	1.58						8.0	D-Gal*p*NAc-1(1→

[a]2,3,4,6-Me_4-Glc represents 1,5-di-O-acetyl-2,3,4,6-tetra-O-methyl-D-glucitol-1-d_1, etc.
[b]Retention times (T_{gm}) are reported relative to 2,3,4,6-Me_4-Glc, and values are not corrected for differences in detector response factors.
[c]Trace amount detected.
[d]All samples showed detectable levels of 2,4,6,-Me_3-Gal. Only the sample from mutant strain *lgtF* is expected to give rise to this methylated derivative (Risberg *et al*., 1999). We suspect a contaminating source contributes to this peak in all samples.

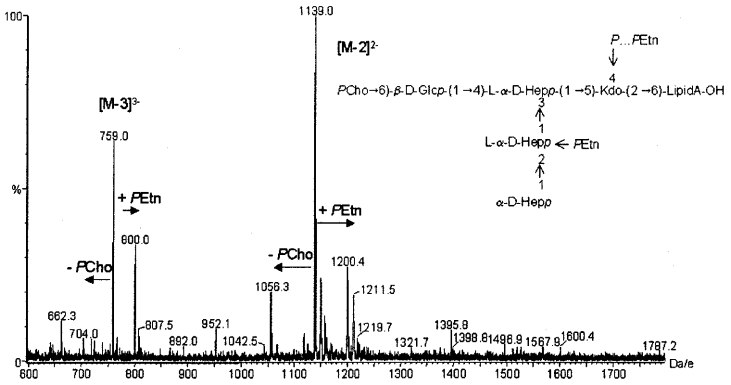


Fig. 2. Negative-ion ESI-MS of the LPS-OH from the *lpsA* mutant of *H. influenzae* RM118 showing doubly and triply charged ions from the major Hex1 glycoform (structure 3).

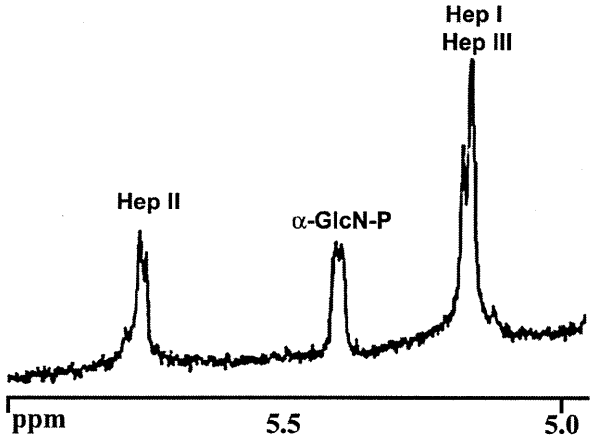
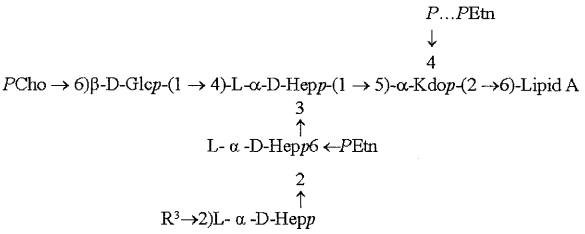


Fig. 3. ^{1}H-NMR spectrum of the LPS-OH from the *lpsA* mutant of *H. influenzae* RM118 showing the α-anomeric proton region between 5.0 and 6.0 ppm. Anomeric resonances corresponding to the 3,4-disubstituted Hep (HepI), 6-*P*Etn substituted Hep (HepII), terminal Hep (HepIII), and phosphorylated α-GlcN in the lipid A region are indicated.

Sugar analysis indicated that LPS from the *lgtC* mutant contained D-galactose, which by linkage analysis was found to be present as a terminal residue (Table III). Linkage analysis also revealed 4-linked D-Glc*p* residues consistent with the major Hex3 glycoform (Table III) being substituted by a lactose moiety at HepIII (structure 5). The ^{1}H-NMR spectrum of the LPS-OH is identical to that previously reported by us (Risberg *et al.*, 1999b) for the lactose-containing Hex3 LPS glycoform which is present in the parent strain. The *lgtC* gene was shown to encode a 1,4-α-galactosyltransferase by examination of the transferase activity of the recombinant enzyme (data not shown).

lgtD mutant. A mixture of Hex3 and Hex4 LPS glycoforms were elaborated by *H. influenzae* RM118*lgtD* (Table II). In accord, two bands were observed on T–SDS–PAGE analysis of the LPS, one corresponding in electrophoretic mobility to that from the *lgtC* mutant and a slower migrating band (Figure 1). LPS from this mutant strain contained terminal and 4-linked D-Gal*p* residues (Table III). A comparison of the 1D

Table IV. Structure of the major LPS glycoforms of increasing oligosaccharide chain length from HepIII in mutant stains of *H. influenzae* RM118



Structure	Glycoform	Mutant Strain(s)	Substitution Pattern (R^3)
3	Hex1	*lpsA*	H
4	Hex2	*lic2A*	β-D-Glc*p*
5	Hex3	*lgtC,lgtD*	β-D-Gal*p*-(1→4)-β-D-Glc*p*
6	Hex4	*lgtD*	α-D-Gal*p*-(1→4)-β-D-Gal-(1→4)-β-D-Glc

^{1}H-NMR spectra with those of the parent strain and its *lgtC* mutant, pointed to the presence of an α-D-Gal*p*-(1→4)-β-D-Gal*p* unit in the Hex4 glycoform, a signal at 5.01 p.p.m. being indicative of the terminal α-D-Gal*p* residue (structure 6) (Table IV). The *lgtD* gene product was shown to have β-Gal*p*NAc transferase activity with the synthetic acceptor FCHASE-p^k, by a comparative assay of the parent and the *lgtD* mutant strains (Figure 4).

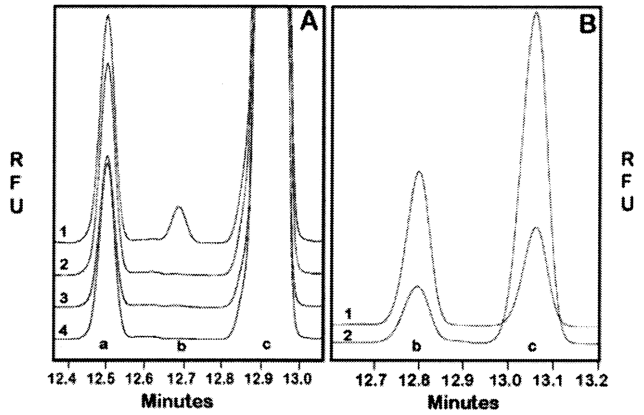


Fig. 4. Capillary electrophoresis analysis of LgtD activity in *H. influenzae* strain RM118. (**A**) Trace 1 is a complete reaction mixture using the 100,000 × *g* pellet of a sonicate as enzyme source; trace 2 is similar to the reaction mixture in 1, except it is missing UDP-GalNAc; trace 3 is a complete reaction mixture from the mutant RM118:lgtD; trace 4 is similar to trace 3 except it is missing UDP-GalNAc. The peak a is an impurity in the FCHASE-P^K preparation, peak b is FCHASE-Globotetraose, peak c is FCHASE-P^K. (**B**) The trace 1 is the purified product for thin-layer chromatography from a reaction as described for **A** trace 1. Trace 2 is the same material as trace 1, but treated with β-hexosaminidase.

lgtF mutant. Mutation of the *lgtF* gene in RM118 gave a strain from which the LPS neither reacted with MAb TEPC-15 (data not shown) nor showed the characteristic *P*Cho methyl proton signal (3.24 ppm) in the ^{1}H-NMR spectrum. Linkage analysis indicated that the LPS lacked the terminal β-D-Glc*p* residue, containing only mono-3-substituted HepI residues (Table III). A similar distribution of glycoforms, as found in the parent strain LPS, differing in the length of the oligosaccharide chain from HepIII was observed for LPS-OH from the *lgtF* mutant in its ESI-MS (Table II). It is noteworthy that full extension of the globotetraose unit, β-D-Gal*p*NAc-(1→3)-α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→4)-β-D-Glc*p* (Hex3.HexNAc) from HepIII can occur in the absence (Figure 5) or the presence (Table II, *lic1*) of the β-D-Glc*p* residue at HepI.

lic1 mutant. ESI-MS analysis of *O*-deacylated LPS from the *lic1* mutant gave a similar heterogeneous mixture of glycoforms (Table II) as that observed in the parent strain but lacking *P*Cho substituents. Examination of the ^{1}H-NMR spectrum of RM118*lic1* LPS-OH revealed the absence of the characteristic *P*Cho methyl proton signal at 3.24 ppm. Additionally, the LPS from this mutant did not react with MAb TEPC-15 (not shown).

Discussion

In a previous study, the complete genome sequence of *H. influenzae* strain Rd facilitated the identification and study of candidate LPS genes in a *H. influenzae* type b strain, RM153 (Hood *et al.*, 1996a). Many probable gene functions were established relevant to LPS biosynthesis in strain RM153, but others remain unidentified. Knowledge of the detailed structure of the LPS from strain RM118 (Risberg *et al.*, 1999b) has shown that the Hex extensions in the oligosaccharide portion are fundamentally different to that of strain RM153. In this study we investigate the genetic basis for the biosynthesis of the major LPS glycoforms expressed by RM118, the index sequenced strain.

In our investigation of the *H. influenzae* LPS inner core, attempts to remove Kdo, the first sugar added to the lipid A, have repeatedly failed presumably because Kdo is required to complete lipid A synthesis and is thus likely essential for cell viability. The Kdo transferase function of KdtA has been demonstrated by complementation experiments in *Escherichia coli* (White and Raetz, 1998). The results from the present study would suggest that *opsX*, *rfaF*, and *orfH* are the genes encoding the enzymes that add the first, second, and third Hep, respectively, to the Kdo, to form the inner core of *H. influenzae* LPS (Figure 6). *opsX*, *rfaF*, and *orfH* have some homology to heptosyl transferases of other bacteria (Hood *et al.*, 1996a). The data from RM118 mutants is consistent with that obtained in the type b strain RM153, where *opsX*, *rfaF*, and *orfH* were proposed as the genes encoding the HepI, HepII, and HepIII transferases, respectively (Hood *et al.*, 1996a). This shows a conservation of the genetic basis for as well as the structure of the triheptosyl inner core of *H. influenzae* LPS.

Each heptose of the inner-core moiety of *H. influenzae* LPS provides a point for elongation by oligosaccharide chains (Masoud *et al.*, 1997; Risberg *et al.*, 1999b). Each of the genes *lpsA*, *lic2A*, *lgtC*, *lgtD*, and *lgtF* are predicted to encode glycosyltransferase enzymes involved in oligosaccharide elongation by homology comparisons with genes of similar function (Hood *et al.*, 1996a; Campbell *et al.*, 1997). The results of the present investigation would indicate that the *lpsA* gene product plays a role in controlling oligosaccharide chain extension from HepIII. A mutation in the *lpsA* gene affords a truncated LPS in which HepIII is devoid of oligosaccharide chain extensions. ESI-MS analysis of the RM118*lpsA*-derived LPS-OH indicated a *P*Cho-containing Hex1 glycoform as the major LPS species (structure 2), confirming that HepI can be substituted in

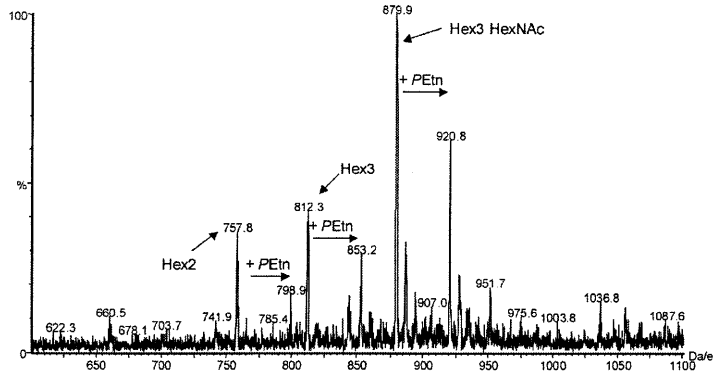


Fig. 5. Negative ion ESI-MS of the triply charged molecular ion region of the LPS-OH from the *lgtF* mutant of *H. influenzae* RM118. Peaks arising from the Hex 2 (β-D-Gal*p*-(1→4)-β-D-Glc*p*), Hex 3 (α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→4)-β-D-Glc*p*), and Hex3•HexNAc (β-D-Gal*p*NAc-(1→3)-α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→4)-β-D-Glc*p*) are indicated.

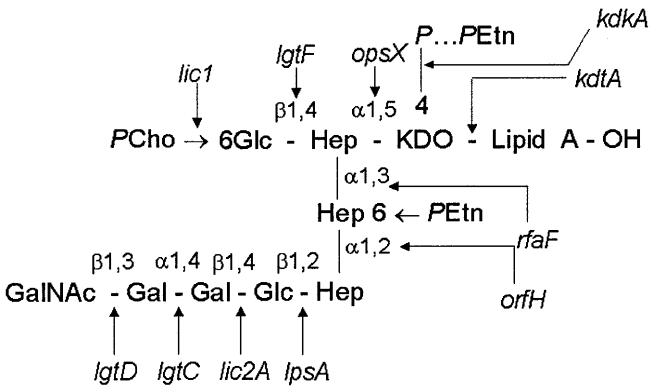


Fig. 6. A schematic representation of the structure of LPS from *H. influenzae* strain RM118 based on the results of the analysis of Risberg *et al.*, (1999b). The proposed site of action in LPS biosynthesis of loci characterized in this study are shown, linked by arrows to the relevant saccharide linkage. The phase variable loci are *lic1*, *lic2A*, and *lgtC*. Represented in the LPS structure: KDO, 2-keto-3-deoxyoctulosonic acid; Hep, L-glycero-D-manno-heptose; Glc, D-glucose; Gal, D-galactose; GalNAc, N-acetylgalactosamine; *P*Etn, phosphoethanolamine; *P*, phosphate; *P*Cho, phosphocholine. For the heptose residues, listed top to bottom are heptose I, heptose II, then heptose III.

the absence of Hex extension from HepIII. The *lic2A*, *lgtC*, and *lgtD* mutants, which contain a functional *lpsA* gene, are capable of adding a β-D-Glc*p* residue in a 1,2-linkage to initiate chain extension from HepIII (Table IV). LpsA has homology to the group of galactosyltransferases typified by Lic2A and LgtB of *Haemophilus* and *Neisseria*, respectively. In strain RM153, it was also found that LPS from a *lpsA* mutant lacked any extension from the third Hep (Hood *et al.*, 1996a). However, in strain RM153, *lpsA* apparently plays a slightly different role being responsible for the addition of Gal as the sole extension from the third Hep (Hood *et al.*, 1996a). Thus, LpsA is likely the transferase for the addition of the first glycose to HepIII in *H. influenzae* LPS biosynthesis.

The RM118*lic2A* mutant showed a *P*Cho-containing Hex2 glycoform as a major LPS species (structure 4) and RM118lgtC, which contains a functional *lic2A* gene, elaborates LPS containing a lactose side chain at HepIII (structure 5). This is consistent with the involvement of the *lic2A* gene to add the β-D-Gal*p* unit in a 1,4 linkage to the terminal β-D-Glc*p* residue attached to HepIII. Lic2A homologues in type b strains have been shown to be involved in expression of the digalactoside-containing Pk epitope (α-D-Gal*p*-[1→4]-β-D-Gal*p*-[1→4]-β-D-Glc*p*). Homology comparisons with other databank sequences support the function of Lic2A as a β-galactosyltransferase; importantly, it has significant homology to the *Neisseria* LgtB and LgtE proteins, both of which are galactosyltransferases (Wakarchuk *et al.*, 1996). Structural analysis of LPS from an RM118 strain mutated in *lgtC* confirmed the loss of α-D-Gal*p*, supporting the α-galactosyltransferase function for this gene. Correspondingly the *lgtD* mutant and the parent strain, which contain a functional *lgtC* gene, are capable of adding a α-D-Gal*p* in a 1,4 linkage to the terminal β-D-Gal*p* of the lactose epitope (structure 6). The function of LgtC was confirmed by demonstrating α-galactosyltransferase activity with the recombinant protein and a synthetic FCHASE-Lac acceptor. In *N. meningitidis*, LgtC is also an α-galactosyltransferase (Gotschlich, 1994; Wakarchuk *et al.*, 1998). It follows that the *lgtC* gene encodes the specific α-galactosyltransferase for the synthesis of the α-D-Gal*p*-(1→4)-β-D-Gal*p* of the RM118 Hex4 LPS glycoform (Figure 6). The parent strain RM118 that contains a functional *lgtD* gene is capable (unlike the mutant) of elaborating the complete globotetraose unit, which is indicative of its role in adding the terminal β-D-Gal*p*NAc. The *H. influenzae lgtD* gene is a homologue of two related *Neisseria* genes, *lgtA* and *lgtD*, which add Glc*p*NAc and Gal*p*NAc, respectively, to *N. gonorrhoeae* LPS (Gotschlich, 1994). Enzyme assays with extracts of RM118 and the RM118*lgtD* mutant confirmed the β-D-Gal*p*NAc transferase activity. The *lgtD* gene was found not to be present in the type b strains RM153 and RM7004 (called *lgtA* in Hood *et al.*, 1996a). Correspondingly, the LPS elaborated by strain RM153 does not contain a Gal*p*NAc moiety (Masoud *et al.*, 1997).

It is noteworthy that, though the activity of glycosyltransferases adding the distal residues of the globotetraose (*lgtD*) and globoside (*lgtC*) oligosaccharide side chains could be assayed with the appropriate synthetic acceptor, similar experiments to assay the activity of the transferases involved in synthesis of the lactose moiety (Lic2A and LpsA) were unsuccessful. It is likely that the latter two enzymes have more stringent specificities that require the acceptor sugar to be

linked to the inner-core Hep residues. Characterization of the initial set of genes and mutant strains available for study of RM118 LPS biosynthesis gave no obvious candidate responsible for addition of the β-D-Glc*p* unit to HepI. An *lgtF* homologue was identified in strain RM118 by searching the strain Rd genome sequence for matches to genes required for the addition of Hex sugars to Hep residues in the LPS of other organisms. These search sequences included the *rfaK* and *lgtF* genes of *Neisseria* (Kahler *et al.*, 1996). Analysis of the LPS from strain RM118*lgtF* supported a role for LgtF in chain extension from HepI. The ESI-MS showed molecular ions corresponding to a mixture of glycoforms having chain extensions, including lactose and globotetraose, from HepIII of a triheptosyl inner-core unit that lacks *P*Cho→6)-β-D-Glc*p* at HepI (Figure 3). Thus, the processes of chain extension from both HepI and HepIII appear to be largely independent in the LPS of strain RM118.

The heterogeneity observed in *H. influenzae* LPS structure may be due in part to intrinsic variation in the biosynthesis of such a complex structure, but the majority of variation observed is presumed to be due to specific LPS biosynthetic genes capable of variable expression (phase variation). This study has allowed us for the first time to confirm the genes involved in the synthesis of an important phase-variable epitope of *H. influenzae* LPS, the digalactoside. In strain RM118, Lic2A adds the proximal β-D-Gal*p* and LgtC the terminal α-D-Gal*p* to the digalactoside (-D-Gal*p*-(1→4)-β-D-Gal*p*) as part of the extension from HepIII, whereas the same epitope is expressed as the terminal extension from a diglucoside on the second Hep in the type b strain RM153 (Masoud *et al.*, 1997). Both *lic2A* and *lgtC* are phase-variable genes (High *et al.*, 1993; Hood *et al.*, 1996b), making the expression of the epitope highly variable within and between organisms. The digalactoside epitope is expressed in the LPS of many related bacteria, including *Neisseria* (Virji *et al.*, 1990). The epitope is potentially immunodominant, and its presence offers the potential for molecular mimicry of host structures and can influence the survival of *Haemophilus* within experimental systems (Hood *et al.*, 1996b; Weiser and Pan, 1998).

In addition to the order and stereochemistry of the sugar residues, the location, type and frequency of substituents such as *P*, *P*Etn, and *P*Cho can have a profound affect on LPS structure and biological function. The *lic1* locus is essential for the phase-variable addition of *P*Cho to the *H. influenzae* LPS molecule (Weiser *et al.*, 1997; Lysenko *et al.*, 2000). DNA sequence polymorphisms in *lic1* direct the different acceptor specificity observed for *P*Cho incorporation and influence the resistance of *H. influenzae* to innate humoral immunity (Weiser and Pan, 1998; Lysenko *et al.*, 2000). The gene encoding a Kdo kinase, *kdkA*, responsible for phosphorylation of Kdo, has been identified (White *et al.*, 1999). This gene has previously been investigated by us as *orfZ* and when mutated was shown to alter bacterial survival in an infant rat model of infection (Hood *et al.*, 1996a). The only remaining substituents in the core oligosaccharide, whose genetic control remains unknown, therefore, are the *P*Etn residues that are attached to the 6-position of the HepII residue stoichiometrically and sometimes to the phosphate group on Kdo.

In summary, the genetic blueprint for synthesis of the major globotetraose-containing oligosaccharide of RM118 LPS has been elucidated. The type b strain RM153 and strain RM118 have a gene pool for LPS biosynthesis that is generally the same for the two strains but with some key differences related to their respective LPS structure and biology.

Materials and methods

Bacterial strains and culture conditions

The *H. influenzae* Rd strain was originally obtained from Alexander and Leidy by Herriot. It was given to H. O. Smith, who named the strain KW-20 and used it in the genome sequencing project (Fleischmann *et al.*, 1995). This same strain obtained from the Smith laboratory has been used by us (RM118). The genotypes of mutants derived from this strain are listed in Table I. *H. influenzae* strains were grown at 37°C in brain heart infusion broth supplemented with hemin (10 μg/ml), nicotinamide adenine dinucleotide (2 μg/ml), and kanamycin (10 μg/ml) when appropriate.

E. coli strain DH5α was used to propagate cloned PCR products and gene constructs and was grown at 37°C in Luria-Bertani broth supplemented with ampicillin (100 μg/ml) or kanamycin (50 μg/ml) as required (Sambrook *et al.*, 1989).

Identification of LPS-related genes from the H. influenzae *genome sequence*

Putative LPS biosynthetic genes had been previously identified by an *in silico* search of the *H. influenzae* genome sequence with heterologous sequences of LPS biosynthetic genes from a wide range of organisms obtained from publicly available databases (Hood *et al.*, 1996a). The RM118*lgtF* locus (HI0653) was identified by searching the Institute for Genomic Research *H. influenzae* strain Rd sequence database (*www.tigr.org/tdb/CMR/ghi/htmls/SplashPage.html*) for matches with the LgtF protein sequence from *N. meningitidis* (GenBank accession no. U58765).

Recombinant DNA methodology, cloning, and mutation

Restriction endonucleases and DNA-modifying enzymes were obtained from Boehringer Mannheim and used according to the manufacturer's instructions. Plasmid DNA preparation, Southern blotting, and hybridization analysis were performed as described by Sambrook *et al.* (1989). Chromosomal DNA was prepared from *Haemophilus* by the method described elsewhere (High *et al.*, 1993).

Apart from *lgtF*, putative *H. influenzae* LPS biosynthetic genes were cloned and mutated as previously reported (Hood *et al.*, 1996a). For the *lgtF* locus, oligonucleotide primers, lgtFa (5′-TGGTGGTGGGCAAGACGC-3′) and lgtFb (5′- AGCCTG-AATTCGACAGCC-3′), amplified a 1461-bp DNA fragment including HI0653 by PCR. PCR conditions were for 1-min periods of denaturation (94°C), annealing (50°C), and polymerization (72°C) for 30 cycles. One microliter of PCR product was ligated with 50 ng of plasmid pT7Blue (Novagen) and transformed into *E. coli* strain DH5α. Recombinant plasmids were confirmed by restriction endonuclease digestion and sequencing from plasmid-specific primers (Hood *et al.*, 1996a). The *lgtF* gene was inactivated by inserting a kanamycin resistance cassette (released by digestion with *Eco*R1 from pUC4Kan, Pharmacia) into a *Mun*I restriction site 257 bp inside the 5′ end of HI0653 to give plasmid pDQ1.

Construction of mutant strains

Two to three micrograms of linearized plasmid, containing mutated LPS biosynthetic genes, was used to transform *H. influenzae* strain RM118 by the MIV procedure (Herriott *et al.*, 1970) and transformants were selected on kanamycin. To construct strain RM118*lic1,* RM118 was transformed with 5 μg of sheared chromosomal DNA isolated from the corresponding RM153 mutant. Strain RM118*lic2A* was constructed by transformation of RM118 with 1 μg of a PCR product including inactivated *lic2A* and the adjacent gene *ksgA* amplified from strain RM153*lic2A*. PCR used the primers L2A (5′-CTCCATATTACATAAT-3′) and L2D (5′-AAACACTTAGGCCATACG-3′) under conditions as described above. All transformants were recultured on appropriate brain heart infusion/antibiotic plates, then were confirmed as mutants by PCR amplification and/or Southern blotting/hybridization of digested chromosomal DNA.

Analysis of LPS by immunoblotting and electrophoresis

LPS isolated from wild type and mutants of *H. influenzae* strain RM118 was analyzed using LPS-specific monoclonal antibodies and by T–SDS–PAGE) as described previously (Hood *et al.*, 1996a).

Structural fingerprinting of LPS

Cells from 10-L batch cultures (10 lots of 1 L) were harvested after overnight growth, LPS was extracted by the hot phenol-water method (Westphal and Jann, 1965), followed by ethanol precipitation as described by Thibault and Richards (2000). LPS was purified and *O*-deacylated as previously described (Holst *et al.*, 1993). Sugars were identified by gas-liquid chromatography mass spectrometry (GLC-MS) as their alditol acetates as previously described (Masoud *et al.*, 1997). Linkage analysis was accomplished following acetylation of the oligosaccharides with acetic anhydride (0.5 ml) and 4-dimethylaminopyridine (0.5 mg) at room temperature for 24 h. Peracetylated material was then treated with methyl iodide in dimethylsulfoxide in the presence of lithium methylsulfinylmethanide to afford the methylated oligosaccharides, which were recovered using a SepPak C18 cartridge and subjected to sugar analysis (Blakeney and Stone, 1985). The relative proportions of the various alditol acetates and partially methylated alditol acetates obtained in sugar and methylation analyses were measured from the detector response of the GLC-MS and are uncorrected. GLC-MS was carried out with a Delsi Di200 chromatograph equipped with a NERMAG R10–10H quadrupole mass spectrometer or with a Varian Iontrap system using a DB-5 fused silica capillary column (25 m × 0.25 mm × 0.25 μm) and a temperature gradient of 160°C (1 min) → 250°C at 3°C/min.

ESI-MS was performed on a VG Quattro Mass Spectrometer (Micromass, Manchester, UK) in the negative-ion mode. Samples were dissolved in water and then mixed in a 1:1 ratio with 50% aqueous acetonitrile containing 1% acetic acid. Sample solutions were injected via a syringe pump into a running solvent of H_2O:CH_3CN (1:1) at a flow rate of 5 μl/min. 1D ^{1}H-NMR spectra were recorded at 500 MHz for solutions in deuterium oxide at 22°C, after several lyophyllizations with D_2O, on a Bruker AMX 500 spectrometer. To enhance spectral resolution, perdeutero-EDTA (2 mM) and perdeutero-SDS (10 mg/ml) were added to the D_2O solutions (Risberg *et al.*, 1997). Chemical shifts are referenced to the methyl proton resonance (δ; 2.225 ppm) of internal acetone.

Analysis of enzymatic activity from LgtC and LgtD

The enzyme encoded by *lgtD* was assayed with the synthetic acceptor FCHASE-P^k using capillary electrophoresis for detection of the product (Wakarchuk *et al.*, 1996). FCHASE-P^k was synthesized from FCHASE-Lac using the *N. meningitidis* LgtC enzyme as previously described (Gotschlich, 1994; Wakarchuk *et al.*, 1998). The reaction conditions were 0.5 mM acceptor, 1 mM UDP-GalNAc, 50 mM HEPES–NaOH (pH 7.0), 10 mM $MgCl_2$, 10 mM $MnCl_2$. Extracts from RM118 and RM118:*lgtD* were made by sonicating the cells, and then collecting the membrane fraction by centrifugation at 100,000 × *g* for 30 min. The product was isolated by thin-layer chromatography as previously described (Wakarchuk *et al.*, 1996). Since the proportion of acceptor converted to product was small, some of the starting material was also isolated. The recovered mixture was divided into two parts and then treated with β-hexosaminidase under conditions recommended by the enzyme supplier (NEB).

Activity for LgtC was below the limits of detection in extracts of RM118, so the *lgtC* gene was cloned into an expression vector and activity was assayed in *E. coli*. The gene was amplified by PCR (as described above) using primers lgtCa (5′GGGGGGCATATGGGACGGACTGTCAGTCAGACAATG) and lgtCb (5′GGGGGGGTCGACTCATTAATTATCTTTTATTCTCTTTCTTAATC). The gene was then inserted into plasmid pCWori plus at the *Nde*I and *Sal*I sites similar to what was described for *lgtC* from *N. meningitidis* (Gotschlich, 1994; Wakarchuk *et al.*, 1998). Crude sonicated extracts of the recombinant clone were assayed with 1 mM FCHASE-Lac, 1 mM UDP-Gal, 10 mM $MnCl_2$, 5 mM dithiothreitol, 50 mM HEPES, pH 7.5. The enzyme was shown to be unstable in *E. coli* and was assayed within a few hours of when the extracts were made. The product of the enzyme reaction was analyzed by specific glycosidase digestion, mass spectrometry, and cochromatography with authentic FCHASE-P^k.

Acknowledgments

We thank Don Krajcarski for assistance in ESI-MS analysis, Jianjun Li for CE-ESI-MS analysis, and Johan Sagemark and Anna-Karin Karlsson for methylation analysis. D.H., S.W., and M.D. were supported by a program grant from the Medical Research Council, UK.

Abbreviations

ESI-MS, electrospray ionization mass spectrometry; GLC-MS, gas-liquid chromatography mass spectrometry; Kdo, 2-keto-3-deoxyoctulosonic acid; LPS, lipopolysaccharide; LPS-OH, *O*-deacylated lipopolysaccharide; MAb, monoclonal antibody; MS-MS, tandem mass spectrometry; NMR, nuclear magnetic resonance; *P*, free phosphate group; PCR, polymerase chain reaction *P*Etn, phosphoethanolamine; *PP*Etn, pyrophosphoethanolamine; *P*Cho, phosphocholine; T–SDS–PAGE, tricine–sodium dodecyl sulfate–polyacrylamide gel electrophoresis.

References

Blakeney, A.B. and Stone, B.A. (1985) Methylation of carbohydrates with lithium methylsulphinyl carbanion. *Carbohydr. Res.*, **140**, 319–324.

Brook, J.S. and Valvano, M.A. (1996) Molecular cloning of the *Haemophilus influenzae gmhA* (*lpcA*) gene encoding a phosphoheptose isomerase required for lipooligosaccharide biosynthesis. *J. Bacteriol.*, **178**, 3339–3341.

Campbell, J.A., Davies, G.J., Bulone, V., and Henrissat, B. (1997) A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. *Biochem. J.*, **326**, 929–942.

Cope, L.D., Yogev, R., Mertsola, J., Argyle, J.C., McCracken, G.H., and Hansen, E.J. (1990) Effect of mutation in lipopolysaccharide biosynthetic genes on virulence of *Haemophilus influenzae* type b. *Infect. Immun.*, **58**, 2343–2351.

Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Butt, C.J., Tomb, J-F., Dougherty, B.A., Merrick, J.M., and others (1995) The genome of *Haemophilus influenzae* Rd. *Science*, **269**, 496–512.

Gibson, B.W., Melaugh, W., Phillips, N.J., Apicella, M.A., Campagnari, A.A., and Griffiss, J.M. (1993) Investigation of the structural heterogeneity of lipooligosaccharides from pathogenic *Haemophilus* and *Neisseria* species and of R-type lipopolysaccharides from *Salmonella typhimurium* by electrospray mass spectrometry. *J. Bacteriol.*, **175**, 2702–2712.

Gotschlich, E.C. (1994) Genetic locus for the biosynthesis of the variable portion of *Neisseria gonorrhoeae* lipooligosaccharide. *J. Exp. Med.*, **180**, 2181–2190.

Helander, I.M., Lindner, B., Brade, H., Altmann, K., Lindberg, K.K., Rietschel, E.T., and Zahringer, U. (1988) Chemical structure of the lipopolysaccharide of *Haemophilus influenzae* strain I-69 Rd^{-}/b^{+}: Description of a novel deep-rough chemotype. *Eur. J. Biochem.*, **177**, 483–492.

Herriott, R.M., Meyer, E.M., and Vogt, M.J. (1970) Defined non-growth media for stage II development of competence in *Haemophilus influenzae*. *J. Bacteriol.*, **101**, 517–524.

High, N.J., Deadman, M.E., and Moxon, E.R. (1993) The role of the repetitive DNA motif (5′-CAAT-3′) in the variable expression of the *Haemophilus influenzae* lipopolysaccharide epitope Galα(1-4)βGal. *Mol. Microbiol.*, **9**, 1275–1282.

Holst, O., Broer, W., Thomas-Oates, J.E., Mamat, U., and Brade, H. (1993) Structural analysis of two oligosaccharide bisphosphates isolated from the lipopolysaccharide of a recombinant strain of *Escherichia coli* F515 (Re chemotype) expressing the genus-specific epitope of *Chlamydia* lipopolysaccharide. *Eur. J. Biochem.*, **214**, 703–710.

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

Hood, D.W., Deadman, M.E., Jennings, M.P., Bisceric, M., Fleischmann, R.D., Venter, J.C., and Moxon, E.R. (1996b) DNA repeats identify novel virulence genes in *Haemophilus influenzae*. *Proc. Natl. Acad. Sci. USA*, **93**, 11121–11125.

Jarosik, G.P. and Hansen, E.J. (1994) Identification of a new locus involved in the expression of *Haemophilus influenzae* type b lipooligosaccharide. *Infect. Immun.*, **62**, 4861–4867.

Kahler, C.M., Carlson, R.W., Rahman, M.M., Martin, L.E., and Stephens, D.S. (1996) Two glycosyltransferase genes, *lgtF* and *rfaK*, constitute the lipooligosaccharide ice (Inner Core Extension) biosynthesis operon of *Neisseria meningitidis*. *J. Bacteriol.*, **178**, 6677–6684.

Kimura, A. and Hansen, E.J. (1986) Antigenic and phenotypic variants of *Haemophilus influenzae* type b lipopolysaccharide and their relationship in virulence. *Infect. Immun.*, **51**, 60–79.

Lysenko, E., Richards, J.C., Cox, A.D., Stewart, A., Martin, A., Kapoor, M., and Weiser, J.N. (2000) The position of phosphorylcholine on the lipopolysaccharide of *Haemophilus influenzae* affects binding and sensitivity to C-reactive protein-mediated killing. *Mol. Microbiol.*, **35**, 234–245.

Masoud, H., Moxon, E.R., Martin, A., Krajcarski, D., and Richards, J.C. (1997) Structure of the variable and conserved lipopolysaccharide oligosaccharide epitopes expressd by *Haemophilus influenzae* serotype b strain Eagan. *Biochem.*, **36**, 2091–2103.

Moxon, E.R. and Maskell, D.J. (1992) "*Haemophilus influenzae* lipopolysaccharide: the biochemistry and biology of a virulence factor." In Hormaeche, C.E., Penn, C.W., Smythe, C.J. (eds.), *Molecular biology of bacterial infection, current status and future perspectives*. Society for General Microbiology Symposium 49. Cambridge University Press, Cambridge, UK.

Phillips, N.J., Apicella, M.A., Griffiss, J.M., and Gibson, B.W. (1992) Structural characterization of the cell surface lipooligosaccharides from a non-typable strain of *Haemophilus influenzae*. *Biochem.*, **31**, 4515–4526.

Phillips, N.J., Apicella, M.A., Griffiss, J.M., and Gibson, B.W. (1993) Structural studies of the lipooligosaccharides from *Haemophilus influenzae* type b strain A2. *Biochem.*, **32**, 2003–2012.

Phillips, N.J., McLaughlin, R., Miller, T.J., Apicella, M.A., and Gibson, B.W. (1996) Characterization of two transposon mutants of *Haemophilus influenzae* type b with altered lipooligosaccharide biosynthesis. *Biochem.*, **35**, 5937–5947.

Preston, A., Mandrell, R.E., Gibson, B.W., and Apicella, M.A. (1996a) The lipopolysaccharides of pathogenic Gram-negative bacteria. *Crit. Rev. Microbiol.*, **22**, 139–180.

Preston, A., Maskell, D., Johnson, A., and Moxon, E.R. (1996b) Altered lipopolysaccharide characteristic of the I69 phenotype in *Haemophilus influenzae* results from mutation in a novel gene, *isn. J. Bacteriol.*, **178**, 396–402.

Rahman, M.M., Gu, X-X., Tsai, C-M., Kolli, V.S.K., and Carlson, R.W. (1999) The structural heterogeneity of the lipooligosaccharide (LOS) expressed by pathogenic non-typeable *Haemophilus influenzae* strain NTHi 9274. *Glycobiology*, **9**, 1371–1380.

Risberg, A., Schweda, E.K.H., and Jansson, P.-E. (1997) Structural studies of the cell-envelope oligosaccharide from the lipopolysaccharide of *Haemophilus influenzae* strain RM.118-28. *Eur. J. Biochem.*, **243**, 701–707.

Risberg, A., Alvelius, G., and Schweda, E.K.H. (1999a) Structural analysis of the lipopolysaccharide oligosaccharide epitopes expressed by *Haemophilus influenzae* strain RM.118-26. *Eur. J. Biochem.*, **265**, 1067–1074.

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

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular cloning; a laboratory manual*, 2d ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Schweda, E.K.H., Hegedus, O.E., Borrelli, S., Lindberg, A.A., Weiser, J.W., Maskell, D.J., and Moxon, E.R. (1993) Structural studies of the saccharide portion of cell envelope lipopolysaccharide from *Haemophilus influenzae* strain AH1-3 ($lic3^{+}$). *Carbohydr. Res.*, **246**, 319–330.

Schweda, E.K.H., Jansson, P.-E., Moxon, E.R., and Lindberg, A.A. (1995) Structural studies of the saccharide part of the cell envelope lipooligosaccharide from *Haemophilus influenzae* strain *galEgalK*. *Carbohydr. Res.*, **272**, 213–224.

Thibault, P. and Richards, J.C. (2000) Application of combined capillary electrophoresis electrospray mass spectrometry in the characterisation of short chain lipopolysaccharides: *Haemophilus influenzae*. In Holst, O. (ed.), *Methods of molecular biology, vol. 45, Bacterial toxins: methods and protocols*. Humana Press, Totowa, NJ.

Virji, M., Weiser, J.N., Lindberg, A.A., and Moxon, E.R. (1990) Antigenic similarities in lipopolysaccharides of *Haemophilus* and *Neisseria* and expression of a digalactoside structure also present on human cells. *Microb. Pathogen.*, **9**, 441–450.

Wakarchuk, W., Martin, A., Jennings, M.P., Moxon, E.R., and Richards, J.C. (1996) Functional relationships of the genetic locus encoding the glycosyl transferase enzymes involved in the expression of the lacto-N-neotetraose terminal lipopolysaccharide structure in *Neisseria meningitidis*. *J. Biol. Chem.*, **271**, 19166–1917.

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

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

Weiser, J.N., Love, J.M., and Moxon, E.R. (1989) The molecular mechanism of phase variation of *Haemophilus influenzae* lipopolysaccharide. *Cell*, **59**, 657–665.

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

Weiser, J.N., Williams, A., and Moxon, E.R. (1990b) Phase-variable lipopolysaccharide structures enhance the invasive capacity of *Haemophilus influenzae*. *Infect. Immun.*, **58**, 3455–3457.

Weiser, J.N., Shchepetov, M., and Chong, S.T.H. (1997) Decoration of lipopolysaccharide with phosphorylcholine: a phase variable characteristic of *Haemophilus influenzae*. *Infect. Immun.*, **65**, 943–950.

Westphal, O. and Jann, K. (1965) Bacterial lipopolysaccharides. Extraction with phenol-water and further applications of the procedure. *Meth. Carbohydr. Chem.*, **5**, 83–91.

White, K.A. and Raetz, C.R.H. (1998) Complementation of a genomic disruption of the *E.coli* bifunctional KDO transferase by the *H. influenzae* monofunctional KDO transferase. *FASEB J.*, **12**, L44.

White, K.A., Lin, S., Cotter, R.J., and Raetz, C.R.H. (1999) A *Haemophilus influenzae* gene that encodes a membrane bound 3-deoxy-D-manno-octulosonic acid (Kdo) kinase: Possible involvement of KDO phosphorylation in bacterial virulence. *J. Biol. Chem.*, **274**, 31391–31400.

Zamze, S.E. and Moxon, E.R. (1987) Composition of the lipopolysaccharide from different capsular serotype strains of *Haemophilus influenzae*. *J. Gen. Microbiol.*, **133**, 1443–1451.