Available online at www.sciencedirect.com
ScienceDirect
Carbohydrate Research 343 (2008) 1424–1434
Carbohydrate RESEARCH

Structural elucidation of the major Hex4 lipopolysaccharide glycoform from the *lgtC* mutant of *Haemophilus influenzae* strain Eagan

Hussein Masoud,[a,b] E. Richard Moxon[c] and James C. Richards[a,*]

[a]*Institute for Biological Sciences, National Research Council of Canada, Ottawa, Ontario, Canada K1A 0R6*
[b]*Department of Biological Sciences, University of Jordan, 11942 Amman, Jordan*
[c]*Molecular Infectious Diseases Group and Department of Paediatric, Institute for Molecular Medicine, John Radcliffe Hospital, Headington, Oxford 0X3 3DU, UK*

Received 28 December 2007; received in revised form 18 February 2008; accepted 1 April 2008
Available online 7 April 2008

Abstract—Lipopolysaccharide (LPS) oligosaccharide epitopes are major virulence factors of *Haemophilus influenzae*. The structure of LPS glycoforms of *H. influenzae* type b strain Eagan containing a mutation in the gene *lgtC* is investigated. *LgtC* is involved in the biosynthesis of globoside trisaccharide [α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→], an LPS epitope implicated in the virulence of this organism. Glycose and methylation analyses provided information on the composition while electrospray ionization mass spectrometry (ESI-MS) on O-deacylated LPS (LPS-OH) indicated the major glycoform to contain 4 hexoses attached to the common *H. influenzae* triheptosyl inner-core unit. The structure of the Hex4 glycoform in LPS-OH and core oligosaccharide samples was determined by NMR. It consists of an L-α-D-HepIII*p*-(1→2)-[*P*Etn→6]-L-α-D-HepII*p*-(1→3)-L-α-D-HepI*p*-(1→5)-[*P*→4]-α-D-Kdo*p*-(2→ to which a β-D-Glc*p*-(1→4)-α-D-Glc*p* disaccharide unit is extended from HepII at the C-3 position, while HepI and HepIII are substituted at the C-4 and C-2 positions with β-D-Glc*p* and β-D-Gal*p*, respectively. This structure corresponds to that expressed as a subpopulation in the parent strain. ^{31}P NMR studies permitted the identification of subpopulations of LPS containing Kdo substituted at the C-4 position with monophosphate or pyrophosphoethanolamine (*PP*Etn). HepIII was found to be substituted with either phosphate at the C-4 position or acetate at the C-3 position, but not both of them together in the same subpopulation. The subpopulations containing phosphate and acetate at HepIII and their location have not previously been reported.

Keywords: *Haemophilus influenzae*; Lipopolysaccharide; Structural analysis; lgtC mutant

1. Introduction

Haemophilus influenzae remains a significant cause of bacterial infections in infants and young children worldwide. Encapsulated organisms, particularly those expressing the type b polysaccharide, cause meningitis and pneumonia,[1,2] while acapsular (non-typeable) strains cause otitis media and upper respiratory tract infections.[3] Lipopolysaccharide (LPS) is an important virulence factor and the structure of this molecule has been extensively investigated in a number of strains (for a review, see Ref. 4). In every strain investigated to date, the low-molecular-mass LPS (sometimes referred to as lipooligosaccharide) is composed of a

Abbreviations: DOC-PAGE, deoxycholate-polyacrylamide gel electrophoresis; GLC–MS, gas liquid chromatography–mass spectrometry; ESI-MS, electrospray ionization-mass spectrometry; COSY, correlated spectroscopy; TOCSY, total correlated spectroscopy; HMQC, heteronuclear multiple quantum coherence; NOE, nuclear Overhauser effect; NOESY, two-dimensional NOE spectroscopy; Kdo, 3-deoxy-D-*manno*-octulosonic acid; LPS, lipopolysaccharide; LPS-OH, O-deacylated LPS; OS, oligosaccharide; Etn, ethanolamine; *P*Etn, phosphoethanolamine; *PP*Etn, pyrophosphoethanolamine: and *P*Cho: phosphocholine.

* Corresponding author. Tel.: +1 613 993 7506; fax: +1 613 957 7867; e-mail: jim.richards@nrc.ca

0008-6215/$ - see front matter Crown Copyright © 2008 Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.carres.2008.04.002

heptosyl trisaccharide unit attached to the lipid A region through a pyrophosphorylated 3-deoxy-D-*manno*-octulosonic acid (Kdo) in which the middle heptose (HepII) is substituted with phosphoethanolamine (*P*Etn) at the C-6 position. Each of the heptose residues in the inner core can provide a point for further extension by hexose residues forming the heterogeneous outer-core region. The proximal heptose (HepI) has been found to be invariably substituted at the C-4 position with a β-D-glucose, and further oligosaccharide extension has been observed in several non-typeable strains.[5] In type b strains Eagan and RM7004, HepII is substituted at the C-3 position with an α-D-Glc, which can be further extended by the globoside trisaccharide [α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→4)-β-D-Glc*p*(1→,[6,7] while in the non-typeable strain Rd, oligosaccharide extension from HepII does not occur.[8] The distal heptose (HepIII) has been found to be substituted at the C-2 or C-3 position with β-D-Gal[6,7,9] or β-D-Glc.[10,11] Further chain extension can occur when a β-D-Glc substitutes HepIII, while β-D-Gal has been found only as a terminal residue. *H. influenzae* LPS can also be substituted by non-carbohydrate substituents such as phosphocholine (*P*Cho), *P*Etn, and O-linked acyl groups.[5–8,10–23] A genetic basis for the biosynthesis of *H. influenzae* LPS has been developed.[24–27] Many of the genes involved in the addition of outer-core extensions are phase variable and this contributes to the observed LPS heterogeneity both within clonal populations and in epidemiologically distinct *H. influenzae* strains. Phase variation is thought to provide an advantage for pathogenic bacteria to avoid host immune responses. In *H. influenzae* strain Rd two phase variable genes, *lic2A* and *lgtC*, have been implicated in the biosynthesis of the trisaccharide, α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→4)-β-D-Glc*p*(1→.[25] *Lic2A* is responsible for the addition of β-D-Gal*p* in a 1,4 linkage to β-D-Glc*p* to give lactose which serves as the acceptor for *lgtC*, which transfers α-D-Gal*p* from UDP-Gal to give the globoside trisaccharide.[25] Both genes are present in *H. influenzae* strain Eagan for which similar functions have been inferred.[24,25] Here, we report the structure of LPS glycoforms expressed by a *lgtC* mutant in strain Eagan, which contains a β-D-Glc*p*-(1→4)-β-D-Glc*p*(1→ as the major oligosaccharide extension from HepII indicating that *lic2A* is phased off in this mutant. Moreover, in this LPS glycoform, major subpopulations containing phosphate or acetyl substituents at C-4 or C-3 of the distal heptose residue, respectively, were characterized.

2. Results and discussion

2.1. Isolation and characterization of LPS

H. influenzae strain Eagan (RM 153) containing a mutation in *lgtC*[24] was cultivated in brain heart infusion (BHI) broth supplemented with NADH, Hemin, and kanamycin (to maintain selection pressure), and the LPS was isolated in a yield of ca. 2–3% from dried bacterial cells by using the hot phenol-water extraction method.[28] Deoxycholate-polyacrylamide gel electrophoresis (DOC-PAGE) analysis of the LPS from *lgtC* mutant revealed a major band having electrophoretic mobility similar to the low-molecular-weight band LPS (Hex4 glycoform) from the parent strain (data not shown). Another minor band with faster electrophoretic mobility was also observed indicating the presence of a minor LPS fraction with lower molecular weight.

Partial acid hydrolysis of the LPS sample with acetic acid afforded core OS and insoluble lipid A. Glycose analysis of the core OS components as their aldital acetate and acetylated (*R*)-2-octyl glycoside derivatives revealed the presence of D-Glc, D-Gal, and L-*glycero*-D-*manno*-heptose (Hep). GLC–MS analysis of the partially methylated glycoses as expected indicated the presence of 2-substituted Hep*p* and 3,4-disubstituted Hep*p* residues from the triheptosyl inner-core unit. In addition, trace amounts of the acetylated derivatives of 4,6,7-trimethyl-Hep*p*, 4,6-dimethyl Hep*p*, and 4,7-dimethyl Hep*p* were observed. This was consistent with the presence of a 2,3-disubstituted Hep*p* substituted by a phosphate substituent at C-6 or C-7 where the latter two derivatives arise from the migration of the *P*Etn substituent known to be present in the parent strain[6] under the strongly alkaline conditions used in the methylation procedure. Terminal non-reducing Glc*p* and Gal*p* residues as well as 4-substituted Glc*p* were also detected as the major glycose units present in the core OS.

2.2. Characterization of O-deacylated LPS

O-Deacylated LPS (LPS-OH) was prepared by the treatment of the LPS sample with anhydrous hydrazine (37 °C, 1 h). Electrospray ionization-mass spectrometry (ESI-MS) in the negative ion mode showed several triply and doubly charged molecular ions indicative of molecular species containing the conserved *P*Etn-substituted inner-core heptose trisaccharide basal unit attached to the lipid A-OH moiety via a phosphorylated Kdo residue (Table 1).[6] The triply charged ion at *m*/*z* 865.9 with the corresponding doubly charged ion at *m*/*z* 1299.4 indicated the presence of a Hex4 glycoform having the composition $Hex_4.Hep_3.PEtn_1{\cdot}P_1{\cdot}Kdo_1{\cdot}$LipidA-OH. A parallel series of triply and doubly charged ions shifted 80, 123, and 203 amu to higher mass could be attributed to Hex4 glycoforms containing additional phosphate, *P*Etn, and both phosphate and *P*Etn groups, respectively (Table 1). Minor signals with triply charged ions at *m*/*z* 811.9 and 852.6 were attributed to Hex3 glycoforms containing one and two *P*Etn groups, respectively.

The structure of the major LPS-OH Hex4 glycoform was elucidated by detailed ^{1}H, ^{13}C, and ^{31}P NMR

Table 1. Negative ion ESI-MS data and proposed compositions for O-deacylated LPS from the *lgtC* mutant of *H. influenzae* strain Eagan

LPS glycoform	Observed	Ion (*m/z*)	Molecular weight (Da)		Proposed composition
	$(M-3H)^{3-}$	$(M-2H)^{2-}$	Observed[a]	Calculated[b]	
Hex 4	865.9	1299.4	2600.7	2601.3	Hex_4 Hep_3 *P*Etn$_1$ P_1 Kdo_1 lipidA-OH
	892.4	—	2680.4	2681.3	Hex_4 Hep_3 *P*Etn$_1$ P_2 Kdo_1 lipidA-OH
	907.0	1360.9	2723.7	2724.4	Hex_4 Hep_3 *P*Etn$_2$ P_1 Kdo_1 lipidA-OH
	933.3	—	2803.8	2804.4	Hex_4 Hep_3 *P*Etn$_2$ P_2 Kdo_1 lipidA-OH
Hex 3[c]	811.9	1218.7	2438.6	2439.2	Hex_3 Hep_3 *P*Etn$_1$ P_1 Kdo_1 lipidA-OH
	852.6	–	2560.8	2562.2	Hex_3 Hep_3 *P*Etn$_2$ P_1 Kdo_1 lipidA-OH

[a] Average of the observed doubly and triply charged ions.
[b] Average mass units were used for the calculation of molecular weight values based on the proposed compositions as follows: Hex, 162.15; Hep, 192.17; Kdo, 220.18; phosphate, 79.98; *P*Etn, 123.05; LipidA-OH, 953.03.[13]
[c] Minor signals.

analysis. Partial assignments of the ^{1}H and ^{13}C NMR resonances of the LPS-OH were made by COSY, TOCSY, and HMQC experiments, and the data are recorded in Tables 2 and 3. Nine major anomeric ^{1}H resonances were observed in the low-field region of the spectrum (Fig. 1) and subspectra corresponding to each of the glycosyl residues were identified in the COSY and TOCSY spectra (Fig. 2A and B), and by comparison with the corresponding glycosyl residues from previous studies.[6,7] Three subspectra corresponding to the heptose residues (HepI-HepIII) were identified on the basis of the observed small $J_{1,2}$ (<1.0 Hz) and $J_{2,3}$ (≈4.0 Hz) values, which were indicative of the *manno*-pyranose ring system. The anomeric ^{1}H and ^{13}C chemical shift values indicated that each of the Hep ring system has the L-α-D-configuration, and this was confirmed by the occurrence of a single intraring NOE between the respective H-1 and H-2 resonances. Three subspectra having the *gluco*-configuration (labeled GI, GII, and GIII) were identified on the basis of the large magnitude of the vicinal coupling constants $J_{2,3}$, $J_{3,4}$, and $J_{4,5}$ (8–10 Hz) of which GI and GIII have the β-configuration ($J_{1,2}$ = 8.1 Hz) and GII, the α-configuration ($J_{1,2}$ = 4.3 Hz). In addition, the subspectrum for a β-D-Gal*p* (labeled Gal) was identified from the magnitude of the vicinal couplings of $J_{1,2}$(8.1 Hz), $J_{2,3}$(9.2 Hz), $J_{3,4}$(4.1 Hz), and $J_{4,5}$ (≈1.0 Hz). The glucosamine disaccharide component of lipid A moiety, attributed to the two residues labeled GNI and GNII, showed similar chemical shifts to the parent strain.[6,7] The methylene proton resonances of Kdo (H-3a: 2.00/1.95 ppm and H-3 e: 2.40/2.38 ppm) gave access to Kdo spin systems which were assigned from the COSY and TOCSY spectra (Table 2). The observed doubling of the methylene proton resonances is most likely due to different populations of LPS having phosphate or *PP*Etn substitution at C-4 of Kdo as previously observed.[6]

The sequence of the glycoses within the core OS was determined from transglycosidic proton NOE connectivities between anomeric and aglyconic protons on adjacent glycosidically linked residues. Part of the NOESY contour plot is shown in Figure 2C. Intense transglycosidic NOE connectivities between HepIII H-1 and HepII H-2, HepII H-1 and HepI H-3, HepI H-1 and Kdo H-5 resonances established the linear sequence of triheptosyl region of the inner core as L-α-D-Hep*p*-(1→2)-L-α-D-Hep*p*-(1→3)-L-α-D-Hep*p*-(1→5)-α-D-Kdo*p*. Intraresidue NOEs were observed between H-1 and H-2 of each of the heptose units, confirming the assigned α–D-configuration. Transglycosidic NOE between H-1 of GlcI and HepI H-4/H-6 protons, in combination with the methylation results, established the substitution of HepI at C-4 by a terminal β-D-Glc*p* residue. A disaccharide β-D-Glc*p*-(1→4)-α-D-Glc*p* substituting HepII at C-3 was established by the observed interresidue NOEs between GlcIII H-1/GlcII H-4 and GlcII H-1/HepII H-3. Moreover, HepIII is substituted at C-2 by a terminal β-D-Gal*p*, which was established from a transglycosidic NOE between Gal H-1 and HepIII H-2. Interresidue NOEs were observed between the anomeric proton resonances (H-1) of the three residues HII, HIII, and Gal (Fig. 2C), an indication of the proximity (<3 Å) of these anomeric protons in this glycoform. Thus, the combined results from compositional, methylation, ESI-MS, and NMR analyses establish the sequence of the Hex 4 LPS glycoform to be

β-D-Glc*p*-(1→4)-L-α-D-Hep*p*-(1→5)-α-Kdo*p*-(2→6)-Lipid A-OH
3
↑
1
β-D-Glc*p*-(1→4)-α-D-Glc*p*-(1→3)-L-α-D-Hep*p*
2
↑
1
β-D-Gal*p*-(1→2)-L-α-D-Hep*p*

Table 2. Proton chemical shifts (ppm) and coupling constants (Hz) for O-deacylated LPS from the *lgtC* mutant from *H. influenzae* strain Eagan

Residue	Glycose unit	H-1 (J_{1-2})	H-2 (J_{2-3})	H-3 (J_{3-4})	H-3′ ($J_{3'-4}$)	H-4 (J_{4-5})	H-5	H-6 (J_{5-6})	H-6′ ($J_{5-6'},J_{6-6'}$)	H-7 (J_{6-7})	H-7′ ($J_{6-7},J_{7-7'}$)
GNI	→6)-α-D-GlcNp-(1→	5.40	3.83		—						
GNII	→4,6)-β-D-GlcNp-(1→	4.59	3.89	3.79	—	3.98					
K	→4,5)-α-Kdop-(2→	—	—	2.00	2.40	4.80	4.32	3.95	—	3.71	
K′	→4,5)-α-Kdop-(2→	—	—	1.95	2.38	4.55	4.28		—		
HI	→3,4)-L-α-D-Hepp-(1→	5.15 (~1.0)	4.18	4.10	—	4.28	4.20	4.06	—	3.95	3.83
HII	→2,3,6)-L-α-D-Hepp-(1→	5.66 (−1.0)	4.30	4.13	—	4.08	3.80	4.58	—	3.87	3.69
HIII	→2)-L-α-D-Hepp-(1→	5.09 (~1.0)	4.01		—				—		
GI	β-D-Glcp-(1→	4.52 (8.1)	3.32 (8.7)	3.52 (10.3)	—	3.42 (9.7)	3.51	3.92 (5.4)	3.74 (−, 10.8)		
GII	→4)-α-D-Glcp-(1→	5.35 (4.3)	3.59 (8.7)	3.89 (8.7)	—	3.69 (8.7)					
GIII	β-D-Glcp-(1→	4.53 (8.1)	3.34 (8.7)	3.46 (10.3)	—	3.42	3.49	3.92	3.74		
Gal	β-D-Galp-(1→	4.35 (8.1)	3.58 (9.2)	3.67 (4.1)	—	3.92 (−1.0)	3.69	3.82	3.75		
	*P*Etn[a]	4.13[c] (5.9, $J_{P,H}$ ~ 10.3)	3.31[d] (9.5)								
	*PP*Etn[b]	4.27[c]	3.37[d]								

Residue labeled with prime is present in glycoform with *PP*Etn at C-4 of Kdo.
[a] *P*Etn attached to C-6 of HepII.
[b] *PP*Etn attached to C-4 of Kdo.
[c] ^{1}H chemical shifts for methylene groups attached to phosphates.
[d] ^{1}H chemical shifts for methylene groups attached to amino groups.

The location of the phosphate subsistent in the LPS-OH Hex 4 glycoform was determined by ^{31}P NMR. Broad band proton decoupled ^{31}P NMR showed seven major ^{31}P resonances at 4.7 ppm, 3.7 ppm, 2.0 ppm, 0.7 ppm, −0.1 ppm, −10.9 ppm (d, $J_{p,p} \approx 20$ Hz), and −11.5 ppm (d, $J_{p,p} \approx 20$ Hz). The chemical shifts and the doublet pattern of the two high-field resonances (−10.9 ppm, and −11.5 ppm) were indicative for pyrophosphate diphosphodiester moieties,[29–31] while the remaining resonances could be attributed to monophosphate groups. In the ^{1}H–^{31}P chemical shift correlation experiment, the ^{31}P-resonance at −10.9 ppm showed a correlation to the methylene protons of *PP*Etn group at 4.27 ppm; however, the sensitivity of the experiment was insufficient to establish a correlation for the diphosphodiester ^{31}P-resonance at −11.5 ppm. In the parent strain, the *PP*Etn has been shown to substitute C-4 of the Kdo residue.[6] The ^{31}P signal at 2.0 ppm was identified as two overlapping resonances for the partial substitution of C-4 of the Kdo residue with monoester phosphate and a glycosidic phosphate group for GlcNI in the lipid A region as evidenced by correlations with H-4 (4.80 ppm) of Kdo and H-1 (5.4 ppm) of GlcNI, respectively. Correspondingly, a correlation was observed between H-1 and H-2 of GlcNI in the ^{31}P 2D X-filtered TOCSY experiment. Thus, the NMR data are consistent with a subpopulation of LPS-OH that carry either phosphate or *PP*Etn at the C-4 position of Kdo. An intense ^{31}P resonance at −0.1 ppm was correlated in ^{1}H-^{31}P HMQC to H-6 (4.58 ppm) of HepII and to the methylene protons (4.13 ppm) of *P*Etn. These assignments were confirmed by the observed correlation between H-6 and H-7 protons of HepII and between methylene protons of the ethanolamine moiety of *P*Etn in ^{31}P 2D X-filtered TOCSY experiment. The ^{31}P-resonance at 4.7 ppm was correlated with H-4 of GlcNII, which was confirmed by the correlation between H-4 and H-3 of GlcNII in the ^{31}P 2D X-filtered TOCSY. The ^{31}P-resonance at 0.7 ppm showed a correlation with an unassigned proton resonance at 4.04 ppm in the ^{1}H-^{31}P HMQC. ESI-MS had revealed the presence of glycoforms containing an additional monophosphate substituent, and NMR analysis of a core oligosaccharide sample (see below) permitted the assignment of its location at C-4 of HepIII. It is therefore likely that the ^{1}H resonance at 4.04 ppm is H-4 of HepIII in the LPS-OH sample and that the ^{31}P resonance at 0.7 ppm maps to that resonance in the ^{1}H–^{31}P HMQC. The ^{31}P resonance at 3.7 ppm was not resolved in ^{1}H–^{31}P HMQC, possibly due to heterogeneity of the LPS-OH structure. Thus the structure of the major Hex 4 glycoform in the LPS-OH sample is

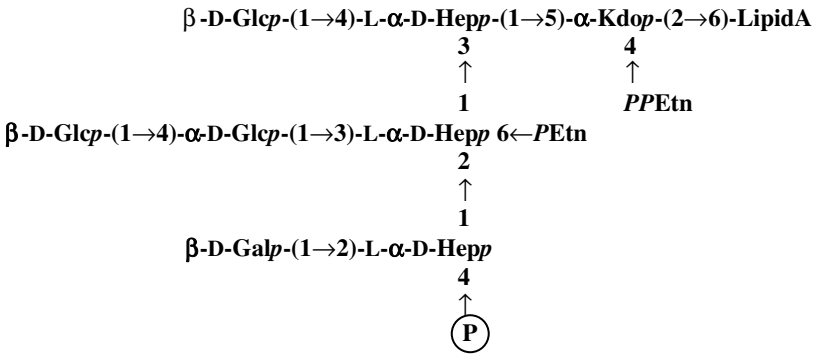


Table 3. ^{13}C-chemical shifts (ppm) for O-deacylated LPS from the *lgtC* mutant of *H. influenzae* strain Eagan

Residue	Glycose unit	C-1	C-2	C-3	C-4	C-5	C-6	C-7	C-8
GNI	→6)-α-D-GlcN*p*-(1→	94.0	55.2						
GNII	→4,6)-β-D-GlcN*p*-(1→	103.4	56.0	71.5					
K	→4,5)-α-Kdo*p*-(2→			35.2		72.6	72.2	70.0	
K′	→4,5)-α-Kdo*p*-(2→					74.1			
HI	→3,4)-L-α-D-Hep*p*-(1→	101.0	71.3	74.6	74.1	73.0	69.5	64.7	
HII	→2,3,6)-L-α-D-Hep*p*-(1→	99.5	79.4	79.2	69.5	72.1	74.5	62.7	
HIII	→2)-L-α-D-Hep*p*-(1→	100.5	79.2						
GI	β-D-Glc*p*-(1→	103.6	74.5	76.8	70.5	77.2	61.8		
GII	→4)-α-D-Glc*p*-(1→	101.1	72.8	72.7	79.5				
GIII	β-D-Glc*p*-(1→	103.6	74.5	76.8	70.5	77.2	61.8		
Gal	β-D-Gal*p*-(1→	103.6	71.6	73.5	69.6	76.3	62.0		
	*P*Etn[a]	63.0[c]	41.5[d]						
	*PP*Etn[b]	63.6[c]	41.5[d]						

Residue labeled with prime is present in glycoform with *PP*Etn at C-4 of Kdo.
[a] *P*Etn attached to C-6 of HepII.
[b] *PP*Etn attached to C-4 of Kdo.
[c] ^{13}C chemical shifts for methylene groups attached to phosphates.
[d] ^{13}C chemical shifts for methylene groups attached to amino groups.

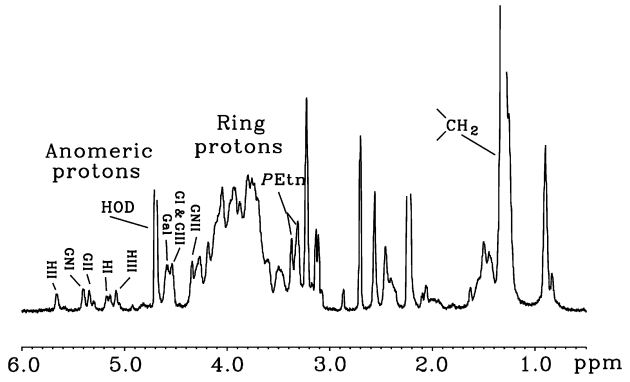


Figure 1. ^{1}H NMR spectrum of O-deacylated LPS from the *H. influenzae* strain Eagan *lgtC* mutant comprising a major Hex 4 glycoform recorded in D_2O at pD = 6–7 and 27 °C. The anomeric protons, ring protons, and methyl resonances are indicated.

2.3. Characterization and location of *O*-acetyl substituents in the core OS

ESI-MS analysis of the core OS obtained following mild acid hydrolysis of LPS from the *lgtC* mutant in the positive ion mode gave doubly and singly charged ions at *m/z* 834.6/1667.7 and 874.5/1747.1 with the compositions $Hex_4{\cdot}Hep_3{\cdot}PEtn_1{\cdot}P_{1-2}{\cdot}Kdo$ corresponding to the major Hex4 glycoform. In addition, abundant ions were observed due to mono- and di-sodiated counterparts. An intense singly charged ion at *m/z* 1630.2 indicated a Hex4 glycoform bearing a single *O*-acetyl group. This species was not detected in the LPS-OH sample since hydrazine removes all *O*-acyl groups. The location of the *O*-acetyl group was determined by the analysis of the core OS by ^{1}H and ^{13}C NMR which also served to confirm the structure of the major Hex4 glycoform in the LPS-OH sample. The ^{1}H- and ^{13}C-resonances were assigned by COSY, TOCSY, and HMQC experiments as described for the LPS-OH sample, and the data are presented in Tables 4 and 5. The anomeric ^{1}H-resonances (5.8–4.3 ppm) revealed heterogeneity particularly for HepIII (Fig. 3), which could be to a large extent attributed to variation on glycoform substitution with phosphate or acetate groups (Table 4). The ^{1}H methyl resonance at 2.20 ppm was indicative of the acetyl substituted glycoform population. The chemical shifts for ^{1}H and ^{13}C resonances of HepIII (HIII′) significantly differed in the *O*-acetyl-free and *O*-acetyl populations.

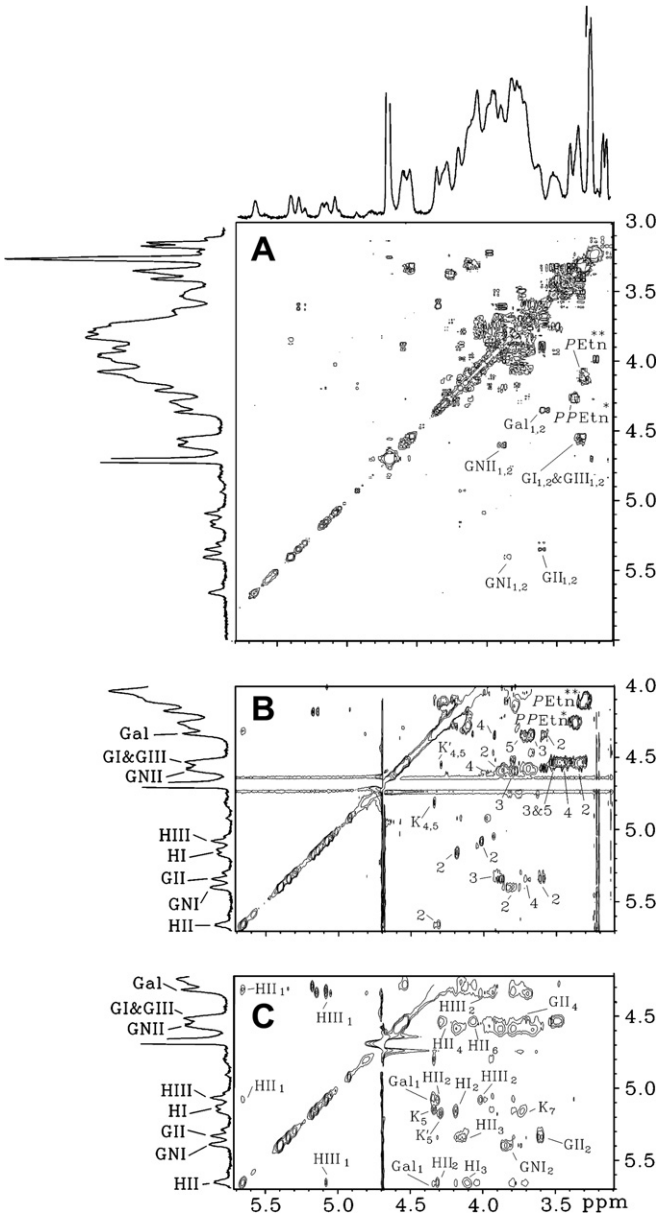


Figure 2. Partial 2D COSY (A), TOCSY (B), and NOESY (C) spectra of the O-deacylated LPS from *H. influenzae* strain Eagan *lgtC* mutant. Cross-peaks relating anomeric and ring protons, and NOE connectivities are indicated in spectra A, B, and C. Cross-peaks labeled by one asterisk are from the correlation of methylene protons of *PP*Etn substituting Kdo. Cross-peaks labeled by two asterisks are from the correlation of methylene protons of *P*Etn substituting C-6 of HepII. Cross peaks relating H-1 and H-2 of the heptose residues (HI, HII, and HIII) were weak in the COSY spectrum and are not visible at the indicated contour level. Abbreviations are as follows: HepI-HepIII, HI-HIII; GlcI-GlcIII,GI-GIII; and Kdo, K. Kdo labeled by K′ corresponds to subpopulation in which Kdo is substituted with *PP*Etn.

Thus the chemical shift of H-3 of HepIII′ (Fig. 3A) in *O*-acetyl population showed a noticeable downfield shift (to 5.03 ppm) indicating the site of substitution.[32] Correspondingly, H-3 of HepIII′ was correlated to a signal in the $^{1}H–^{13}C$ HMQC spectrum with a chemical shift at 74.5 ppm (Fig. 3B) which is downfield shifted from the non-acetylated carbon.[6,7,33]

Major ^{31}P-resonances were observed at −0.69, 0.33, and 3.78 ppm in the ^{31}P NMR of the core OS. The high intensity ^{31}P-resonance at −0.69 was due to the mono-phosphodiester from the *P*Etn group at *O*-6 of HepII″ since a correlation was observed between the ^{31}P-resonance and both H-6 resonances (4.59 ppm) of HepII and the methylene protons (4.13 ppm) of the *P*Etn moiety while the ^{31}P-resonance at 0.33 ppm was correlated to H-4 of HepIII″ in $^{1}H–^{31}P$ HMQC spectrum (3.99 ppm) which established the location of a phospho-monoester group at C-4 of HepIII″. This assignment was also confirmed by the observed correlation between H-4 and H-3 of HepIII″ in X-filtered COSY. The ^{31}P signal at 3.78 ppm showed a correlation with a ^{1}H-resonance at 3.90 ppm in the $^{1}H–^{31}P$ HMQC spectrum

Table 4. Proton chemical shifts (ppm) and coupling constants (Hz) for core OS from LPS from the *lgtC* mutant of *H. influenzae* strain Eagan

Residue	Glycose unit	H-1 (J_{1-2})	H-2 (J_{2-3})	H-3 (J_{3-4})	H-3′ ($J_{3'}$−4)	H-4 (J_{4-5})	H-5	H-6 (J_{5-6})	H-6′ ($J_{5-6'}$,$J_{6-6'}$)	H-7 (J_{6-7})	H-7′ (J_{6-7},$J_{7-7'}$)
K	→4,5)-α-Kdo*p*-(2→	—	—			4.61	4.18		—		
HI	→3,4)-L-α-D-Hep*p*-(1→	5.05 (∼1.0)	4.00 (5.3)	3.98 (9.8)	—	4.28	3.63	4.12	—	3.71	3.78
HII	→2,3,6)-L-α-D-Hep*p*-(1→	5.69 (∼1.0)	4.27 (5.3)	4.10 (9.8)	—	4.23	3.74	4.59	—	3.72	3.89
HIII	→2)-L-α-D-Hep*p*-(1→	5.10 (∼1.0)	4.02		—				—		
GI	β-D-Glc*p*-(1→	4.52 (7.1)	3.30 (8.8)	3.52 (8.9)	—	3.42	3.50	3.91	3.75 (5.3, 12.4)		
GII	→4)-α-D-Glc*p*-(1→	5.31 (3.6)	3.60	3.88	—	3.69					
GIII	β-D-Glc*p*-(1→	4.52 (7.1)	3.35 (8.9)	3.46 (8.9)	—	3.40 (8.9)	3.46	3.96	3.75 (5.3, 12.4)		
Gal	β-D-Gal*p*-(1→	4.36 (7.1)	3.58	3.68	—	3.92 (∼1.0)	3.69	3.82 (3.5)	3.75 (5.3, 11.5)		
HII′	→2,3,6)-L-α-D-Hep*p*-(1→	5.71 (−1.0)	4.32 (5.3)	4.09 (9.8)	—	4.22	3.70	4.59	—	3.72	3.89
HIII′	→2)-L-α-D-Hep*p*-(1→ 3 ↑ Ac	5.08 (∼1.0)	4.19	5.03	—	4.00	4.17		—		
Gal′	β-D-Gal*p*-(1→	4.26 (7.1)	3.57 (8.9)	3.67 (8.9)	—	3.96 (∼1.0)	3.61	3.71 (6.4)	3.78 (6.4, 11.4)		
HII″	→2,3,6)-L-α-D-Hep*p*-(1→	5.70	4.29	4.11	—			4.60	—	3.72	3.90
HIII″	→2)-L-α-D-Hep*p*-(1→ 4 ↑ P	5.11	4.02	3.78	—	3.99			—		
Gal″	β-D-Gal*p*-(1→	4.38	3.58	3.68	—	3.92					
	*P*Etn[a]	4.13[b]	3.29[c]								
	Acetate	2.20[d]									

Residues labeled with prime are present in glycoform with acetate on C-3 of HepIII.
Residues labeled with double primes are present in glycoform with phosphate on C-4 of HepIII.
[a] *P*Etn attached to C-6 of HepII.
[b] ^{1}H chemical shift for methylene group attached to phosphate.
[c] ^{1}H chemical shift for methylene group attached to amino group.
[d] ^{1}H chemical shift for methyl group of acetate.

Table 5. ^{13}C-chemical shifts (ppm) for core OS from LPS from the *lgtC* mutant of *H. influenzae* strain Eagan

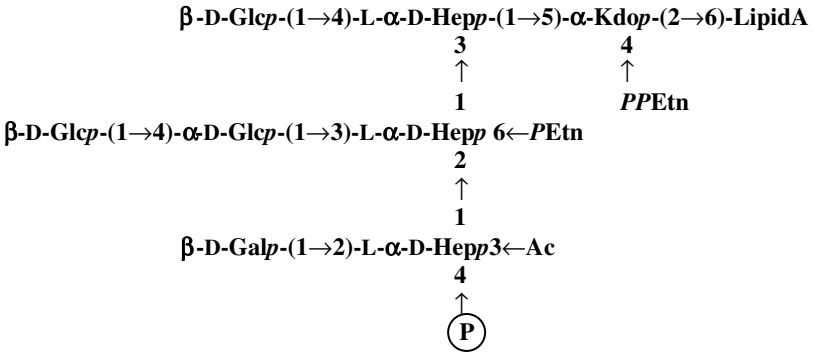
Residue	Glycose unit	C-1	C-2	C-3	C-4	C-5	C-6	C-7	C-8
K	→4,5)-α-Kdo*p*-(2→				76.9	76.1			
HI	→3,4)-L-α-D-Hep*p*-(1→	97.8	71.0	74.2	74.3	72.5	68.9		
HII	→2,3,6)-L-α-D-Hep*p*-(1→	99.4	79.3	79.3	66.5	73.0	75.2		
HIII	→2)-L-α-D-Hep*p*-(1→	100.4	79.2						
GI	β-D-Glc*p*-(1→	103.8	74.5	76.9	70.5	77.2			
GII	→4)-α-D-Glc*p*-(1→	101.0	72.6	72.5	79.7				
GIII	β-D-Glc*p*-(1→	103.8	74.5	77.0	70.9	77.2			
Gal	β-D-Gal*p*-(1→	104.1	71.3	73.5	69.6	77.0			
HII′	→2,3,6)-L-α-D-Hep*p*-(1→	99.2	79.2	79.3	66.5	73.0	75.2		
	→2)-L-α-D-Hep*p*-(1→								
HIII′	3 ↑ Ac	100.0	81.9	74.5	64.9	73.2			
Gal′	β-D-Gal*p*-(1→	103.6	71.3	73.5	69.1	75.7			
HII″	→2,3,6)-L-α-D-Hep*p*-(1→	99.5	79.3	79.2			75.3		
	→2)-L-α-D-Hep*p*-(1→								
HIII″	4 ↑ P	100.5	79.1	73.0	70.8				
Gal″	β-D-Gal*p*-(1→	103.6	71.3	73.6	69.8				
	*P*Etn[a]	63.0[b]	41.0[c]						
	Acetate	20.8[d]							

Residues labeled with prime are present in glycoform with acetate on C-3 of HepIII.
Residues labeled with double primes are present in glycoform with phosphate on C-4 of HepIII.
[a] *P*Etn attached to C-6 of HepII.
[b] ^{13}C chemical shift for methylene group attached to phosphate.
[c] ^{13}C chemical shift for methylene group attached to amino group.
[d] ^{13}C chemical shift for methyl group of acetate.

which is likely due to the heterogeneity in the structure of the core region.

Thus, the major Hex 4 glycoform is a mixture of subpopulations that contain either phosphate or *O*-acetyl groups at HepIII.

extends from HepIII of the triheptosyl inner-core unit.[8] The same trisaccharide extension is expressed as a terminal unit in *H. influenzae* strain Eagan, but in that strain it extends from C-4 of the α-D-Glc*p* off HepII.[6] We have now demonstrated that the major glycoform in strain



This structure is identical to the major Hex 4 glycoform expressed by the parent strain,[6] however the locations of the monophosphate substituent and *O*-acetyl group at HepIII had not been previously determined.

In *H. influenzae* strain Rd, two genes, *lic2A* and *lgtC*, have been shown to be responsible for phase variable expression of the galabiose component of the globoside trisaccharide [α-D-Gal*p*-(1→4)-β-D-Gal*p*- (1→4)-β-D-Glc*p*-(1→].[25] In strain Rd, this oligosaccharide epitope

Eagan bearing a mutation in *lgtC* lacks the terminal α-D-Gal*p*-(1→4)-β-D-Gal*p* unit, expressing only the β-D-Glc*p*-(1→4)-α-D-Glc*p*- (1→ as the major chain extension from C-3 of HepII in the Hep 4 glycoform. Higher glycoforms containing terminal D-Gal*p* residues were not detectable in the *lgtC* mutant strain.

While being consistent with *lgtC* having α-galactosyltransferase activity in strain Eagan as it does in strain Rd the absence of a detectable glycoform containing

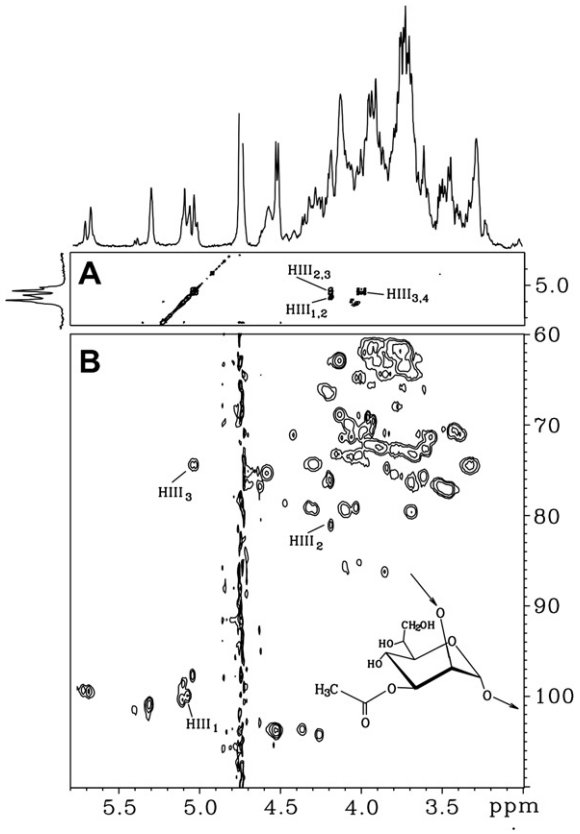


Figure 3. Partial 2D COSY (A) and HMQC (B) spectra of the core OS fraction of LPS from *H. influenzae* strain Eagan *lgtC* mutant. Cross-peaks relating anomeric and H-2 proton, and H-2 and H-3 are indicated in spectrum A. ^{13}C–^{1}H correlations of C-1, C-2 and C-3 of HepIII′ are indicated in spectrum B. Prime labeled HIII labeled with asterisks (HIII′) depicts substitution at C-3 with *O*-acetyl group.

the galobiose component of the globoside trisaccharide as a chain extension from HepII in the LPS would suggest that *lic2A* is phased off in the *lgtC* mutant strain. The globoside epitope is expressed in the LPS of many related bacteria, including *Neisseria*.[33] The α-D-Gal*p*-(1→4)-β-D-Gal*p* epitope is potentially immunodominant and its presence offers the potential for molecular mimicry of host structures.[24]

3. Experimental

3.1. Bacterial growth and LPS preparation

The *lgtC* mutant of *H. influenzae* strain Eagan (RM 153)[24] was cultivated aerobically at 36 °C in BHI broth (3.7% w/v Difco) enriched with 2 mg/L each of NADH and Hemin (Sigma) as described previously[6] except that kanamycin (10 mg/L) was added to the medium to maintain selection pressure. Bacterial cells were harvested in stationary phase by centrifugation and the obtained pellet was washed successively, once with ethanol, twice with acetone, and twice with petroleum ether (35–60 °C). Crude LPS was obtained from the air dried cellular material by the hot phenol–water extraction procedure,[28] followed by extensive dialysis, and lyophilization. LPS was purified by repeated ultracentrifugation (105,000*g*, 4 °C, 5 h).

3.2. Deoxycholate-polyacrylamide gel electrophoresis (DOC-PAGE)

PAGE was performed using the buffer system of Laemmli and Favre[34] as modified by Komuro and Galanos[35] with DOC as detergent. LPS bands were visualized by silver staining procedure as described by Tsai and Frasch.[36]

3.3. O-Deacylated LPS and core oligosaccharide preparations

LPS was O-deacylated with anhydrous hydrazine under mild condition[37] as previously described.[6,38]

Core oligosaccharide (OS) was obtained by mild acid hydrolysis as described previously.[6,38] The OS product was fractionated on the Bio-Gel P2 gel filtration column (2.6 × 140 cm, 200–400 mesh, BioRad) elution with pyridinium acetate (0.05 M, pH 4.5). Column eluants were continuously monitored for changes in refractive index by a Water R403 differential refractometer, and collected fractions (4.5 mL) were assayed colorimetrically for neutral glycoses by phenol-sulfuric acid method.[39] The fractions containing OS were collected and lyophilized.

3.4. Analytical methods and methylation analysis

Neutral glycoses were determined by GLC as their alditol acetate derivatives as described.[6] Briefly, samples (0.2–0.5 mg) were hydrolyzed with 2 M TFA (90 min at 125 °C) and evaporated to dryness under a stream of nitrogen, and the liberated glycoses were reduced ($NaBH_4$) and acetylated (Ac_2O). The hexoses were established to have D-configurations by GLC analysis of their acetylated (*R*)-2-octyl glycoside derivatives.[40] The Kdo absolute configuration was previously reported to be D-*manno* configuration.[41] Methylation analysis was performed on core OS sample (2 mg) with iodomethane in dimethyl sulfoxide containing an excess of potassium (methylsulfinyl) methanide.[42] Excess iodomethane was evaporated under a stream of nitrogen, water (3 mL) was added, and the methylated OS was purified on a Sep-Pak C-18 cartridge. Purified methylated oligosaccharide was hydrolyzed with 0.25 M H_2SO_4 in 95% acetic acid at 85 °C overnight, reduced ($NaBD_4$), and acetylated according to the acetolysis procedure of Stellner et al.[43] Partially methylated glycose derivatives were separated by GLC as described.[6]

3.5. NMR spectroscopy

NMR spectra were obtained on Bruker AMX 500 or 600 MHz spectrometers using standard Bruker software. All measurements were made at 27 °C in 0.5 mL of D_2O solution for core OS sample or 1.5 mM EDTA in 0.5 mL of D_2O solution for LPS-OH sample (pD 6-7), subsequent to several lyophilizations with D_2O.

Proton spectra were recorded by using a spectral width of 10.6 kHz and a 90° pulse. ^{31}P NMR spectra were measured at 202 MHz using spectral width of 12.1 and 41.6 kHz, and phosphoric acid (85%) was used as the external standard (δ_p, 0.0 ppm). Acetone was used as the internal standard for ^1H and ^{13}C spectra, and chemical shifts were referenced to the methyl resonance (δ_H, 2.225 ppm; δ_C, 31.07 ppm). Homonuclear (COSY, TOCSY, NOESY) and heteronuclear (^{13}C–^1H and ^{31}P–^1H HMQC) correlation experiments were performed as previously described.[6,44] 2D ^{31}P–^1H NMR X-filtered COSY was performed for core OS sample over spectral width of 1.8 kHz by using a data set of 256×1024 points; X-filtered TOCSY was performed for O-deacylated LPS samples over a spectral width of 4.8 kHz by using a data set of 128×2048 points.

3.6. Electrospray mass-spectrometry

Samples were analyzed on a VG Quattro Mass Spectrometer (Fisons Instruments) fitted with an electrospray ion source. Core OS and LPS-OH samples were dissolved in water which was then mixed in a 1:1 ratio with 50% aqueous acetonitrile containing 0.4% acetic acid for mass spectral analysis in either the negative or positive ion mode as describes previously.[6,7]

Acknowledgments

We thank D. Griffith for large-scale growth of cells, F. Cooper for GLC–MS analysis, and D. Krajcarski for ESI-MS analysis. We also thank Dr. J.-R. Brisson for assistance in the acquisition of NMR experiments. ERM was supported by a programme grant from the Medical Research Council, UK.

References

1. Moxon, E. R.; Vaughn, K. A. *J. Infect. Dis.* **1981**, *143*, 517–524.
2. Turk, D. D. In *Haemophilus influenzae, Epidemiology, Immunology and Prevention of Disease*; Sell, S. H., Wright, P. F., Eds.; Elsevier: New York, 1981; pp 3–9.
3. Murphy, T. F.; Apicella, M. A. *Rev. Infect. Dis.* **1987**, *9*, 1–15.
4. Schweda, E. K. H.; Richards, J. C.; Hood, D. W.; Moxon, E. R. *Int. J. Med. Microbiol.* **2007**, *297*, 297–306.
5. Schweda, E. K.; Li, J.; Moxon, E. R.; Richards, J. C. *Carbohydr. Res.* **2002**, *337*, 409–420.
6. Masoud, H.; Moxon, E. R.; Martin, A.; Krajcarski, D.; Richards, J. C. *Biochemistry* **1997**, *36*, 2091–2103.
7. Masoud, H.; Martin, A.; Thibault, P.; Moxon, E. R.; Richards, J. C. *Biochemistry* **2003**, *42*, 4463–4475.
8. Risberg, A.; Masoud, H.; Martin, A.; Richards, J. C.; Moxon, E. R.; Schweda, E. K. H. *Eur. J. Biochem.* **1999**, *261*, 171–180.
9. Landerholm, M. K.; Li, J.; Richards, J. C.; Hood, D. W.; Moxon, E. R.; Schweda, E. K. *Eur. J. Biochem.* **2004**, *271*, 941–953.
10. Risberg, A.; Alvelius, G.; Schweda, E. K. H. *Eur. J. Biochem.* **1999**, *265*, 1067–1074.
11. Månsson, M.; Bauer, S. H. J.; Hood, D. W.; Richards, J. C.; Moxon, E. R.; Schweda, E. K. H. *Eur. J. Biochem.* **2001**, *268*, 2148–2159.
12. Phillips, N. J.; Apicella, M. A.; Griffiss, J. M.; Gibson, B. W. *Biochemistry* **1992**, *31*, 4515–4526.
13. Gibson, B. W.; Melaugh, W.; Phillips, N. J.; Apicella, M. A.; Campagnari, A. A.; Griffiss, J. M. *J. Bacteriol.* **1993**, *175*, 2702–2712.
14. Phillips, N. J.; Apicella, M. A.; Griffiss, J. M.; Gibson, B. W. *Biochemistry* **1993**, *32*, 2003–2012.
15. Schweda, E. K. H.; Hegedus, O. E.; Borrelli, S.; Lindberg, A. A.; Weiser, J. W.; Maskell, D. J.; Moxon, E. R. *Carbohydr. Res.* **1993**, *246*, 319–330.
16. Schweda, E. K. H.; Jansson, P.-E.; Moxon, E. R.; Lindberg, A. A. *Carbohydr. Res.* **1995**, *272*, 213–224.
17. Phillips, N. J.; McLaughlin, R.; Miller, T. J.; Apicella, M. A.; Gibson, B. W. *Biochemistry* **1996**, *35*, 5937–5947.
18. Risberg, A.; Schweda, E. K. H.; Jansson, P.-E. *Eur. J. Biochem.* **1997**, *243*, 701–707.
19. Rahman, M. M.; Gu, X.-X.; Tsai, C.-M.; Kolli, V. S. K.; Carlson, R. W. *Glycobiology* **1999**, *9*, 1371–1380.
20. Månsson, M.; Hood, D. W.; Li, J.; Richards, J. C.; Moxon, E. R.; Schweda, E. K. H. *Eur. J. Biochem.* **2002**, *269*, 808–818.
21. Månsson, M.; Hood, D. W.; Moxon, E. R.; Schweda, E. K. H. *Eur. J. Biochem.* **2003**, *270*, 610–624.
22. Månsson, M.; Hood, D. W.; Moxon, E. R.; Schweda, E. K. H. *Eur. J. Biochem.* **2003**, *270*, 2979–2991.
23. Yildirim, H. H.; Hood, D. W.; Moxon, E. R.; Schweda, E. K. H. *J. Biochem.* **2003**, *270*, 3153–3167.
24. Hood, D. W.; Deadman, M. E.; Allen, T.; Masoud, H.; Martin, A.; Brisson, J.-R.; Fleischmann, R.; Venter, J. C.; Richards, J. C.; Moxon, E. R. *Mol. Microbiol.* **1996**, *22*, 951–965.
25. Hood, D. W.; Cox, A. D.; Gilbert, M.; Makepeace, K.; Walsh, S.; Deadman, M. E.; Cody, A.; Martin, A.; Mansson, M.; Schweda, E. K.; Brisson, J.-R.; Richards, J. C.; Moxon, E. R.; Wakarchuk, W. W. *Mol. Microbiol.* **2001**, *39*, 341–350.
26. Hood, D. W.; Deadman, M. E.; Cox, A. D.; Makepeace, K.; Martin, A.; Richards, J. C.; Moxon, E. R. *Microbiology* **2004**, *150*, 2089–2097.
27. Hood, D. W.; Randle, G. A.; Cox, A. D.; Makepeace, K.; Li, J.; Schweda, E. K.; Richards, J. C.; Moxon, E. R. *J. Bacteriol.* **2004**, *186*, 7429–7439.
28. Westphal, O.; Lüderitz, O.; Bister, F. Z. *Naturforschg* **1952**, *7b*, 148–155.
29. Batley, M.; Packer, N. H.; Redmond, J. W. *Biochim. Biophys. Acta* **1985**, *821*, 179–194.
30. Masoud, H.; Urbanik-Sypniewska, T.; Lindner, B.; Weckesser, J.; Mayer, H. *Arch. Microbiol.* **1991**, *156*, 167–175.
31. Masoud, H.; Perry, M. B.; Richards, J. C. *Eur. J. Biochem.* **1994**, *220*, 209–216.

32. Schweda, E. K.; Brisson, J. R.; Alvelius, G.; Martin, A.; Weiser, J. N.; Hood, D. W.; Moxon, E. R.; Richards, J. C. *Eur. J. Biochem.* **2000**, *267*, 3902–3913.
33. Virji, M.; Weiser, J. N.; Lindberg, A. A.; Moxon, E. R. *Microb. Pathog.* **1990**, *9*, 441–450.
34. Laemmli, U. K.; Favre, M. *J. Mol. Biol.* **1973**, *80*, 575–599.
35. Komuro, T.; Galanos, C. *J. Chromatogr.* **1988**, *450*, 381–387.
36. Tsai, C.-M.; Frasch, C. E. *Biochemistry* **1982**, *119*, 115–119.
37. Holst, O.; Brade, L.; Kosma, P.; Brade, H. *J. Bacteriol.* **1991**, *173*, 1862–1866.
38. Masoud, H.; Altman, E.; Richards, J. C.; Lam, J. S. *Biochemistry* **1994**, *33*, 10568–10578.
39. Dubois, M.; Gilles, K. A.; Hamilton, J. K.; Rebers, P. A.; Smith, F. *Anal. Chem.* **1956**, *28*, 350–356.
40. Leontein, K.; Lindberg, B.; Lönngren, J. *Carbohydr. Res.* **1978**, *62*, 359–362.
41. Helander, I. M.; Lindner, B.; Brade, H.; Altmann, K.; Lindberg, A. A.; Rietschel, E. Th.; Zähringer, U. *Eur. J. Biochem.* **1988**, *177*, 483–492.
42. Hakomori, H. I. *J. Biochem.* **1964**, *55*, 205–208.
43. Stellner, K.; Saito, H.; Hakomori, S.-I. *Arch. Biochem. Biophys.* **1973**, *155*, 464–472.
44. Masoud, H.; Perry, M. B.; Brisson, J.-R.; Uhrin, D.; Richards, J. C. *Can. J. Chem.* **1994**, *72*, 1466–1477.