Carbohydrate Research 344 (2009) 632–641

Contents lists available at ScienceDirect

Carbohydrate Research

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

A dual role for the *lex2* locus: identification of galactosyltransferase activity in non-typeable *Haemophilus influenzae* strains 1124 and 2019

Mikael K. R. Engskog [a], Håkan H. Yildirim [a], Jianjun Li [b], James C. Richards [b], Mary Deadman [c], Derek W. Hood [c], Elke K. H. Schweda [a,*]

[a] *Clinical Research Centre, Karolinska Institutet and Södertörn University, NOVUM, S-141 86 Huddinge, Sweden*
[b] *Institute for Biological Sciences, National Research Council of Canada, Ottawa, Ontario, Canada K1A OR6*
[c] *Molecular Infectious Diseases Group, University of Oxford, Department of Paediatrics, Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, Headington, Oxford 0X3 9DS, UK*

ARTICLE INFO

Article history:
Received 20 November 2008
Accepted 8 January 2009
Available online 17 January 2009

Keywords:
Haemophilus influenzae
Lipopolysaccharide
Structure analysis
lex2
Galactosyltransferase

ABSTRACT

Lipopolysaccharide (LPS) of *Haemophilus influenzae* comprises a conserved tri-L-*glycero*-D-*manno*-heptosyl inner-core moiety (L-α-D-Hep*p*-(1→2)-[*P*Etn→6]-L-α-D-Hep*p*-(1→3)-[β-D-Glc*p*-(1→4)]-L-α-D-Hep*p*-(1→5)-α-Kdo*p*) to which addition of β-D-Glc*p* to O-4 of GlcI in serotype b strains is controlled by the gene *lex2B*. In non-typeable *H. influenzae* strains 1124 and 2019, however, a β-D-Gal*p* is linked to O-4 of GlcI. In order to test the hypothesis that the *lex2* locus is involved in the expression of β-D-Gal*p*-(1→4-β-D-Glc*p*-(1→ from HepI, *lex2B* was inactivated in strains 1124 and 2019, and LPS glycoform populations from the resulting mutant strains were investigated. Detailed structural analyses using NMR techniques and electrospray-ionisation mass spectrometry (ESIMS) on O-deacylated LPS and core oligosaccharide material (OS), as well as ESIMS^n on permethylated dephosphorylated OS, indicated both *lex2B* mutant strains to express only β-D-Glc*p* extensions from HepI. This provides strong evidence that Lex2B functions as a galactosyltransferase adding a β-D-Gal*p* to O-4 of GlcI in these strains, indicating that allelic polymorphisms in the *lex2B* sequence direct alternative functions of the gene product.

1. Introduction

Haemophilus influenzae is a non-motile and non-spore-forming Gram-negative bacterium, which resides in the upper respiratory tracts of humans, mainly in the nasopharynx. *H. influenzae* strains are divided into capsulated (types a–f) and non-capsulated (non-typeable, NT*Hi*) strains, determined by the presence or absence of capsular polysaccharide.

The bacterial pathogen is a major cause of human diseases worldwide, including chronic lower respiratory tract infections and otitis media, as well as potentially fatal diseases such as meningitis and epiglottitis.[1] The potential of *H. influenzae* to cause disease depends on its surface-expressed carbohydrate antigens, capsular polysaccharide and lipopolysaccharide (LPS).[2,3]

Since the development of conjugate vaccines specific for the capsule of type b strains, the incidence of invasive childhood diseases caused by these strains has declined. However, no vaccine has been successfully developed against NT*Hi* strains. Extensive structural studies of LPS from *H. influenzae* have led to the identification of a conserved triheptosyl inner-core moiety (L-α-D-Hep*p*-(1→2)-[*P*Etn→6]-L-α-D-Hep*p*-(1→3)-[β-D-Glc*p*-(1→4)]-L-α-D-Hep*p*-(1→5)-α-Kdo*p*), in which each of the heptosyl residues can provide a point for further attachment of oligosaccharide and non-carbohydrate substituents.[4] *H. influenzae* LPS exhibits significant outer-core structural variation within and between strains. The addition of phosphate-containing substituents, including monophosphate (*P*), phosphoethanolamine (*P*Etn), pyrophosphoethanolamine (*PP*Etn) and phosphocholine (*P*Cho) as well as *O*-acyl substituents (acetate and glycine), contributes to the structural variability of these molecules.

The genes involved in LPS biosynthesis have been investigated extensively in the type b strain Eagan and in the genome reference strain Rd.[5,6] Gene functions that are responsible for most of the steps in the biosynthesis of the oligosaccharide portion of their LPS molecules have been identified. Several genetic mechanisms

Abbreviations: *An*Kdo-ol, reduced anhydro Kdo; CE, capillary electrophoresis; Gly, glycine; Hep, L-*glycero*-D-*manno*-heptose; Hex, hexose; HexNAc, *N*-acetylhexosamine; Kdo, 3-deoxy-D-*manno*-oct-2-ulosonic acid; Lipid A-OH, O-deacylated lipid A; LPS, lipopolysaccharide; LPS-OH, O-deacylated lipopolysaccharide; MS^n, multiple-step tandem mass spectrometry; Neu5Ac, *N*-acetylneuraminic acid; NT*Hi*, non-typeable *Haemophilus influenzae*; OS, oligosaccharide; *P*, monophosphate; *P*Cho, phosphocholine; *P*Etn, phosphoethanolamine; *PP*Etn, pyrophosphoethanolamine; tHep, terminal heptose.

* Corresponding author. Tel.: +46 8585 838 23; fax: +46 8585 838 20.
E-mail address: Elke.Schweda@ki.se (E. K. H. Schweda).

0008-6215/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.carres.2009.01.005

are known that contribute to the considerable LPS structural heterogeneity observed within and between *H. influenzae* strains. The expression of a number of *H. influenzae* LPS biosynthetic genes can be switched on and off—the so-called phase variation. The genes *lgtC* and *lic2A* are responsible for the phase-variable expression of the galabiose component of the globoside oligosaccharide [α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→] when extending from the distal heptose (HepIII), β-D-Glc*p* from HepI and α-D-Glc from HepII.[5,7] The phase-variable *lex2* locus comprises two closely positioned genes, *lex2A* and *lex2B*. *Lex2B* has been shown to encode a glucosyltransferase adding a second β-D-Glc*p* to O-4 of the glucose (GlcI) linked to the proximal heptose (HepI) in the inner-core of type b strains RM7004 and RM153 (Eagan).[8] A further mechanism promoting LPS structural heterogeneity has been identified in the gene required for initiation of oligosaccharide extensions from HepIII in the inner-core, *lpsA*.[9] The encoded glycosyltransferase can add a glucose or galactose through either a β-(1→2)- or a β-(1→3)-linkage, and this functional polymorphism is dependent upon the sequence of the gene present in any given strain.

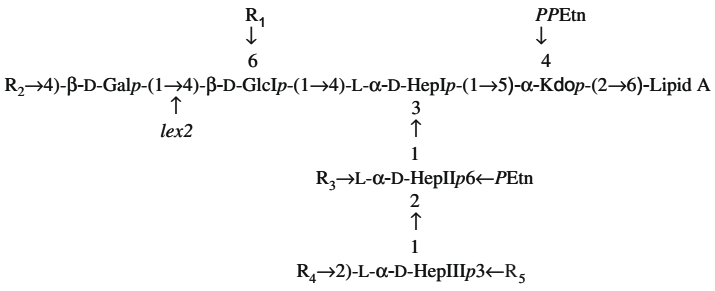
We have recently determined the distribution of glycoforms displayed by *H. influenzae* non-typeable strain 1124 (Fig. 1).[10] Interestingly, a β-D-Gal substitutes GlcI at O-4, a structure that had previously only been observed for one other NT*Hi* strain, 2019.[10,11] Moreover, subsequent addition of α-D-Gal in the biosynthesis of galabiose as part of the extension from HepI was shown to involve only *lgtC*. The galactosyltransferase involved in the attachment of β-D-Gal*p* to GlcI was found not to be Lic2A since globoside was still present as an OS extension from HepI in the mutant strain NT*Hi* 1124*lic2A*. NT*Hi* strain 1124 possesses the *lex2* locus, but no β-D-Glc*p*-(1→4-β-D-Glc*p*-(1→ OS extension has been identified.

In order to test the hypothesis that *lex2* is a phase-variable locus that might display functional polymorphisms, we inactivated *lex2B* in strains 1124 and 2019, and LPS glycoform populations from the resulting mutant strains were investigated. Detailed structural analyses indicated that both *lex2B* mutant strains express only β-D-Glc*p* extensions from HepI, providing strong evidence that Lex2B can also function as a galactosyltransferase, adding a β-D-Gal*p* to O-4 of GlcI in these strains.

2. Results

2.1. Structural characterisation of the LPS oligosaccharide from NT*Hi* strain 1124*lex2*

Strain 1124*lex2*, in which the *lex2* locus has been inactivated, was grown in liquid culture, and LPS was isolated by extraction using the phenol–chloroform–light petroleum method. O-Deacylation of LPS by treatment with anhydrous hydrazine under mild conditions afforded water-soluble material (LPS-OH), which was subjected to ESIMS analyses. The ESIMS spectrum (negative mode) showed the presence of a heterogeneous mixture of glycoforms, consistent with each molecular species containing the conserved *P*Etn-substituted triheptosyl inner-core moiety attached via a phosphorylated Kdo linked to the O-deacylated lipid A (lipid A-OH). In agreement with the structure of LPS from the wild-type strain (Fig. 1), glycoforms with a third *P*Etn were also observed (Table S1). The spectrum was dominated by multiply charged ions at *m/z* 642.0, 802.7, 1070.8 ($[M-5,4,3H]^{5,4,3-}$), and their sodiated adducts corresponding to a major glycoform with the composition *P*Cho·Hex_4·Hep_3·*P*Etn·*P*·Kdo·lipidA-OH. Minor triply charged ions at *m/z* 895.0, 962.1, 1003.0, 1015.7 and 1029.8 indicated the presence of *P*Cho·Hex_2·Hep_3·*P*Etn·*P*·Kdo·lipidA-OH, *P*Cho·Hex_4·Hep_3· *P*Etn·*P*·Kdo·lipidA-OH, *P*Cho·Hex_4·Hep_3·*P*Etn·*P*·Kdo·lipidA-OH, HexNAc·Hex_4·Hep_3·*P*Etn·*P*·Kdo·lipidA-OH and *P*Cho·HexNAc·Hex_4·Hep_3·*P*Etn·*P*·Kdo·lipidA-OH, respectively. The wild-type strain was also shown to express sialylated and disialylated glycoforms by capillary electrophoresis (CE–ESIMS/MS, negative mode) using precursor ion-monitoring experiments scanning for the diagnostic ions at *m/z* 290 (Neu5Ac) and 581 ($Neu5Ac_2$). When using *m/z* 290 as the precursor ion, the resulting spectrum of LPS-OH from 1124*lex2* revealed triply charged ions at *m/z* 1143.5, 1102.5, 1046.5 and 1005.5, together with their respective quadruply charged ions (Table S2). The ions corresponded to glycoforms having the respective compositions Neu5Ac2·*P*Cho·Hex3·Hep_3· *P*Etn·*P*·Kdo·lipid A-OH, Neu5Ac2·*P*Cho·Hex3·Hep_3·*P*Etn·*P*·Kdo·lipid A-OH, Neu5Ac·*P*Cho·Hex3·Hep_3·*P*Etn·*P*·Kdo·lipid A-OH and Neu5Ac·*P*Cho·Hex3·Hep_3·*P*Etn·*P*·Kdo·lipid A-OH. When using *m/z* 581 as the precursor ion, the resulting spectrum of LPS-OH from 1124*lex2* only



R	NT*Hi* 1124	NT*Hi* 2019
R_1	*P*Cho	H
R_2	β-D-Gal*p*NAc-(1→3)-α-D-Gal*p*-(1	H
R_3	H	Hex-Hex-Hex
R_4	β-D-Gal*p*NAc-(1→3)-α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1	H
R_5	*P*Etn	*P*Etn[a]

[a] exact position has not been reported.

Figure 1. Schematic representation of major LPS glycoforms reported for NT*Hi* strains 1124 and 2019. The various extensions R_1–R_5 are given in table inset. A triheptosyl inner-core unit is substituted at the O-4 position of HepI by a β-D-Glc*p* residue (GlcI) which, in turn, is substituted in both strains by β-D-Gal*p* at O-4. The proposed role for *lex2* in LPS synthesis is indicated by an arrow at the appropriate linkage.

revealed triply charged ions at *m/z* 1143.5 and 1102.5, together with their quadruply charged ions.

Mild acid hydrolysis of LPS with dilute aqueous acetic acid afforded insoluble lipid A and core oligosaccharide material, which after purification by gel-filtration chromatography resulted in one oligosaccharide (OS) sample, OS1124*lex2*.

OS1124*lex2* was subjected to ESIMS analysis (negative mode) and revealed, inter alia, doubly charged ions at *m/z* 765.9, 846.9, 927.8 and 1029.4, indicating the presence of *P*Cho·Hex_{2-4}·Hep_3·*P*Etn_2·*An*Kdo-ol and *P*Cho·HexNAc·Hex_4·Hep_3·*P*Etn_2·*An*Kdo-ol, respectively. In agreement with data from the parent strain, ions of their glycinated and diglycinated counterparts were also observed (Table S1). In the parent strain, glycine had been found to substitute HepIII.[10] No indications of glycoforms above HexNAc·Hex_4 were observed, thus suggesting a truncated version of glycoforms compared to the NT*Hi* 1124 wild-type strain.

Investigation of dephosphorylated and permethylated OS material by $ESIMS^n$ gave conclusive information about sequence and branching patterns within the glycoforms.[10] In the ESIMS spectrum of dephosphorylated and permethylated OS1124*lex2*, four sodiated adduct ions ($[M+Na]^+$) were observed at *m/z* 1468.2, 1671.8, 1876.0 and 2121.4 corresponding to Hex2, Hex3, Hex4 and Hex4HexNAc

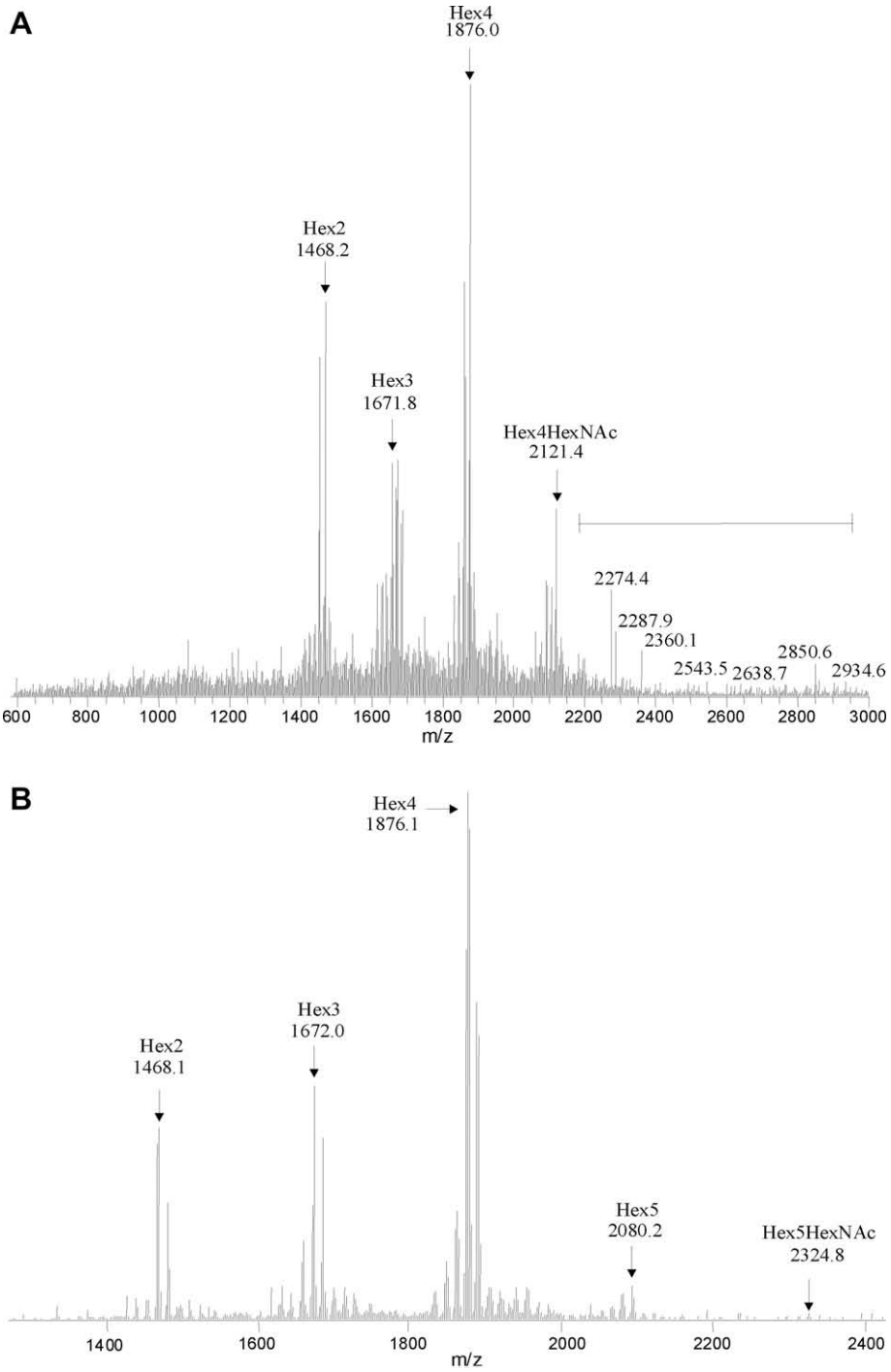
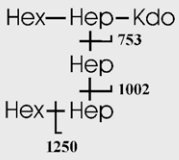
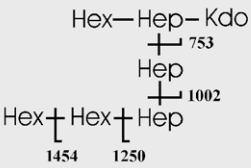
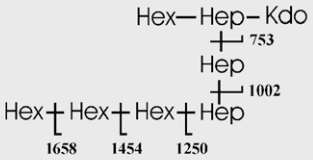


Figure 2. ESIMS spectra showing sodiated adducts $[M+Na]^+$ of permethylated OS1124*lex2* (A) and OS2019*lex2* (B). Ions indicated in (A) in the higher *m/z* region do not correspond to higher glycoforms.

Table 1
Glycoforms observed in *H. influenzae* strain 1124*lex2* as identified by ESIMS^n after permethylation[a]

Hex2	Hex3
	
Hex4	**Hex4HexNAc**
	

[a] Product ions of significant importance and the corresponding fragments are indicated.

glycoforms. Ions corresponding to higher glycoforms were not detected (Fig. 2A). One Hex2 isomer was determined by performing MS^2 on ion *m/z* 1468.2. The resulting spectrum gave, inter alia, the ion at *m/z* 1001.5 resulting from the loss of terminal (t)Hex-HepIII. Subsequent MS^3 on this ion gave the ion at *m/z* 753.4 corresponding to the loss of a HepII residue. This confirmed the structure of the Hex2 glycoform, in which terminal hexoses substitute both HepI and HepIII (Table 1).

The single Hex3 structure was determined by performing MS^2 on the parent ion at *m/z* 1671.8, resulting in fragment ions *m/z* 1453.7, 1249.6 and 1001.5 due to consecutive losses of tHex-Hex-HepIII. When the ion at *m/z* 1001.5 was further fragmented, it gave a product ion at *m/z* 753.3 due to the loss of a HepII element. Thus, in this glycoform, a disaccharide-substituted HepIII and a hexose-substituted HepI exist. Similarly, a single Hex4 glycoform was identified comprising a trisaccharide unit substituting HepIII and a hexose substituted HepI. This structure was characterised by a MS^2 experiment on the molecular ion at *m/z* 1876.0 resulting in fragment ions at *m/z* 1657.6, 1453.8, 1249.6 and 1001.5 due to the consecutive losses of tHex-Hex-Hex-HepIII. This was further supported by the MS^3 experiment on *m/z* 1001.5, in which the resulting fragment ion *m/z* 753.4 is due to loss of HepII. Performing MS^2 on the parent ion at *m/z* 2121.4 gave fragment ions at *m/z* 1861.8, 1657.9, 1454.1, 1250.1 and 1002.9 due to consecutive losses of tHexNAc-Hex-Hex-Hex-HepIII (Fig. 3). The product ion at *m/z* 1250.2 is due to the loss of tHexNAc-Hex-Hex-Hex, and further MS^3 fragmentation on this ion gave a product ion at *m/z* 1002.1 representing the loss of HepIII. Thus, the structure of the major Hex4-HexNAc glycoform comprised a tHexNAc-Hex-Hex-Hex unit substituting HepIII and tHex substituting HepI (Fig. 3).

Compositional sugar analysis of OS1124*lex2* revealed glucose (Glc), galactose (Gal), 2-amino-2-deoxygalactose (GalN), L-*glycero*-D-*manno*-heptose (Hep) in the ratio of 25:34:9:32, as identified by GLC–MS analysis of their corresponding alditol acetate derivates. Methylation analysis of OS1124*lex2* showed terminal Glc, terminal Gal, 4-substituted Gal, 4-substituted Glc, 3-substituted Gal, 2-substituted Hep, 3,4-substituted Hep, terminal GalN and 2,3-substituted Hep in the relative proportions 12:9:8:16:8:26:18:2:1.

Structural details of OS1124*lex2* were provided by ^1H–^1H chemical shift correlation experiments (DQF-COSY, TOCSY and NOESY) and heteronuclear ^1H–^{13}C and ^1H–^{31}P NMR correlation studies in the ^1H-detected mode (HMQC) as described previously. The chemical shift data are given in Table 2.

The anomeric resonances corresponding to the triheptosyl moiety, HepI, HepII and HepIII, were identified at δ 5.10–5.15, 5.75–5.83 and 5.18, respectively (Fig 4A). As observed previously, several anomeric signals were observed for HepI and HepII due to the microheterogeneity caused by the anhydro-forms of Kdo.[12] Spin systems corresponding to a terminal and substituted α-D-Gal*p* residue were identified at δ 5.04 (GalIIa) and δ 5.00 (GalIIb), respectively. The anomeric resonances at δ 4.77, 4.70, 4.60 and 4.59 could be assigned to residues GlcII, Gal*p*NAc, GlcI and GalI, respectively (Fig 4B). The interresidue NOE connectivities, observed in the NOESY spectrum (Fig 4C), between proton pairs HepIII H-1/HepII H-1,H-2 and HepII H-1/HepI H-3 confirmed the presence of the common inner-core triheptosyl moiety of NT*Hi* strain 1124*lex2*.

Interresidue NOE between Gal*p*NAc H-1/GalIIb H-3, GalIIb H-1/GalI H-4, GalI H-1/GlcII H-4 and GlcII H-1/HepIII H-1,2 established the sequence of a globotetraose unit and its attachment point to HepIII as β-D-Gal*p*NAc-(1→3)-α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→2)-L-α-D-Hep*p*-(1→ Interresidue NOE between GlcI H-1/HepI H-4,6 established the sequence β-D-Glc*p*-(1→4-L-α-D-Hep*p*-(1→ (Fig 4C).

Interresidue NOE between GalIIa H-1/GalI H-4, GalI H-1/GlcII H-4 and GlcII H-1/HepIII H-1, 2 (not visible in Fig. 4C due to low sensitivity) established the sequence of a globoside unit substituting HepIII at O-2 position in the Hex4 glycoform.

The linkage positions of the *P*EtnI, *P*EtnII and *P*Cho moieties in 1124*lex2* were as in the wild-type strain (Fig. 1).[10] The ^1H–^{31}P HMQC spectrum showed the respective ^{31}P resonances of *P*EtnI, *P*EtnII and *P*Cho at δ 0.63, −0.46 and 0.69, which correlated to H-6 of HepII at δ 4.66, H-3 of HepIII at δ 4.42 and H-$6_{a,b}$ at δ 4.34 and 4.23, respectively.

2.2. Sequence analysis of dephosphorylated and permethylated oligosaccharide derived from NT*Hi* strain 2019*lex2*

Following batch culture and LPS extraction, mild acid hydrolysis of LPS with dilute aqueous acetic acid afforded insoluble lipid A and OS material, which after purification by gel-filtration

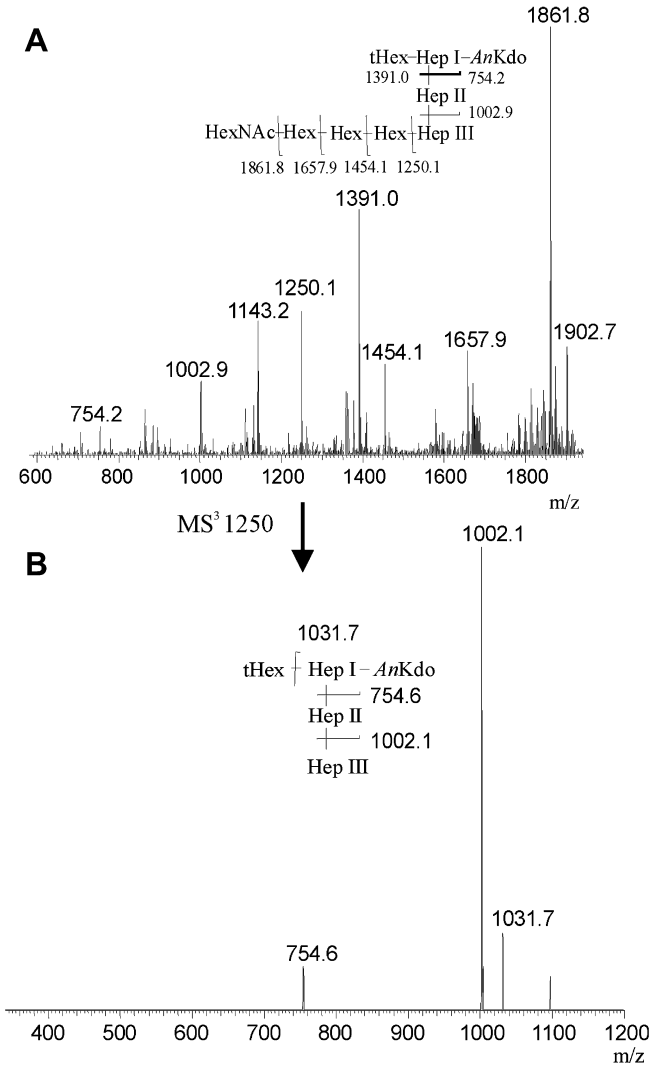


Figure 3. ESIMS^n analysis of permethylated OS1124*lex2*. (A) MS^2 spectrum of *m/z* 2121.4 corresponding to the sodium adduct of the Hex4HexNAc glycoform. (B) MS^3 spectrum of the fragment ion at *m/z* 1250.1. Proposed key fragments are indicated in the structures.

chromatography resulted in OS2019*lex2*. In the ESIMS spectrum of dephosphorylated and permethylated OS2019*lex2* (Fig. 2B), five sodiated adduct ions ([M+Na]$^+$) were observed at *m/z* 1468.1, 1672.0, 1876.1, 2080.2 and 2324.8 corresponding to Hex2, Hex3, Hex4, Hex5 and Hex5HexNAc glycoforms. As with 1124*lex2*, only isomeric glycoforms having a single Hex off HepI were detected (Table 3). Performing MS^2 on ion *m/z* 1468.1, and subsequent MS^3 gave identical results as described for 1124*lex2* (see above), that is, the Hex2 glycoform comprised a monosaccharide extension from both HepI and HepIII. Two Hex3 structures were determined from MS^2 experiments on the parent ion at *m/z* 1672.0, resulting in fragment ions *m/z* 1453.7, 1249.5, 1001.4 and 753.3 due to consecutive losses of tHex-Hex-HepIII-HepII. This confirmed the structure of a Hex3 glycoform with a mono- and disaccharide extension from HepI and HepIII, respectively. Another ion in the MS^2 spectrum at *m/z* 1409.6 corresponded to the loss of tHep (HepIII). Further MS^3 fragmentation of this ion gave ions at *m/z* 1191.4, 987.4 and 753.2 due to consecutive losses of tHex-Hex-HepII. Thus, this Hex3 glycoform comprised a disaccharide unit attached to HepII and a monosaccharide extension from HepI. Two Hex4 isomers were determined by MS^2 fragmentation of the ion at *m/z* 1876.1. In the resulting spectrum, two sets of fragment ions representing the two different isomeric structures of the Hex4 glycoform were observed (Fig. 5A). The first, represented by the consecutive losses of tHex-HepIII at *m/z* 1657.7 and 1409.6 and further MS^3 fragmentation of the ion 1409.6 gave ions at *m/z* 1191.4, 987.4 and 753.3 due to consecutive losses of tHex-Hex-HepII (Fig. 5B). This confirmed a Hex4 glycoform comprising monosaccharide extensions from both HepIII and HepI and a disaccharide unit at HepII. In the MS^2 spectrum the ion at *m/z* 1613.7 corresponded to the loss of tHep (HepIII). The second Hex4 glycoform was evident from MS^3 experiments on this ion. The MS^3 spectrum obtained contained ions at *m/z* 1396.6, 1191.5, 987.4 and 753.2 due to consecutive losses of tHex-Hex-Hex-HepII. This confirmed a Hex4 glycoform with a trisaccharide extension from HepII and monosaccharide unit at HepI. Two Hex5 isomers were determined by MS^2 fragmentation of the ion at *m/z* 2080.2. The resulting spectrum contained, inter alia, ions at *m/z* 1861.7 and 1613.9 due to consecutive losses of tHex-HepIII. The ion at *m/z* 1613.9 was fragmented further to give ions at *m/z* 1395.0, 1192.1, 988.0 and 754.4 due to consecutive loss of tHex-Hex-Hex-HepII. Thus, one of the Hex5 glycoforms comprised monosaccharides at HepIII and HepI and a trisaccharide extension from HepII. The MS^2 spectrum also contained ions at *m/z* 1861.7, 1658.0 and 1410.0 resulting from consecutive losses of tHex-Hex-HepIII. The ion at *m/z* 1410.0 was fragmented further to give ions at *m/z* 1191.8, 988.5 and 754.4 due to losses of tHex-Hex-HepII. This provided evidence of a Hex5 glycoform with disaccharide extensions at HepIII and HepII and a monosaccharide at HepI. The sole Hex5HexNAc isomer was determined by MS^2 fragmentation of the ion at *m/z* 2324.8. In the resulting spectrum, fragment ions at *m/z* 2106.1, 1861.6, 1657.5 and 1454.2 came from the consecutive losses of tHex-HexNAc-Hex-Hex, and the ion at *m/z* 2062.9 resulted from loss of tHep (HepIII). When the ion at *m/z* 2062.9 was fragmented further, it gave fragment ions at *m/z* 1846.7, 1599.6, 1395.3 and 1191.3 resulting from the loss of tHex-HexNAc-Hex-Hex. Consecutive losses of a tHex-Hex-HepII unit were indicated in ions *m/z* 1846.7, 1643.4 and 1407.0. This evidenced a glycoform having a disaccharide attached to HepII and a HexHexNAcHexHex unit linked to HepI.

3. Discussion

A characteristic feature of *H. influenzae* LPS is the extensive inter- and intra-strain heterogeneity of glycoform structure. This diversity of LPS structure is key to the role of the molecule in both the commensal and disease-causing behaviour of the bacterium. As a part of our ongoing studies on the role of *H. influenzae* LPS in disease pathogenesis and in eliciting protective immune responses, we have undertaken a systematic analysis of LPS from a genetically diverse set of NT*Hi* strains obtained from children with otitis media in Finland. Recently, we determined the structure of LPS from NT*Hi* strain 1124 in detail. It contains the conserved inner-core element L-α-D-Hep*p*-(1→2)-[*P*Etn→6]-L-α-D-Hep*p*-(1→3)-[β-D-Glc*p*-(1→4)]-L-α-D-Hep*p*-(1→5)-[*PP*Etn→4]-α-Kdo*p*-(2→6)-Lipid A of *H. influenzae* in which both the proximal and terminal heptose (HepI and HepIII, respectively) are substituted by globotetraose chains [β-D-Gal*p*-NAc-(1→3)-α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→4-β-D-Glc*p*-(1→] or the truncated versions thereof (Fig. 1).[10] Lactose [β-D-Gal*p*-(1→4-β-D-Glc*p*-(1→] attached to HepI had previously only been observed in NT*Hi* strain 2019.[11] The genes *lgtC* and *lic2A* have been shown to be responsible for the phase-variable biosynthesis of the galabiose component of the globoside oligosaccharide [α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→] when extending from the distal heptose (HepIII), GlcI from HepI and α-D-Glc from HepII.[5,7] Following the observation that in NT*Hi* strain 1124 the gene *lic2A* was not involved in the addition of α-D-Gal*p* to O-4 of β-D-Glc*p*-(1→4)-L-α-D-

Table 2
^{1}H and ^{13}C NMR chemical shifts for NT*Hi* 1124*lex2* recorded in D_2O at 22 °C and 25 °C, respectively[a]

Residue	Glycose unit	H-1/C-1	H-2/C-2	H-3/C-3	H-4/C-4	H-5/C-5	H-6_A/C-6	H-6_B	H-7_A/C-7	H-7_B
Hep I	→3, 4)-L-α-D-Hep*p*(1→	5.10–5.15 (n.r.) 97.6	4.05	4.11	4.33[b] 74.3	–[c]	4.22[b] 62.8		–	–
Hep II	→2)-L-α-D-Hep*p*(1→ 6 ↑ *P*Etn	5.75–5.83 (n.r.) 99.5	4.37 79.7	–	–	–	4.66 75.6		–	–
Hep III	→2)-L-α-D-Hep*p*(1→ 3 ↑ *P*Etn	5.18 (n.r) 100.7	4.46 77.7	4.42 75.3	4.13	–	–		–	–
Glc I	β-D-Glc*p* (1→ 6 ↑ *P*Cho	4.60 (7.8) 104.2	3.47 74.0	3.53 77.0	3.66 75.4	3.61 75.6	4.34 65.2	4.23 65.2		
Glc II	→4)-β-D-Glc*p* (1→	4.77 (7.8) 102.6	3.47 74.0	3.77 75.4	3.76 79.7	3.67 75.5	4.07 64.1	3.88 64.1		
Gal I	→4)-β-D-Gal*p*(1→	4.59 (7.8) 104.3	3.68 71.7	3.85 72.8	4.14 78.3	3.88[d] 76.3	–	–		
Gal IIa	α-D-Gal*p*(1→	5.04 (4.1) 101.4	3.91 76.2	3.97 68.6	4.10 70.1	–	–	–		
Gal IIb	→3)-α-D-Gal*p*(1→	5.00 (4.1) 101.3	3.97 –	4.05 79.7	4.33 69.8	–	–	–		
GalNAc	β-D-Gal*p*NAc(1→	4.70 (8.6) 104.3	4.02 53.4	3.82 71.7	4.02 68.6	–	–	–		
*P*Cho		4.46 60.1	3.70 66.5							
*P*EtnI		4.24 61.5	3.31 40.4							
*P*EtnII		4.29 62.5	3.37 40.4							

[a] $^{3}J_{H,H}$ values for anomeric ^{1}H resonances are given in parentheses; n.r., not resolved (small coupling). Signals corresponding to *P*Cho methyl protons and carbons occurred at 3.23/55.0 ppm.
[b] H4/H6 of HepI were identified by NOE from GlcI.
[c] –, Not determined.
[d] Tentative assignment from NOE data.

Hep*I*p-(1→, we investigated the possible role of *lex2*, which in other strains adds a β-D-Glc*p* in a (1→4)-linkage to β-D-Glc*p*-(1→4)-L-α-D-Hep*I*p.[8] The phase-variable *lex2* locus comprises *lex2A* and *lex2B*, of which the latter has been shown to encode a β-(1→4)-glucosyltransferase. We inactivated *lex2B* in NT*Hi* strains 1124 and 2019 and investigated the corresponding LPS glycoforms for truncated extensions from HepI. Strain 2019 had not previously been included in our studies of NT*Hi* LPS structure, but has been studied extensively by others.[11,13] The structure of LPS from NT*Hi* strain 1124*lex2* was determined in detail using both NMR and mass spectrometric techniques, which indicated an identical pattern of isomeric glycoforms as the parent strain (Fig. 1) with the notable lack of chain extension beyond β-D-Glc*I*p. This was particularly evident from multiple-step tandem ESIMS (ESIMS^n) experiments performed on dephosphorylated and permethylated oligosaccharide material, which is a powerful method for obtaining sequence and branching information. For most glycoforms, the presence of several isomeric compounds is revealed, and due to the increased MS response obtained by permethylation in combination with added sodium acetate, several glycoforms can be observed in the full-scan MS spectra that are not detected in underivatized samples.[10] With this sensitive method, no glycoforms having further chain extensions from GlcI in NT*Hi* strain 1124*lex2* were observed. Similarly, when we applied this method to NT*Hi* strain 2019*lex2*, no glycoforms having disaccharide extensions from HepI were observed. However, a minor glycoform could be detected having a HexHexNAcHexHex unit linked to HepI. The identity of this structure could not be established by other methods due to its limited abundance. However, previously the tetrasaccharide unit α-Neu5Ac-(2→3)-β-D-Gal*p*-(1→4)-β-D-Glc*p*NAc-(1→3)-β-D-Gal*p*-(1→ linked to O-4 of GlcI had been reported.[14] The addition of this tetrasaccharide unit to O-4 of GlcI follows a biosynthetic mechanism distinct from that of the rest of the LPS molecule and involves *en bloc* addition, similar to that seen in O-antigen biosynthesis.[14] Genetic analysis has confirmed that NT*Hi* strain 2019 contains the encoding *hmg* locus (Hood, unpublished). Thus, it is reasonable to assume that a α-Neu5Ac-(2→3)-β-D-Gal*p*-(1→4)-β-D-Glc*p*NAc-(1→3)-β-D-Gal*p*-(1→ unit substitutes O-4 of GlcI in NT*Hi* strain 2019lex2 and that *lex2* is not involved in this. The LPS structure of NT*Hi* strain 2019 has been described in a number of publications;[11,13] however, detailed structural information has only been provided for the major lactose-containing glycoform substituting

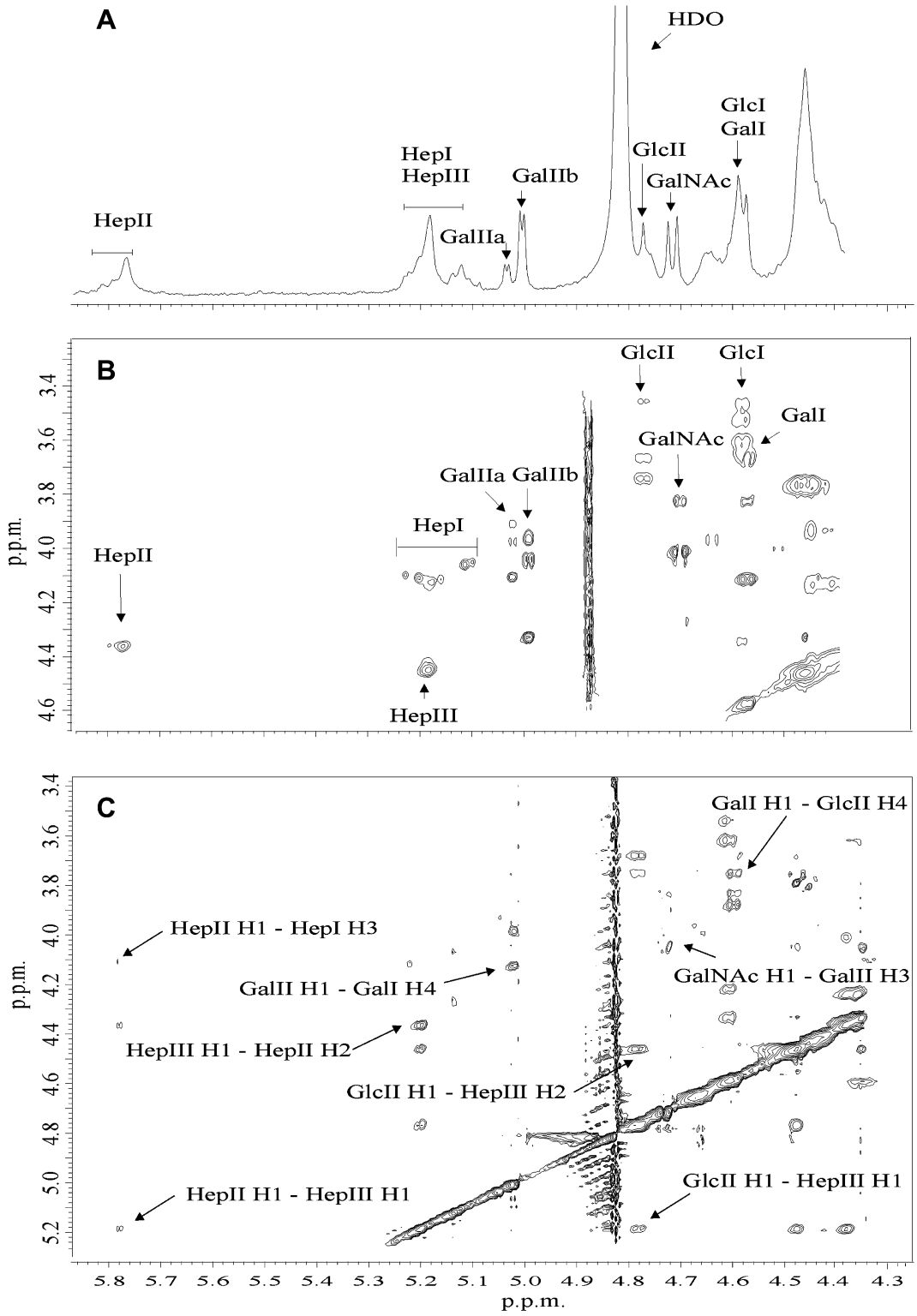
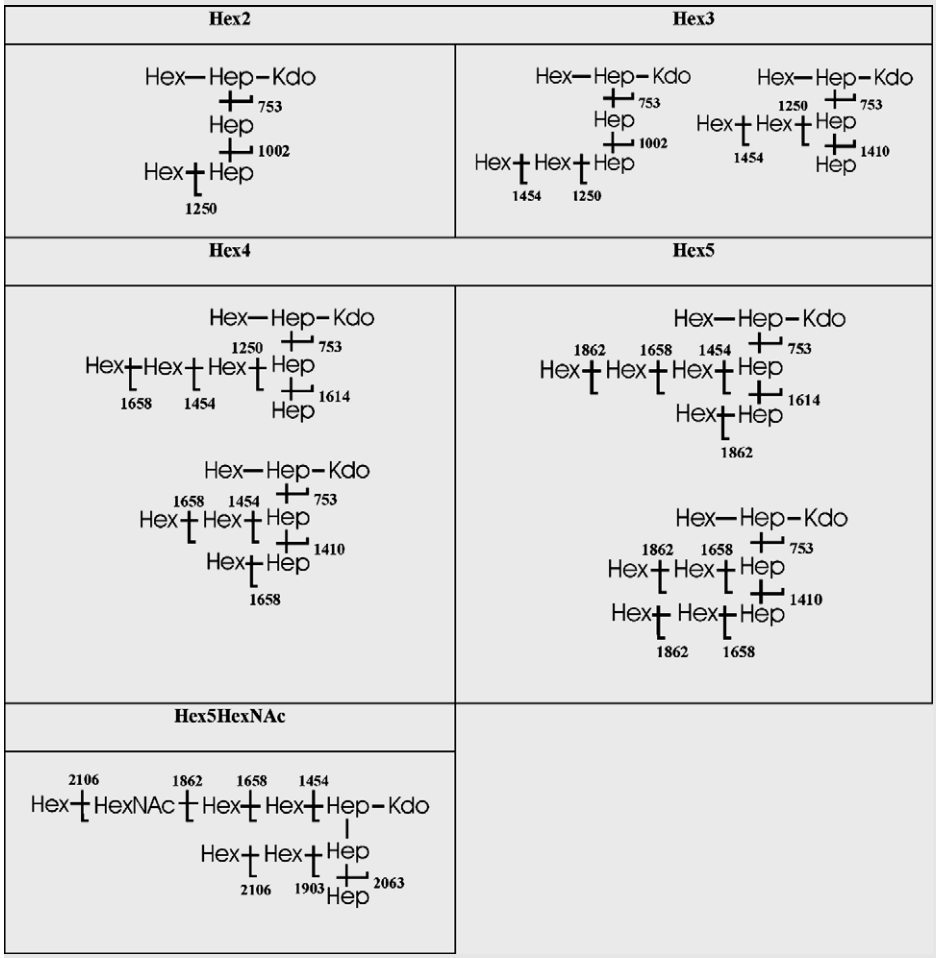


Figure 4. Selected region of 500 MHz NMR spectra of OS derived from LPS of 1124*lex2* obtained in D_2O. (A) ^1H NMR spectrum recorded at 25 °C. The anomeric resonances are indicated. (B) Two-dimensional TOCSY spectrum recorded at 22 °C. (C) Two-dimensional NOESY spectrum (mixing time 250 ms) of OS recorded at 25 °C. Cross-peaks of significant importance are labelled.

HepI (Fig. 1).[11] In some studies, HepII has also been indicated to be substituted by up to three hexoses.[15] Interestingly, our results would also indicate that HepIII can also carry oligosaccharide extensions. This has prompted us to fully evaluate the manifold of LPS isomeric glycoforms of the parent strain, which will be reported separately.

To summarize, our results indicate that *lex2* is involved in adding a β-D-Gal*p* residue to the Glc linked to HepI in NT*Hi* strains 1124 and 2019. Thus, Lex2B has β-(1→4)-galactosyltransferase activity in these strains, whereas in other strains it had been previously shown to be a β-(1→4)-glucosyltransferase. A similar observation was previously made for LpsA, which can be responsible for

Table 3
Glycoforms observed in *H. influenzae* strain 2019*lex2* as identified by ESIMS^n after permethylation



Product ions of significant importance and the corresponding fragments are indicated.

adding either a glucose or a galactose to HepIII through either a β-(1→2)- or a β-(1→3)-linkage. Each *H. influenzae* strain uniquely produces only one of the four possible combinations of linked sugars to HepIII in its LPS. We have recently shown that in any given strain, a specific allelic variant of LpsA directs the anomeric linkage and the added hexose, glucose or galactose.[16] In this study, it was found that the nature of a single key amino acid at position 151 governed the addition of either a glucose or a galactose. Since *lex2B* also appears to be polymorphic in function, a similar study of allelic variants has been carried out, details of which will be published separately.

4. Experimental

4.1. Construction of mutant strains, bacterial cultivation and preparation of LPS

NT*Hi* strains 1124 and 2019 had been described previously.[10,11,17] Strains 1124 and 2019 were mutated by transformation[18] with linearised plasmid containing the *H. influenzae lex2B* gene interrupted by insertion of a kanamycin resistance cassette to give strains 1124*lex2* and 2019*lex2*, respectively. The mutant strains were confirmed by PCR and Southern analyses as described before.[5] Bacteria were grown in brain–heart infusion broth supplemented with haemin (10 μg mL^{-1}) and NAD (2 μg mL^{-1}). LPS was extracted by the phenol–chloroform–light petroleum method, as described previously.[12]

4.2. Preparation of OS material

4.2.1. O-Deacylation of LPS

O-Deacylation of LPS was achieved with anhydrous hydrazine as described previously.[19]

4.2.2. Mild acid hydrolysis of LPS

Reduced core OS material was obtained after mild acid hydrolysis (1% aqueous acetic acid, pH 3.1, 100 °C, 2 h) as described previously.[10]

4.2.3. Dephosphorylation

Dephosphorylated OS material was obtained after treatment with 48% hydrogen fluoride as described previously.[10]

4.3. Mass spectrometry and NMR spectroscopy

GLC–MS, ESIMS and ESIMS^n were performed as described previously.[10] CE–ESIMS and CE–ESIMS/MS were carried out in the

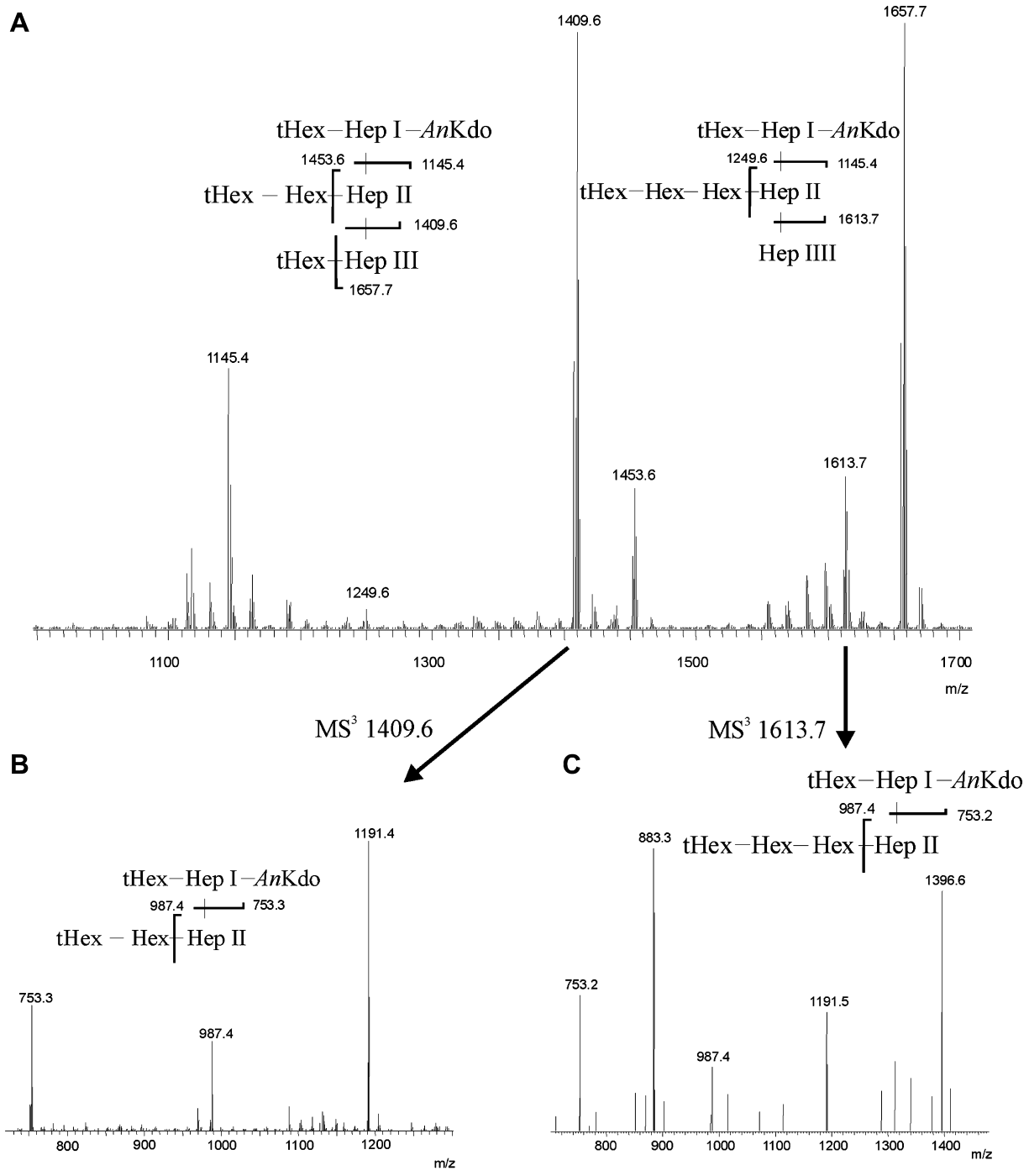


Figure 5. ESIMS^n analysis of permethylated OS2019*lex2*. (A) Product ion spectrum of $[M+Na]^+$ *m/z* 2121.4 corresponding to the sodium adduct of the Hex4 glycoform. (B) MS^3 spectrum of the fragment ion at *m/z* 1409.6. (C) MS^3 spectrum of the fragment ion at *m/z* 1613.7. Proposed key fragments are indicated in the structures.

negative-ion mode with a PrinCE-C 660 instrument (Prince Technologies, The Netherlands) coupled to a 4000 Q-Trap instrument (Applied Biosystems/MDS Sciex, Concord, Canada) via a MicroIonspray interface.

NMR spectra were obtained at 22 or 25 °C on a JEOL JNM-ECP500 spectrometer using the previously described experiments.[10]

4.4. Analytical methods

Sugars were identified as their alditol acetates as previously described.[20] Methylation analysis was performed as described earlier.[10] The relative proportions of the various alditol acetates and partially methylated alditol acetates obtained in sugar and methylation analyses correspond to the detector response of the GLC–MS.

Acknowledgements

These studies were supported by a grant from the Swedish Research Council (EKHS).

D.W.H. and M.D. were supported by a grant from the Medical Research Council, UK. We are grateful to Pascale Hermant for help with strain construction.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.carres.2009.01.005.

References

1. Campagnari, A. A.; Gupta, M. R.; Dudas, K. C.; Murphy, T. F.; Apicella, M. A. *Infect Immun.* **1987**, *55*, 882–887.
2. Anderson, P.; Johnston, R. B., Jr.; Smith, D. H. *J. Clin. Invest.* **1972**, *51*, 31–38.
3. Zwahlen, A.; Rubin, L. G.; Moxon, E. R. *Microb. Pathogenesis* **1986**, *1*, 465–473.
4. Schweda, E. K.; Richards, J. C.; Hood, D. W.; Moxon, E. R. *Int. J. Med. Microbiol.* **2007**, *297*, 297–306.
5. 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.
6. Hood, D. W.; Cox, A. D.; Wakarchuk, W. W.; Schur, M.; Schweda, E. K.; Walsh, S. L.; Deadman, M. E.; Martin, A.; Moxon, E. R.; Richards, J. C. *Glycobiology* **2001**, *11*, 957–967.
7. Masoud, H.; Martin, A.; Thibault, P.; Moxon, E. R.; Richards, J. C. *Biochemistry* **2003**, *42*, 4463–4475.
8. Griffin, R.; Cox, A. D.; Makepeace, K.; Richards, J. C.; Moxon, E. R.; Hood, D. W. *Microbiology* **2003**, *149*, 3165–3175.
9. Hood, D. W.; Deadman, M. E.; Cox, A. D.; Makepeace, K.; Martin, A.; Richards, J. C.; Moxon, E. R. *Microbiology* **2004**, *150*, 2089–2097.
10. Yildirim, H. H.; Li, J.; Richards, J. C.; Hood, D. W.; Moxon, E. R.; Schweda, E. K. *Biochemistry* **2005**, *44*, 5207–5224.
11. Phillips, N. J.; Apicella, M. A.; Griffiss, J. M.; Gibson, B. W. *Biochemistry* **1992**, *31*, 4515–4526.
12. Schweda, E. K.; Hegedus, O. E.; Borrelli, S.; Lindberg, A. A.; Weiser, J. N.; Maskell, D. J.; Moxon, E. R. *Carbohydr. Res.* **1993**, *246*, 319–330.
13. Greiner, L. L.; Watanabe, H.; Phillips, N. J.; Shao, J.; Morgan, A.; Zaleski, A.; Gibson, B. W.; Apicella, M. A. *Infect Immun.* **2004**, *72*, 4249–4260.
14. 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.
15. Gaucher, S. P.; Cancilla, M. T.; Phillips, N. J.; Gibson, B. W.; Leary, J. A. *Biochemistry* **2000**, *39*, 12406–12414.
16. Deadman, M. E.; Lundström, S. L.; Schweda, E. K.; Moxon, E. R.; Hood, D. W. *J. Biol. Chem.* **2006**, *281*, 29455–29467.
17. Borrelli, S.; Hegedus, O.; Shaw, D. H.; Jansson, P. E.; Lindberg, A. A. *Infect Immun.* **1995**, *63*, 3683–3692.
18. Herriott, R. M.; Meyer, E. M.; Vogt, M. *J. Bacteriol.* **1970**, *101*, 517–524.
19. Holst, O.; Brade, L.; Kosma, P.; Brade, H. *J. Bacteriol.* **1991**, *173*, 1862–1866.
20. Sawardeker, J. S.; Sloneker, J. H.; Jeanes, A. *Anal. Chem.* **1965**, *37*, 1602–1604.