Carbohydrate Research 343 (2008) 2763–2770

Contents lists available at ScienceDirect

Carbohydrate Research

journal homepage: http://www.elsevier.com/locate/carres

Identification of a novel structural motif in the lipopolysaccharide of the *galE*/*galK* double mutant of *Haemophilus influenzae* strain Eagan

Hussein Masoud [a,b], Dusan Uhrin [a,†], E. Richard Moxon [c], 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 Paediatrics, Institute for Molecular Medicine, John Radcliffe Hospital, Headington, Oxford 0X3 3DU, UK*

ARTICLE INFO

Article history:
Received 28 January 2008
Received in revised form 12 April 2008
Accepted 21 April 2008
Available online 29 April 2008

Keywords:
Haemophilus influenzae
Lipopolysaccharide
LPS-backbone
Galactose biosynthetic pathway
galE/*galK* mutant

ABSTRACT

Defined mutants of the galactose biosynthetic (Leloir) pathway were employed to investigate lipopolysaccharide (LPS) oligosaccharide expression in *Haemophilus influenzae* type b strain Eagan. The structures of the low-molecular-mass LPS glycoforms from strains with mutations in the genes that encode galactose epimerase (*galE*) and galactose kinase (*galK*) were determined by NMR spectroscopy on O- and N-deacylated and dephosphorylated LPS-backbone, and O-deacylated oligosaccharide samples in conjunction with electrospray mass spectrometric, glycose and methylation analyses. The structural profile of LPS glycoforms from the *galK* mutant was found to be identical to that of the galactose and glucose-containing Hex5 glycoform previously identified in the parent strain [Masoud, H.; Moxon, E. R.; Martin, A.; Krajcarski, D.; Richards, J. C. *Biochemistry* **1997**, *36*, 2091–2103]. LPS from the *H. influenzae* strain bearing mutations in both *galK* and *galE* (*galE*/*galK* double mutant) was devoid of galactose. In the double mutant, Hex3 and Hex4 glycoforms containing di- and tri-glucan side chains from the central heptose of the triheptosyl inner-core unit were identified as the major glycoforms. The triglucoside chain extension, β-D-Glc*p*-(1→4)-β-D-Glc*p*-(1→4)-α-D-Glc*p*, identified in the Hex4 glycoform has not been previously reported as a structural element of *H. influenzae* LPS. In the parent strain, it is the galactose-containing trisaccharide, β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→4)-α-D-Glc*p*, and further extended analogues thereof, that substitute the central heptose. When grown in galactose deficient media, the *galE* single mutant was found to expresses the same population of LPS glycoforms as the double mutant.

1. Introduction

Haemophilus influenzae remains a significant cause of upper and lower respiratory tract infections. Encapsulated organisms, particularly serogroup b, cause life-threatening invasive diseases such as meningitis and pneumonia, while acapsulated organisms (nontypable) are associated with localized upper and lower respiratory tract infections and otitis media.[1,2] Lipopolysaccharide (LPS) is an essential component of the cell wall of this pathogen and is implicated as a major virulence factor.[3]

The rough type or low-molecular-weight LPS of *H. influenzae* (sometimes referred to as lipooligosaccharide) is composed of a membrane-anchoring lipid A and short oligosaccharide core region. The core region consists of a conserved inner-core unit composed of a phosphorylated 2-keto-3-deoxy-D-*manno*-octulosonic acid (Kdo)-triheptosyl unit. Each heptose residue can serve as a point for glycosyl extensions that comprise the outer core. The structures of *H. influenzae* LPS have been extensively investigated in a number of strains (for a review, see Schweda et al.[4]). The outer core displays extensive structural and antigenic diversity, within the same strain and between closely related strains.[5–13] This is due to genetic variation in hexose substitution and/or the presence of sialic acid and LD- and DD-heptoses, as well as noncarbohydrate groups including phosphate, acetate, phosphocholine (*P*Cho), glycine, and aminosugars[8–11,13–19] which are controlled by a large number of genes possessed by the microorganism.[6,19] Structural diversity is important for the pathogen to avoid host defense

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*-octulsonic acid; LPS, lipopolysaccharide; LPS-OH, O-deacylated LPS; Lipid A-OH, O-deacylated lipid A; OS, oligosaccharide; Etn, ethanolamine; *P*Etn, phosphoethanolamine; *PP*Etn, pyrophosphoethanolamine; *P*Cho, phosphocholine; UDP, uridine diphosphate; EDTA, ethylenediaminetetraacetic acid; SDS, sodium dodecyl sulfate

* Corresponding author. Tel.: +1 613 993 7506; fax: +1 613 957 7867.
E-mail address: jim.richards@nrc.ca (J. C. Richards).
† Present address: Department of Chemistry, The University of Edinburgh, Edinburgh, UK.

0008-6215/$ - see front matter Crown Copyright © 2008 Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.carres.2008.04.027

mechanism and a number of phase-variable genes have been identified that are contributing factors.

As part of our ongoing interest in the expression, structural diversity, and immunobiology of this molecule we have undertaken a systematic study of mutant strains in which mutants of genes important in LPS biosynthesis have been constructed.[4] In the present study, we report the structures of LPS from single and double *galE* and *galK* mutants generated in *H. influenzae* strain Eagan.[20] These genes encode enzymes of the Leloir pathway for catabolism of exogenous galactose; where GalK is the kinase necessary for 1-phosphorylation of exogenous galactose, and GalE is the *O*-4 epimerase that catalyzes interconversion of UDP-glucose to UDP-galactose. Under the growth conditions employed, *H. influenzae* strain Eagan bearing a single mutation in *galK* elaborates galactose-containing LPS glycoforms similar to those of the parent strain,[7] while LPS from organisms that carry a mutation in *galE* or in both *galE* and *galK* lack galactose. Detailed structural analysis of LPS from the *galE*/*galK* and *galE* mutants uncovered a major glycoform bearing a unique triglucoside extension from the conserved triheptosyl inner-core unit.

2. Results

2.1. Characterization of LPS from *H. influenzae* strain Eagan *gal* mutants

Deletion–insertion mutations constructed in *galE* (containing a kanamycin resistance cassette) and *galK* (containing a tetracycline resistance cassette) were moved into the *H. influenzae* type b strain Eagan (RM153) chromosome to generate each of the single mutants as well as the double mutant (*galE/galK*).[20] Mutant strains were grown in the presence of the appropriate antibiotic(s) and LPS was isolated by using the hot phenol–water extraction method[21] in yields of ca. 2–3% from dried bacterial cells. Deoxycholate-PAGE analysis of LPS from the *galE* and *galE*/*galK* mutants (Fig. 1, lanes 3 and 4) revealed two distinct bands in the low-molecular-mass region having electrophoretic mobility similar to those of LPS from the parent strain (Fig. 1, lane 2), while LPS from the *galK* mutant (Fig. 1, lane 5) showed a single band corresponding in electrophoretic mobility to the upper band from wild-type strain Eagan.

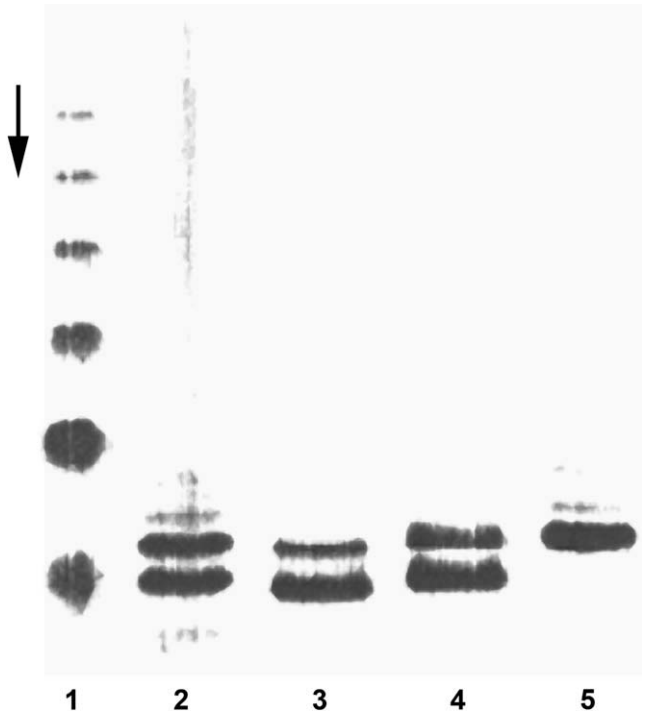


Figure 1. DOC-PAGE pattern of LPS from *Salmonella milwaukee* and *Haemophilus influenzae* RM 153 (Eagan). Lane 1, *S. milwaukee* (S-type LPS, 10 μg); lane 2, *H. influenzae* strain Eagan (4 μg); lane 3, *H. influenzae* strain Eagan *galE/galK* mutant (2 μg); lane 4, *H. influenzae* strain Eagan *galE* mutant (2 μg); and lane 5, *H. influenzae* strain Eagan *galK* mutant (2 μg).

Backbone OS samples were obtained from each of the LPS samples by a procedure involving deacylation, dephosphorylation, and reduction of the lipid A terminal GlcN residue.[22] Following fractionation by size exclusion chromatography on a Biogel P-2 gel filtration system, a single fraction was obtained from *galK* LPS, whereas *galE*/*galK* and *galE* LPS afforded high- and low-molecular-weight (HMW and LMW, respectively) backbone OS fractions. Glycose analysis of both the LMW and HMW backbone OS fractions revealed the presence of D-glucose (D-Glc), 2-amino-2-deoxy-D-glucose (D-GlcN), and L-*glycero*-D-*manno*-heptose (LD-Hep). As expected, D-galactose (D-Gal) was not detectable in the backbone OS fractions from the *galE*/*galK* double mutant. Similarly, D-Gal was not detected in the OS fractions from the *galE* single mutant but was present in the *galK* single mutant.

2.2. Structures of LMW and HMW backbone oligosaccharides from *galE*/*galK*

The structures of the backbone OS fractions from the *galE/galK* double mutant were determined in detail. Methylation analysis confirmed the absence of D-Gal residues in the two oligosaccharide fractions, revealing terminal D-Glc and O-4 substituted D-Glc as the only hexose units. The data indicated an additional O-4 substituted D-Glc in the HMW fraction. The presence of O-3 and O-4 disubstituted LD-Hep, O-2 and O-3 disubstituted LD-Hep, and terminal LD-Hep in both the LMW and HMW fractions is consistent with the LD-HepIII-(1→2)-LD-HepII-(1→3)-LD-HepI(1→ triheptosyl inner-core unit in which there is further substitution at O-4 of HepI and O-3 of HepII.[7] This was confirmed by detailed NMR analysis of the two fractions.

The ^{1}H NMR resonances of the LMW and HMW fractions from the *galE*/*galK* double mutant were fully assigned using homo- and heteronuclear correlation experiments (COSY and HMQC; 37 °C, pD ~ 6) as described earlier[7,12] and the data are presented in Tables 1 and 2. Six anomeric ^{1}H resonances were observed in the low-field region (5.7–4.5) of the LMW-OS fraction; five of approximately equal signal area, and one (~4.57 ppm) of double signal area indicating two overlapping signals which were resolved in the COSY spectrum (Fig. 2A). Correspondingly, seven anomeric ^{13}C resonances were identified (90–110 ppm) in the ^{13}C–^{1}H correlation (HMQC) spectrum (Fig. 3A). The three subspectra (labeled HI–HIII) corresponding to the triheptosyl inner-core unit was readily identified from their chemical shift and coupling constant values (Tables 1 and 2) and by comparison with those of the parent strain.[7] Four ^{1}H subspectra attributed to hexopyranosyl residues having the *gluco* configuration (labeled GI–GIII and GNII) were characterized on the basis of the observed large vicinal proton coupling constant ($J_{2,3}$, $J_{3,4}$, and $J_{4,5}$ = 8–10 Hz) (Table 1). From the magnitude of the $J_{1,2}$ couplings, one of these residues (GI) was assigned to have the α-D-configuration ($J_{1,2}$ = 3.8 Hz), while the other three residues (GII, GIII, and GNII) were identified as β-D-anomers ($J_{1,2}$ = 8.7 Hz). The anomeric configurations of the glycopyranosyl residues were confirmed from the observed $^{1}J_{C\text{-}1,H\text{-}1}$ values which showed 171 Hz for the α-configuration (GI) and ~160 Hz for the β-configuration (GII, GIII, and GNII) (Table 2).[23] The H-2 resonance (3.05 ppm) from the residue GNII was directly correlated to a ^{13}C-resonance at 56.9 ppm in the HMQC (Fig. 3B), which is characteristic of an amino-substituted carbon (50–60 ppm),[24] indicating it to

Table 1
Proton chemical shifts (ppm) and coupling constants (Hz) for LMW and HMW backbone OS fractions from *H. influenzae* strain Eagan *galE/galK*

Oligo-saccharide	Residue unit	Glycose	H-1 ($J_{1,2}$)	H-2 ($J_{2,3}$)	H-3a[b] ($J_{3a,4}$)	H-3e[b] ($J_{3e,4}$, $J_{3e,3a}$)	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′ ($J_{6,7'}$, $J_{7,7'}$)	H-8 ($J_{7,8}$)	H-8′ ($J_{7,8'}$, $J_{8,8'}$)
LMW-OS fraction	GNI^r	→6)-D-GlcN*ol*	3.9 (4.4)	3.55 (3.8)	4.12 (1.0)	—	3.67 (8.8)	3.99	4.18 (3.2)	3.84 (7.3,10.5)	—	—	—	—
	GNII	→6)-β-D-Glc*p*N- (1→	4.73 (8.7)	3.05 (8.7)	3.63 (9.0)	—	3.52 (8.8)	3.63	—	—	—	—	—	—
	K	→5)-α-Kdo*p*-(2→		—	1.83 (12.0)	2.18 (4.5,12.0)	4.18 (4.4)	4.13	3.70 (1.0)	—	3.80 (8.8)	—	3.98 (4.4)	3.67 (5.9,11.7)
	HI	→3,4)-L-α-D-Hep*p*-(1→	5.10 (1.0)	4.16 (3.8)	4.08 (9.4)	—	4.32 (10.3)	4.18	4.07 (1.0)	—	3.75	—	—	—
	HII	2,3)-L-α-D-Hep*p*-(1→	5.69 (1.0)	4.25 (3.2)	4.06 (9.6)	—	4.13 (9.6)	3.68	4.10 (1.0)	—	3.63 (5.8)	3.74 (8.7, 10.2)	—	—
	HIII	L-α-D-Hep*p*-(1→	5.04 (1.0)	3.88 (3.5)	3.89	—	3.82	3.82	4.03 (1.0)	—	3.76 (4.4)	3.70 (8.7, 11.6)	—	—
	GI	→4)-α-D-Glc*p*-(1→	5.29 (3.8)	3.63 (9.6)	3.89 (8.8)	—	3.74 (8.8)	3.91	3.98 (1.0)	3.88 (—, 11.6)	—	—	—	—
	GII	β-D-Glc*p*-(1→	4.55 (8.7)	3.42 (8.7)	3.48 (9.0)	—	3.31 (8.8)	3.48	3.98 (1.0)	3.77 (7.3, 11.7)	—	—	—	—
	GIII	β-D-Glc*p*-(1→	4.59 (8.8)	3.39 (8.8)	3.67 (9.0)	—	3.68		3.94 (1.0)	3.76 (5.9, 11.7)	—	—	—	—
HMW-OS fraction[a]	GIII′	→4)-β-D-Glc*p*-(1→	4.59 (8.8)	3.39 (8.8)	3.67	—	3.63	—	—	—	—	—	—	—
	GIV	β-D-Glc*p*-(1→	4.52 (8.8)	3.32 (8.8)	3.52 (9.0)	—	3.43 (8.8)	3.50	3.92 (2.9)	3.74 (5.8, 11.7)	—	—	—	—

[a] Residue labeled with prime is the same residue present in Hex 3 glycoform but differs in chemical shifts because of the substitution.

Table 2
^{13}C-chemical shifts (ppm) for LMW and HMW backbone OS fractions from *H. influenzae* strain Eagan *galE/galK*

Oligosaccharide	Residue unit	Glycose	C-1 ($J_{C,H}$)	C-2	C-3	C-4	C-5	C-6	C-7	C-8
LMW-OS fraction	GNI^r	→6)-D-GlcN*ol*	60.0	56.5	66.9	72.5	70.5	72.6		
	GNII	→6)-β-D-Glc*p*N-(1→	101.0 (163)	56.9	72.5	71.2	75.5	62.8		
	K	→5)-α-Kdo*p*-(2→			35.9	66.8	75.5	72.7	70.4	64.6
	HI	→3,4)-L-α-D-Hep*p*-(1→	101.9 (171)	71.3	74.5	73.9	72.8	69.5	64.4	
	HII	→2,3)-L-α-D-Hep*p*-(1→	99.9 (175)	80.3	80.3	66.9	72.7	69.5	64.4	
	HIII	L-α-D-Hep*p*-(1→	103.0 (169)	71.9	71.9	67.4	72.9	70.5	64.4	
	GI	→4)-α-D-Glc*p*-(1→	101.6 (171)	72.5	72.6	79.4	72.6	61.1		
	GII	β-D-Glc*p*-(1→	103.6 (161)	74.7	77.4	71.6	77.2	62.5		
	GIII	β-D-Glc*p*-(1→	103.6 (161)	74.2	76.7	70.7	77.0	61.6		
HMW-OS fraction[a]	GIII′	→4)-β-D-Glc*p*-(1→	103.5	74.2	75.4	79.0				
	GIV	β-D-Glc*p*-(1→	103.7	74.2	76.7	70.7	77.2	61.7		

[a] Residue labeled with prime is the same residue present in Hex 3 glycoform but differs in chemical shifts because of the substitution.

be a glucosamine residue. Another ^{13}C resonance in that region (56.5 ppm) was observed in the HMQC spectrum (Fig. 3B), which correlated to H-2 (3.55 ppm) from the ^{1}H spin-system due to the glucosaminitol end-group (GNI^r).[22,25] The Kdo residue (K) was readily identified from its H-3 and C-3 chemical shift values.

Thus, the data were consistent with a nonasaccharide structure comprising one α-D-Glc*p* (GI), two β-D-Glc*p* (GII and GIII), three L-α-D-Hep*p* (HI, HII, and HIII), one α-D-Kdo*p* (K), one β-D-Glc*p*N (GNII), and one D-GlcN*ol* (GNI^r) for the LMW-OS. A similar analysis of the HMW-OS fraction indicated the presence of an additional β-D-Glc*p* (GVI) (Tables 1 and 2) indicating this oligosaccharide to have ten sugar residues. Apart from GIII, the ^{1}H and ^{13}C chemical shift values of the other glycosyl residues in the HMW-OS closely corresponded to those of the LMW-OS (±0.04 ppm for ^{1}H resonances; ±0.40 ppm for ^{13}C resonances).

The sequences of the glycosyl residues within the LMW- and HMW-OS fractions were established from the observed NOEs between anomeric and aglyconic protons of contiguous residues.[7,12] This is shown in Figure 2B for the LMW-OS. NOEs were observed between H-1 of HIII and H-2 of HII, H-1 of HII and H-3 of HI, and H-1 of HI and H-5 of K, which confirmed the linear sequence of triheptosyl inner-core unit and its linkage to Kdo 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*. An inter-residue NOE between the anomeric proton of GII and H-4 and H-6 of HepI established the linkage of this β-D-Glc*p* residue to the O-4 of HepI as previously observed in the parent strain.[7] Trans-glycosidic NOEs were observed between H-1 of GIII and H-4 of GI, and H-1 of GI and H-3 of HII establishing the side chain at O-3 of HepII as β-D-Glc*p*-(1→4)-α-D-Glc*p*(1→3)-L-α-D-Hep*p*II. In the HMW-OS, the extra β-D-Glc*p* (GIV) showed a transglycosidic NOE between its H-1 and the overlapping H-3/H-4 resonances of the GIII′ residue. A comparison of the ^{13}C NMR data of the terminal β-D-Glc*p* residue in the LMW-OS (GIII) and its substituted analogue (GIII′) in the HMW-OS (Table 2) pointed to substitution at the O-4

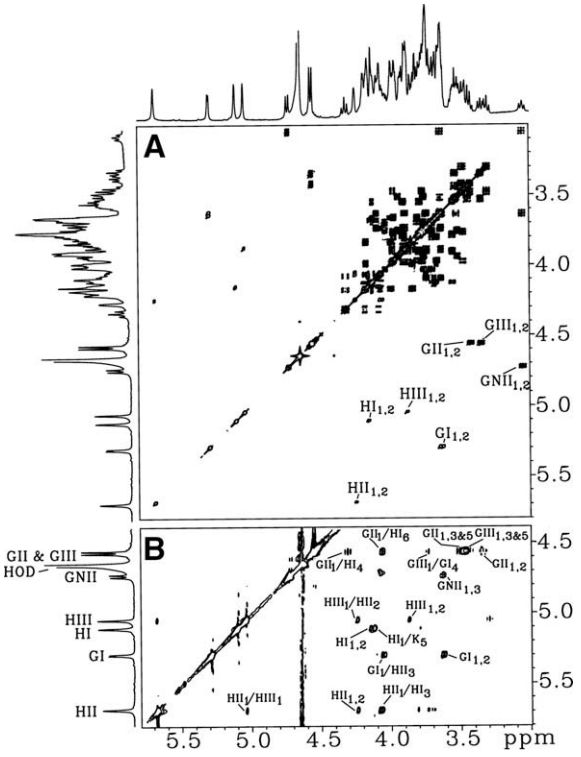
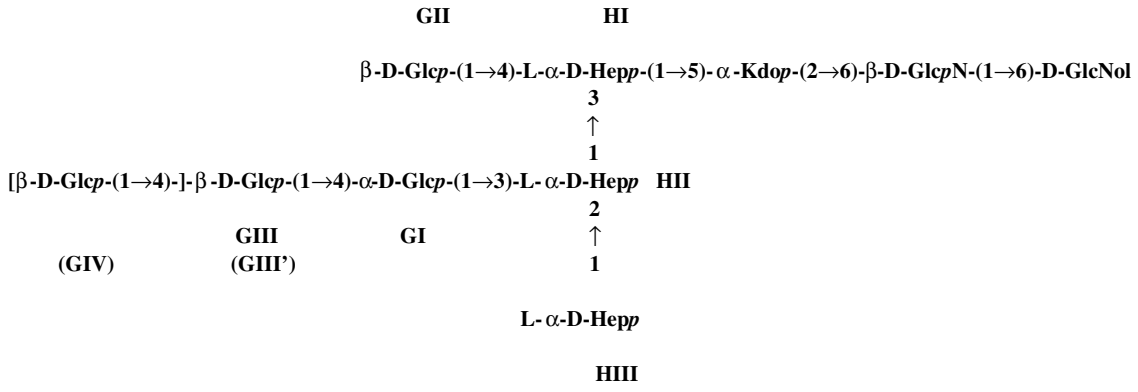


Figure 2. Partial 2D COSY (A) and NOESY (B) spectra of the low-molecular-weight (LMW) backbone OS from the *H. influenzae* strain Eagan *galE*/*galK* mutant. Cross peaks relating anomeric and ring protons, and NOE connectivities are indicated.

of the GIII′ residue in the HMW-OS. This was readily apparent from the observed downfield-shifted value (8.3 ppm) of *C*-4 with the concomitant upfield-shifted value (1.3 ppm) for the adjacent *C*-3 in the HMW-OS.[24] Moreover, methylation analysis indicated only O-4 substituted Glc residues in the OS.

The linkage positions of the O-5 linked Kdo containing fragment to deacylated and reduced lipid A backbone region (D-GlcN*p*-(1→6)-GlcNol) were deduced from a comparison of the NMR data with the parent strain.[7] Thus, based on the combined NMR and methylation analysis, the structures of the *galE*/*galK* backbone HMW- and LMW-OS fractions can be assigned as follows:



2.3. ESI-MS analysis of LPS from *galE*/*galK*

ESI-MS of LPS-OH that was prepared by treatment of LPS with anhydrous hydrazine (37 °C, 1 h) revealed a series of related structures that varied in the number of hexoses and the presence or absence of phosphoethanolamine (*P*Etn). Doubly and triply charged molecular ions corresponding to the *P*Etn-containing triheptosyl inner-core unit attached to the O-deacylated lipid A moiety via a phosphorylated Kdo residue with 3–5 attached hexose residues (Table 3). The Hex3 (53.1%) and Hex4 (28.5%) glycoforms were most prominent and their backbone structures are represented by those of the derived LMW and HMW-OSs, respectively. The presence of a minor Hex5 glycoform (3.8%) was also detected from the respective triply and doubly charged ions at *m*/*z* 920.2 and *m*/*z* 1381.5. Interestingly, a parallel series of triply and doubly charged ions shifted 123 amu to lower masses could be attributed to the LPS-OH glycoforms lacking the *P*Etn group (Table 3).

2.4. Characterization of the phosphate substitution patterns of LPS from *galE*/*galK*

The phosphate substitution patterns of LPS from the double *galE*/*galK* mutant were established on core OS and LPS-OH samples. Core OS was methylated using iodomethane,[27] dephosphorylated (48% aqueous hydrogen fluoride, 0 °C, 48 h) and remethylated using iodomethane-d_3. Total acid hydrolysis, reduction ($NaBD_4$), and acetylation followed by GLC–MS analysis afforded the following derivatives: 2,3,4,6-tetra-*O*-methyl-glucose, 2,3,6-tri-*O*-methyl-glucose, 2,3,4,6,7-penta-*O*-methyl-heptose, 2,6,7-tri-*O*-methyl-heptose, and 4,6,7-tri-*O*-methyl-heptose. The mass spectrum of 4,6,7-tri-*O*-methyl-heptose (not shown data) afforded fragments of approximately equal intensity at *m*/*z* 48 and 45 indicating 1:1 distribution of deuterium labeled and nonlabeled methyl groups at the O-7 position, respectively. Additionally, the absence of a fragment at *m*/*z* 89 taken together with the observed signal at *m*/*z* 92 indicated the presence of a labeled methyl group at either the O-6 or O-7 position of the heptose derivative. It could therefore be concluded from these data that phosphate substitution is either at the *O*-6 or *O*-7 position of middle heptose (HII) of the triheptosyl unit.

As previously noted for the parent strain,[7] phosphate migration between C-6 and C-7 under the alkaline conditions employed in the methylation analysis procedure precluded definitive assignment of the location of the *P*Etn moiety on HepII. The location of the *P*Etn unit was unambiguously established at *O*-6 of HepII from ^{1}H, ^{13}C, and ^{31}P NMR analysis of the LPS-OH sample. The NMR spectra of LPS-OH were made with 1.5 mM EDTA-d_{12} and 10% SDS-d_{25} in D_2O. The ^{13}C and ^{1}H NMR resonances were partially assigned by COSY, TOCSY, HMQC, and selective 1D experiments and the data from the major Hex3 glycoform are presented in Table 4.

The anomeric signal of HepII could be identified at 5.78 ppm. Due to the small $J_{1,2}$ and fast spin–spin relaxation times caused by the presence of fatty acid chains in the lipid A-OH, selective TOCSY transfer from H-1 of HepII stopped at H-2 (4.29 ppm). Advantage was therefore taken of the presence of the neighboring α-D-Glcp residue (GI) which is attached at the C-3 position of HepII.

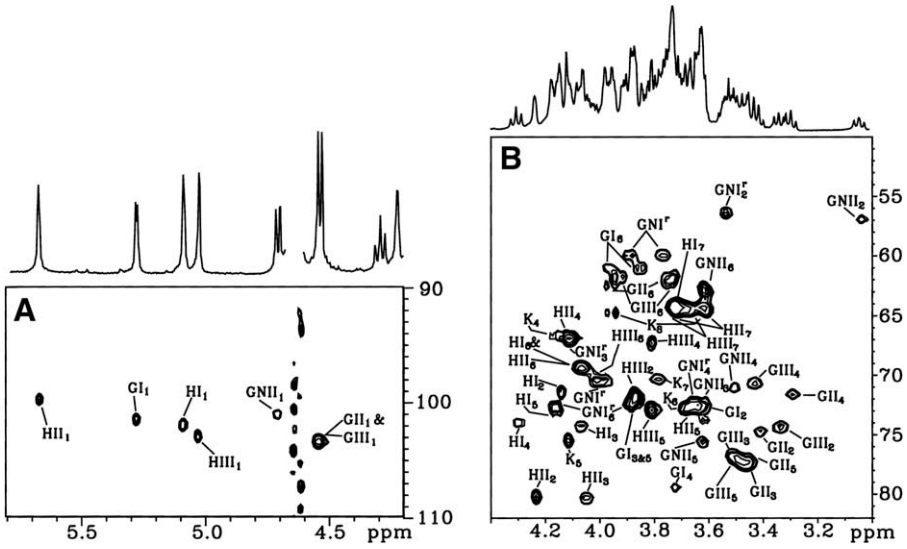


Figure 3. Heteronuclear 2D ^{13}C-^{1}H chemical shifts correlation map of the anomeric (A) and ring (B) regions of the low-molecular-weight backbone (LMW) OS from the *H. influenzae* strain Eagan *galE*/*galK* mutant.

Table 3
Negative ion ESI-MS data and proposed compositions of O-deacylated LPS from *H. influenzae* stain Eagan *galE/galK*

LPS glycoform	Observed ions (*m*/*z*)		Molecular mass (Da)		Relative intensity (%)	Proposed composition
	$(M-2H)^{2-}$	$(M-3H)^{3-}$	Observed[a]	Calculated[b]		
Hex 3	1157.0	—	2316.0	2316.2	1.9	P_1 Hex_3Hep_3 Kdo_1 lipidA-OH
	1218.4	812.1	2438.1	2439.2	53.1	P_1*P*Etn$_1$$Hex_3Hep_3$ Kdo_1 lipidA-OH
Hex 4	1237.3	824.7	2476.9	2478.3	7.0	P_1 Hex_4 Hep_3 Kdo_1 lipidA-OH
	1299.5	866.1	2601.2	2601.3	28.5	P_1 *P*Etn$_1$$Hex_4$ Hep_3 Kdo_1 lipidA-OH
Hex 5	1318.8	878.8	2639.5	2640.4	5.7	P_1 Hex_5 Hep_3 Kdo_1 lipidA-OH
	1381.5	920.2	2764.3	2763.5	3.8	P_1 *P*Etn$_1$$Hex_5$ Hep_3 Kdo_1 lipidA-OH

[a] Calculated average values of doubly and triply charged ions.
[b] Average mass units were used for calculation[26] of molecular mass values based on proposed compositions as follows: 162.15 for Hex, 192.17 for Hep, 220.18 for Kdo, 79.98 for phosphate, 123.05 for *P*Etn, and 953.03 for lipid A-OH ($(GlcN)_2 + (3\text{-}OH\text{-}C_{14})_2 + P_2$).

A selective NOE transfer from H-1 of GI disclosed the chemical shift values of H-2, H-3, and H-4 of HepII (Table 4). In a subsequent 1D NOESY–TOCSY experiment, a relayed selective TOCSY transfer from H-3 of HepII permitted assignment of HepII H-5 (3.78 ppm); and, a subsequent NOESY in a 1D NOESY–TOCSY–NOESY experiment led to assignment of H-6 (4.60 ppm) (Fig. 4). A ^{31}P-resonance at −0.8 ppm was correlated to this resonance (4.60 ppm) and to that of the methylene protons (4.12 ppm) of the *P*Etn residue in a ^{31}P–^{1}H HMQC experiment establishing the site of substitution. ^{31}P–^{1}H HMQC experiment also confirmed phosphate substitution

Table 4
Proton and ^{13}C chemical shifts (ppm) for major glycoform from O-deacylated LPS from *H. influenzae* strain Eagan *galE/galK*

Residue unit	Glycose	H-1 (C-1)	H-2 (C-2)	H-3a[b] (C-3)	H-3e[b]	H-4 (C-4)	H-5 (C-5)	H-6 (C-6)	H-6′	H-7 (C-7)	H-7′	H-8 (C-8)	H-8′
GNI	→6)-α-D-Glc*p*N*p*-(1→	5.40	3.85	—		—	—	—	—				
		(92.8)	(54.0)	—		—	—	—	—				
GNII	→6)-β-D-Glc*p*N*p*-(1→	4.62	3.89	3.70	—	3.82	—	—	—				
		(55.0)											
K	→4,5)-α-Kdo*p*-(2→						4.28	3.80		3.74	—	—	—
		—	—	—		—		(72.4)					
HI	→3,4)-L-α-D-Hep*p*-(1→	5.19	4.19	4.12	—	4.30	—	—		—	—		
		(99.8)	—	—		—	—	—		—			
HII	→2,3)-L-α-D-Hep*p*-(1→	5.78	4.29	4.10		4.15	3.78	4.60		3.89	3.72		
		(98.1)	(78.4)	(78.3)		(65.2)	(71.2)	(73.2)		(61.3)			
HIII	L-α-D-Hep*p*-(1→	5.02	3.84	3.95	—	—	—	—	—	—	—		
		(101.2)	(70.8)	(70.2)		—	—	—		—			
GI	→4)-α-D-Glc*p*-(1→	5.31	3.61	3.91	—	3.74	—	—	—				
		(100.1)	(71.5)	—		(77.8)							
GII	β-D-Glc*p*-(1→	4.56	3.35	3.54	—	3.44	3.51	3.94	3.76	—	—		
		(102.2)	(73.0)	—		—	—	—	—				
GIII	β-D-Glc*p*-(1→	4.52	3.32	3.52	—	3.43	3.51	3.94	3.76	—	—		
		(102.3)	(73.0)	—				(60.5)	—	—	—		

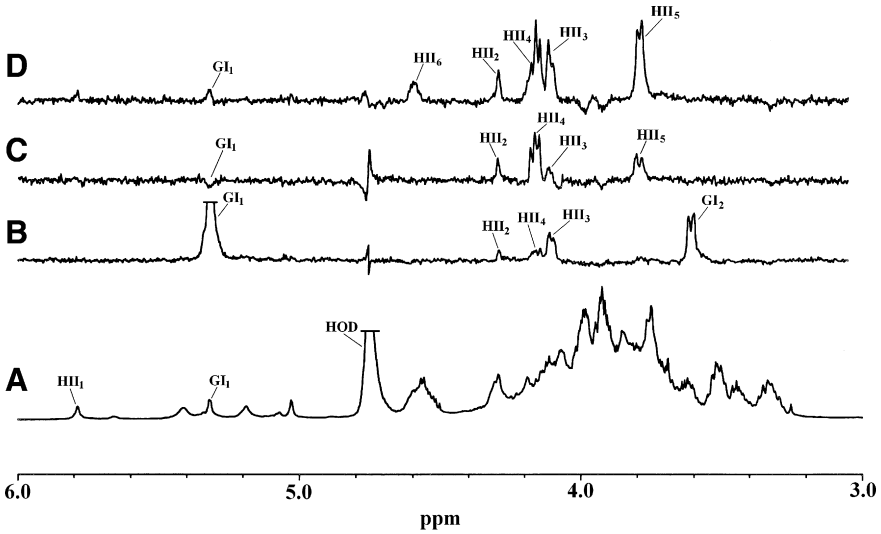


Figure 4. 1D NMR spectra of LPS-OH from *galE/galK* mutant showing ^{1}H NMR spectra of the LPS-OH sample (A), a selective NOE transfer from H-1 of GI (B), a selective NOESY–TOCSY from H-3 of HII (C), and selective NOESY–TOCSY–NOESY from H-5 of HII (D). Relevant resonances in each experiment are indicated.

at O-4 of Kdo, O-4 of GlcNII, and O-1 of GlcNI in the LPS-OH sample (data not shown), in agreement with previously reported phosphate substitutions positions in *H. influenzae* LPS.[7,12,16,28–30]

2.5. Structures of backbone oligosaccharides from *galE* and *galK* single mutants

Detailed NMR analyses revealed the structures of the HMW and LMW backbone OSs from the *galE* mutant to be identical to those of the double mutant *galE*/*galK* (data not shown).

The ^{1}H and ^{13}C NMR spectra of the backbone OS obtained from LPS of *galK* single mutant were similar to those obtained for the parent strain.[7] The OS had a similar glucose-containing basal structure to that of the LMW-OS from the double mutant but with additional substitution of HepIII at the O-2 by a β-D-*galacto*-pyranosyl residue corresponding to the OS derived from the Hex4 glycoform of the parent strain.[7] In the OS from the *galK* mutant the β-D-Glc*p*-(1→4)-α-D-Glc*p*-(1→3)- unit substituting HepII is capped by a β-1,4-linked D-Gal*p* as found in the Hex5 glycoform of the parent strain. This OS was the only one obtained from LPS of the *galK* mutant. Structures of the major LPS glycoforms of *H. influenzae gal* mutants from strain Eagan are presented in Figure 5.

3. Discussion

Recent studies have identified the genetic basis for biosynthesis of *H. influenzae* LPS oligosaccharide in the type b strain, Eagan.[6,19,31] Genes involved in galactose metabolism, *galK* and *galE*, are among those required for its utilization and interconversions. *GalK* encodes the kinase that mediates synthesis of galactose-1-phosphate from exogenous galactose and, in *H. influenzae*, is located in an operon consisting of *galT*, *galM*, and *galR*,[20] whereas *galE* is located in a different locus that includes the sialyltransferase gene, *lic*3.[32] The *galE* gene encodes the enzyme that mediates the conversion of UDP-Glc to UDP-Gal through C-4 epimerization. Mutation of either *galE* or *galK* may effect the incorporation of galactose into *H. influenzae* LPS. Thus, as previously observed,[20] a mutant strain bearing a *galE* deletion could still synthesize LPS glycoforms containing galactose when grown in rich media (galactose-containing media). Similarly, a mutant strain bearing a single *galK* deletion, synthesizes LPS molecules containing galactose. The *galK* mutant is unable to synthesize UDP-Gal from exogenous galactose via galactose kinase,[20] but still can convert internal UDP-Glc into UDP-Gal by the functional galactose epimerase. In the case of the *galK* mutant, it is interesting to note that the lactose (β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→) containing Hex5 glycoform previously identified in the parent strain[7] was found as the major LPS structural motif. The inability of the *galK* mutant to utilize an exogenous source of galactose coupled with a *galE* mutation that prevents the conversion of UDP-Glc to UDP-Gal would render the double mutant incapable of incorporating galactose into its LPS as was evident by simple sugar analysis.

We have previously reported the structure of the LPS elaborated by strain Eagan.[7] It is comprised of a conserved triheptosyl inner-core moiety in which each of the three heptose units pro-

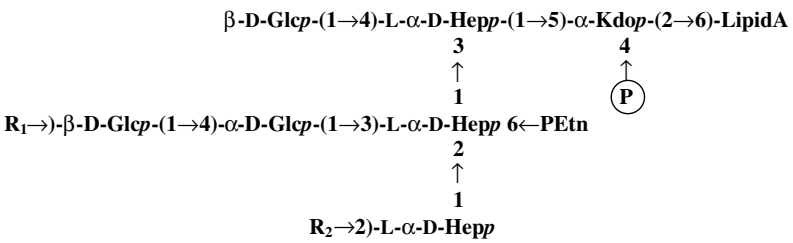


Figure 5. Structures of major LPS glycoforms of *H. influenzae gal* mutants. In double mutant *galE*/*galK* and single mutant *galE*, R_1 = H or β-D-Glc*p*, and R_2 = H, and in the single mutant *galK*, R_1 = β-D-Gal*p*, and R_2 = β-D-Gal*p*.

vides a point for oligosaccharide extension. In the present study we have now determined that the *galE*/*galK* double mutant elaborates two major glycoforms in which there is substitution from only the proximal (HepI) and middle (HepII) heptose residues. The glycoform structures were determined in detail on derived OS samples using 1D and 2D NMR methods which also served to provide unambiguous location of the inner-core *P*Etn substituent. Thus, the major glycoform (Hex3) is substituted by a single β-D-Glc*p*-(1→ residue at O-4 of HepI, a substitution pattern that has been found in every *H. influenzae* strain investigated to date,[4] and a β-D-Glc*p*-(1→4)-α-D-Glc*p*-(1→ disaccharide extension from O-3 of HepII, a structural motif previously identified in the parent strain.[7] A major Hex4 glycoform was also characterized in which the disaccharide extension off HepII is elongated by another D-Glc*p* residue in a β-1,4 linkage. The triglucoside extension (β-D-Glc*p*-(1→4)-β-D-Glc*p*-(1→4)-α-D-Glc*p*-(1→) from HepII in this glycoform has not previously been observed in *H. influenzae* LPS. In the parent strain, a D-Gal*p* residue is attached in a β-1,4 linkage to form the trisaccharide β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→4)-α-D-Glc*p*-(1→. In other type b *H. influenzae* strains, notably the *galE*/*galK* mutant of RM7004[5] and strain A2,[33] a terminal β-D-Glc*p*-(1→4)-β-D-Glc*p*-(1→ unit was found as an extension off HepI. The gene that encodes the β-1,4-glucosyltransferase leading to the synthesis of this disaccharide in strain RM7004 has been recently identified as *lex*2.[34] Although this gene is present in strain Eagan, it is nonfunctional[34] consistent with the absence of further chain elongation from HepI.[7] In strains Eagan and RM7004,[19] the galactosyltransferase *lic*2A mediates the transfer of D-Gal*p* from UDP-Gal to β-D-Glc*p*-(1→4)-α-D-Glc*p*-(1→ in a β-1,4 linkage to give the trisaccharide extension unit. It is interesting to hypothesize that, in the absence of UDP-Gal, UDP-Glc can act as a donor for *lic*2A, thereby accounting for the synthesis of the triglucoside extension from HepII in the *galE*/*galK* mutant. The observation of a Hex5 LPS glycoform in the LPS of this mutant may suggest further chain extension of this glucan to give a tetraglucoside. However, detailed structural data would need to be made to support this assertion. Promiscuity in the donor specificity of galactosyltransferases is well documented.[35] It is noteworthy that the *galE* mutant expressed the same LPS glycoform profiles as that of the double mutant when grown in the absence of an exogenous source of galactose.

4. Experimental

4.1. Organism and growth conditions

Defined *gal* mutants from *H. influenzae* type b[20] were cultivated aerobically using brain heart infusion broth (Difco) medium supplemented with 2 mg/L each of NADH (Sigma) and Hemin (Sigma) in 28- and 75-liter fermentors (New Brunswick Scientific) at 36 °C. Antibiotics were added, to maintain selection pressure, in a concentration of 1 mg/L tetracycline into mutant *galK* medium, 10 mg/L kanamycin into mutant *galE* medium, and both antibiotics into the double mutant *galE*/*galK* medium. The cells were killed, in the stationary phase, with 1% phenol (final concentration) and the cell pellets obtained by centrifugation were successively washed, once with ethanol, twice with acetone, and twice with petroleum ether (35–60 °C).

4.2. LPS extraction and OS preparation

LPS was extracted from the dried cell pellets by the hot phenol–water extraction procedure.[21] Core OS, *O*-deacylated (LPS-OH) and deacylated, dephosphorylated, and reduced LPS (LPS-backbone OS) was obtained from LPS as described earlier.[7,22,25]

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

PAGE was performed using the buffer system of Laemmli and Favre[36] as modified by Komuro and Galanos[37] with DOC as detergent. LPS bands were stained and visualized by silver staining as described by Tsai and Frasch.[38]

4.4. Analytical methods

The determinations of glycose compositions, monosaccharide absolute configurations, and linkage analysis were carried out by GLC and GLC–MS as described earlier.[7,12,22,25] For linkage analyses, OS samples (2–5 mg) were methylated with iodomethane in dimethyl sulfoxide containing an excess of potassium (methylsulfinyl) methanide, hydrolyzed with 0.25 M H_2SO_4 in 95% acetic acid at 85 °C overnight, reduced ($NaBD_4$), and acetylated.[7]

4.5. Electrospray ionization-mass spectrometry (ESI-MS)

LPS samples were analyzed on a VG Quattro (Fisons Instruments) with an electrospray ion source following O-deacylation. LPS-OH was dissolved in acetonitryl/H_2O (1:1 ratio) containing 1% acetic acid and injected into the mass spectrometer through the electrospray interface by direct infusion at 4 μl/min with a Harvard syringe pump 22. The electrospray tip voltage was 3.5 kV. The mass spectrometer was scanned from m/z 50–2500 with a scan time of 10 s. Data were collected in multichannel analysis mode, and data processing was handled by the VG data system (Masslynx).

4.6. NMR spectroscopy

For backbone OS, NMR spectra were obtained with a Bruker AMX 500 spectrometer using standard Bruker software on solutions at 27 or 37 °C in 0.5 ml D_2O, subsequent to several lyophilizations with D_2O. Proton spectra were obtained by using a spectral width of 2.40 kHz and a 90° pulse. Acetone was used as the internal standard, and chemical shifts were referenced to the methyl resonance (δ_H, 2.225 ppm; δ_C, 31.07 ppm). COSY,[39] TOCSY,[40] and NOESY[41] experiments were performed as described earlier.[7] Mixing times of 66.1 and 400 ms were used for TOCSY and NOESY, respectively. Heteronuclear 2D ^{13}C-^{1}H chemical shift correlations were measured in the ^{1}H-detected mode via HMQC[42] with proton decoupling in the ^{13}C domain. For LPS-OH, spectra were obtained with a Bruker 600 spectrometer using standard Bruker software on a solution in D_2O at 27 °C with 1.5 mM EDTA-d_{12} and 10% SDS-d_{25}, subsequent to several lyophilizations with D_2O. The measurements of the 1D and 2D spectra were recorded using the above parameters. Selective 1D-TOCSY, 1D NOESY, 1D-TOCSY–NOESY, and 1D-NOESY–TOCSY were performed for assignment of HepII. ^{31}P spectra were measured at 202 MHz by employing spectral widths of 10.2 kHz and a 90° pulse, and phosphoric acid (85%) was used as the external standard (δ_P, 0.0 ppm). ^{1}H–^{31}P correlations (HMQC) were made in the ^{1}H-detected mode by using a data matrix of 16 × 1024 points, and spectral width of 2.25 and 10.2 kHz for ^{1}H and ^{31}P domains, respectively. 2048 scans were acquired for each t_1 value, and a mixing time of 60 ms.

Acknowledgments

We thank Dr. Jean-Robert Brisson for expert assistance in conducting NMR experiments, D. Griffith for large-scale production of cells, F. Cooper for GLC–MS analysis, and D. Krajcarski for ESI-MS analysis. E.R.M. was supported by a program 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, influenzae,* Epidemiology, Immunology and Prevention of Disease; Sell, S. H., Wright, P. F., Eds.; Elsevier: New York, 1981; pp 3–9.
3. Moxon, E. R.; Maskell, D. *Haemophilus influenzae* Lipopolysaccharide: The Biochemistry of a Virulence Factor. In *Molecular Biology of Bacterial Infection: Current Status and Future Perspectives*; Harmaeche, C. E., Penn, C. W., Smyth, C. J., Eds.; Cambridge University Press: Cambridge, UK, 1992; pp 75–96.
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. H.; Jansson, P.-E.; Moxon, E. R.; Lindberg, A. A. *Carbohydr. Res.* **1995**, *272*, 213–224.
6. 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.
7. Masoud, H.; Moxon, E. R.; Martin, A.; Krajcarski, D.; Richards, J. C. *Biochemistry* **1997**, *36*, 2091–2103.
8. 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.
9. Månsson, M.; Hood, D. W.; Li, J.; Richards, J. C.; Moxon, E. R.; Schweda, E. K. H. *Eur. J. Biochem.* **2002**, *269*, 808–818.
10. Schweda, E. K.; Li, J.; Moxon, E. R.; Richards, J. C. *Carbohydr. Res.* **2002**, *337*, 409–420.
11. Månsson, M.; Hood, D. W.; Moxon, E. R.; Schweda, E. K. H. *Eur. J. Biochem.* **2003**, *270*, 610–624.
12. Masoud, H.; Martin, A.; Thibault, P.; Moxon, E. R.; Richards, J. C. *Biochemistry* **2003**, *42*, 4463–4475.
13. Yildirim, H. H.; Hood, D. W.; Moxon, E. R.; Schweda, E. K. H. *J. Biochem.* **2003**, *270*, 3153–3167.
14. Hood, D. W.; Makepeace, K.; Deadman, M. E.; Rest, R. F.; Thibault, P.; Martin, A.; Richards, J. C.; Moxon, E. R. *Mol. Microbiol.* **1999**, *33*, 679–692.
15. Risberg, A.; Alvelius, G.; Schweda, E. K. H. *Eur. J. Biochem.* **1999**, *265*, 1067–1074.
16. Risberg, A.; Masoud, H.; Martin, A.; Richards, J. C.; Moxon, E. R.; Schweda, E. K. H. *Eur. J. Biochem.* **1999**, *261*, 171–180.
17. Cox, A. D.; Hood, D. W.; Martin, A.; Makepeace, K. M.; Deadman, M. E.; Li, J.; Brisson, J.-R.; Moxon, E. R.; Richards, J. C. *Eur. J. Biochem.* **2002**, *269*, 4009–4019.
18. Li, J.; Bauer, S. H. J.; Månsson, M.; Moxon, E. R.; Richards, J. C.; Schweda, E. K. H. *Glycobiology* **2001**, *11*, 1009–2001.
19. Hood, D. W.; Deadman, M. E.; Cox, A. D.; Makepeace, K.; Martin, A.; Richards, J. C.; Moxon, E. R. *Microbiology* **2004**, *150*, 2089–2097.
20. Maskell, D.; Szabo, M. J.; Deadman, M. E.; Moxon, E. R. *Mol. Microbiol.* **1992**, *6*, 3051–3063.
21. Westphal, O.; Jann, K. *Methods Carbohydr. Chem.* **1965**, *5*, 83–91.
22. Masoud, H.; Altman, E.; Richards, J. C.; Lam, J. S. *Biochemistry* **1994**, *33*, 10568–10578.
23. Bock, K.; Pedersen, C. *J. Chem. Soc., Perkin Trans. 2* **1974**, 293–297.
24. Bock, K.; Pederson, C. *Adv. Carbohydr. Chem. Biochem.* **1983**, *41*, 27–66.
25. Masoud, H.; Perry, M. B.; Brisson, J.-R.; Uhrin, D.; Richards, J. C. *Can. J. Chem.* **1994**, *72*, 1466–1477.
26. Gibson, B. W.; Melaugh, W.; Phillips, N. J.; Apicella, M. A.; Campagnari, A. A.; Griffiss, J. M. *J. Bacteriol.* **1993**, *175*, 2702–2712.
27. Hakomori, H. I. *J. Biochem.* **1964**, *55*, 205–208.
28. Zamze, S. E.; Ferguson, M. A. J.; Moxon, E. R.; Owek, R. A.; Rademacher, T. W. *Biochem. J.* **1987**, *245*, 583–587.
29. Helander, I. M.; Lindner, B.; Brade, H.; Altmann, K.; Lindberg, A. A.; Rietschel, E. Th.; Zähringer, U. *Eur. J. Biochem.* **1988**, *177*, 483–492.
30. Schweda, E. K. H.; 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.
31. Hood, D. W.; Randle, G.; Cox, A. D.; Makepeace, K.; Li, J.; Schweda, E. K.; Richards, J. C.; Moxon, E. R. *J. Bacteriol.* **2004**, *186*, 7429–7439.
32. Maskell, D.; Szabo, M. J.; Butler, P. D.; Williams, A. E.; Moxon, E. R. *Mol. Microbiol.* **1991**, *5*, 1013–1022.
33. Phillips, N. J.; Apicella, M. A.; Griffiss, J. M.; Gibson, B. W. *Biochemistry* **1993**, *32*, 2003–2012.
34. Griffin, R. A.; Cox, A. D.; Makepeace, K.; Richards, J. C.; Moxon, E. R.; Hood, D. W. *Microbiology* **2003**, *149*, 3165–3175.
35. Palcic, M. M.; Hindsgaul, O. *Glycobiology* **1991**, *1*, 205–209.
36. Laemmli, U. K.; Favre, M. *J. Mol. Biol.* **1973**, *80*, 575–599.
37. Komuro, T.; Galanos, C. *J. Chromatogr.* **1988**, *450*, 381–387.
38. Tsai, C.-M.; Frasch, C. E. *Anal. Biochem.* **1982**, *119*, 115–119.
39. Bax, A.; Freeman, R.; Morris, G. *J. Magn. Reson.* **1981**, *42*, 164–168.
40. Subramanian, S.; Bax, A. *J. Magn. Reson.* **1987**, *71*, 325–330.
41. Kumar, A.; Ernst, R. R.; Wüthrich, K. *Biochem. Biophys. Res. Commun.* **1980**, *95*, 1–6.
42. Bax, A.; Griffey, R. H.; Hawkins, B. L. *J. Magn. Reson.* **1983**, *55*, 301–315.