Invited review

Innate Immunity

14(4) (2008) 199–211

Los Angeles, London,
New Delhi and Singapore
ISSN 1753-4259 (print)
10.1177/1753425908095958

Profiling structural elements of short-chain lipopolysaccharide of non-typeable *Haemophilus influenzae*

Elke K.H. Schweda[1], Brigitte Twelkmeyer[1], Jianjun Li[2]

[1]*Clinical Research Centre, Karolinska Institutet and University College of South Stockholm, NOVUM, Huddinge, Sweden*
[2]*Institute for Biological Sciences, National Research Council of Canada, Ottawa, Ontario, Canada*

Lipopolysaccharide (LPS) is a major virulence determinant of the human bacterial pathogen *Haemophilus influenzae*. A characteristic feature of *H. influenzae* LPS is the extensive intra- and inter-strain heterogeneity of glycoform structure which is key to the role of the molecule in both commensal and disease-causing behaviour of the bacterium. The chemical composition of non-typeable *Haemophilus influenzae* (NTHi) LPS is highly diverse. It contains a number of different monosaccharides (Neu5Ac, L-*glycero*-D-*manno* heptose, D-*glycero*-D-*manno* heptose, Kdo, D-Glc, D-Gal, D-GlcNAc, D-GalNAc) and non-carbohydrate substituents. Prominent non-carbohydrate components are *O*-acetyl groups, glycine and phosphates. We now know that sialic acid (*N*-acetylneuraminic acid or Neu5Ac) and certain oligosaccharide extensions are important in the pathogenesis of NTHi; however, the biological implications for many of the various features are still unknown. Electrospray ionization mass spectrometry in combination with separation techniques like CE and HPLC is an indispensable tool in profiling glycoform populations in heterogeneous LPS samples. Mass spectrometry is characterized by its extreme sensitivity. Trace amounts of glycoforms expressing important virulence determinants can be detected and characterized on minute amounts of material. The present review focuses on LPS structures and mass spectrometric methods which enable us to profile these in complex mixtures.

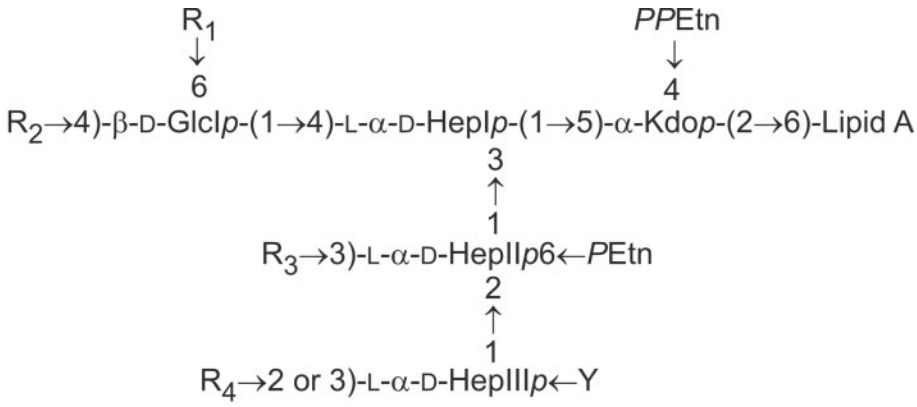
Keywords: *Haemophilus influenzae*, lipopolysaccharide, mass spectrometry, sialic acid, phosphocholine

Introduction

Haemophilus influenzae is an important commensal and pathogen of humans which commonly colonises the upper respiratory tract and is found in both encapsulated (types a–f) and non-encapsulated, so-called non-typeable (NTHi), forms. Encapsulated, mostly type b strains, can be invasive and cause bacteraemic infections such as meningitis and septicaemia, whereas NTHi are a relatively common cause of respiratory tract infections (*e.g.* otitis media, pneumonia), especially in young children. Acute otitis media is the most common cause of paediatric visits in developed countries and also the most common indication for antibiotic prescription and surgical intervention.[1,2] Glycoconjugate vaccines based on the specific capsular polysaccharide of type b *H. influenzae* have proven successful in the control of invasive *H. influenzae* type b disease in infants. Nevertheless, these vaccines do not provide protection against diseases caused by NTHi because they are only protective against infections caused by *H. influenzae* strains bearing the type b capsule. In the absence of an antigenic capsule, development of an effective vaccine seems much less straightforward. Success in these efforts requires a clear understanding of the mechanisms by which the organism is able to cause disease.

Received 30 April 2008; Revised 29 June 2008; Accepted 30 June 2008

Correspondence to: Elke Schweda, Karolinska Institutet and University College of South Stockholm, Clinical Research Center, NOVUM, S-141 86 Huddinge, Sweden. Tel: +46 8 585 838 23; Fax: +46 8 585 838 20; E-mail: elke.schweda@crc.ki.se



Scheme 1. Where R_1 = H, *P*Cho, D,D-Hep or L,D-Hep; R_2 = H, Glc, Gal or Ac; R_3 = H or Glc; R_4 = H, Glc, Gal or Ac; and Y = H, Ac, Gly, *P* or *P*Etn.

Lipopolysaccharide (LPS) is a major and essential component of the cell wall of *H. influenzae* and is also a virulence determinant for both encapsulated and non-encapsulated strains. LPS can be involved at each stage of the pathogenesis of *H. influenzae* infections, causes cytotoxic injury to target tissues and is a target for host immune responses.[3] *H. influenzae* expresses short-chain LPS, lacking O-specific polysaccharide chains and is often referred to as lipooligosaccharide (LOS).

The structural diversity of LPS expression and the genetic basis for that diversity in a representative set of NTHi clinical isolates obtained from otitis media patients has been studied in great detail over the last decade.[4] As a consequence, *H. influenzae* is now an exemplar Gram-negative bacterium for which the most extensive and detailed LPS structural data and functional correlates are available. The *H. influenzae* LPS has previously been reviewed by Apicella and co-workers in 1996 and again in 2003.[5,6] The structure, genetics and expression of variant LPS glycoforms and their implications on virulence was reviewed by ourselves in 2007.[4]

The heterogeneity and structural diversity of *H. influenzae* short-chain LPS pose significant analytical challenges. In the present review, we focus on use of recent mass spectrometric methods to profile these in complex mixtures.

The common structural element of H. influenzae *LPS*

Lipopolysaccharide from *H. influenzae* contains a glucose-substituted 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* linked to lipid A via Kdo 4-phosphate.[7–23] This inner-core unit provides the template for attachment of oligosaccharide and non-carbohydrate substituents as depicted in Scheme 1, where R_1 = H, *P*Cho, D,D-Hep or L,D-Hep; R_2 = H, Glc, Gal or Ac; R_3 = H or Glc; R_4 = H, Glc, Gal or Ac; and Y = H, Ac, Gly, *P* or *P*Etn.

H. influenzae LPS comprises extremely heterogeneous glycoform populations (see Figs 2, 4 and 6). Intra-strain variability results from variable lengths of the oligosaccharide chains attached to HepI–HepIII. This is due to incomplete biosynthesis during stepwise addition of the sugar residues and phase variable expression of particular genes.[24,25] Phase variation is the high frequency on-off switching of gene expression that can result in reversible loss or gain of specific structural elements within the LPS molecule and is further discussed below. Another source of intra-strain variation in LPS structure results from competition during biosynthesis, *i.e*. when a sugar residue is the acceptor for two or more transferase enzymes giving rise to multiple variant glycoforms. This was especially obvious for NTHi strains 1209 and 1233 in which substitution with *P*Cho competes with substitution of an external heptose to GlcI (see below).[18] Inter-strain variation primarily results from differences in the genetic blueprint for biosynthesis available in each strain. This is reflected in the presence or absence of particular biosynthetic genes and the expression pattern of these genes at any given point in time. Lipopolysaccharide of *H. influenzae* is highly substituted by non-carbohydrate substituents. This adds considerably to the heterogeneity of LPS since certain substituents can be located at several positions in the same LPS molecule and in non-stoichiometric abundances (see Fig. 4).

The lipid A backbone of *H. influenzae* LPS was shown by Helander *et al*.[26] to consist of a β-(1→6)-linked D-glucosamine disaccharide, which is phosphorylated at positions 1 and 4′. The disaccharide is primarily acylated with four 3-hydroxytetradecanoic acids. *N*-Acylation occurs at C-2 and C-2′ and *O*-acylation at C-3 and C-3′. In addition, the 3-hydroxy groups of the fatty acids at 2′ and 3′ are acylated with myristic acids.[26] Recently, the structure of lipid A isolated from 22 non-typeable and 2 type f *H. influenzae* strains were investigated.[19] The major and hexa-acyl lipid A structure comprised of a β-

Glc*p*N-(1→6)-α-Glc*p*N with phosphate group at C-1 and C-4′ positions. The N-2/N-2′ and O-3/O3′ positions were substituted by amide-linked and ester-linked 3-hydroxytetradecanoic acid chains (14:0(3-OH)), respectively. The hydroxy group of 14:0(3-OH) fatty acid chains on N-2′ and O-3′ positions were further esterified by tetradecanoic acid chains (14:0). In all strains, minor amounts of lipid A molecules with different acylation patterns were detected. In addition, lipid A with an acetyl group attached to the 14:0(3-OH) at N-2 or O-3 position was detected in two non-typeable *H. influenzae* strains.[27]

A characteristic of *H. influenzae* LPS is that terminal oligosaccharide moieties mimic host structures, generally glycosphingolipids found on human cells. This mechanism is referred to as molecular mimicry and contributes to evasion of host immune response. These host-mimicking structures include sialyllactose (α-Neu5Ac-(2→3)-β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→) and globotetraose (β-D-Gal*p*NAc-(1→3)-α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→) and truncated versions thereof (Table 1). Interestingly, the same oligosaccharide extension or non-carbohydrate substitution can occur at more than one of the heptose residues, *i.e*. globotetraose (see Fig. 1), whereas others only extend from the proximal (*i.e*. β-D-Gal*p*NAc-(1→3)-α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→6)-L-α-D-Hep*p*-(1→6)-β-D-Glc*p*-(1→) or terminal (β-D-Gal*p*-(1→) heptose residues (HepI and HepIII, respectively).

The distal heptose (HepIII) in the inner-core region has been found, depending on strain, to be substituted at either the O-2 or O-3 position (Scheme 1, R_4) by β-D-Gal*p*,[28,29] or β-D-Glc*p*.[14,30] Whereas β-D-Gal*p* has been found only as a terminal residue, strains that display a β-D-Glc*p* at this position show further extension at O-4 by β-D-GalNAc*p*-(1→3)-α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→ or α-Neu5Ac-(2→8)-α-Neu5Ac-(2→3)-β-D-Gal*p*-(1→ and/or truncated versions of these.[14,30,31]

The middle heptose of the triheptosyl moiety (HepII) has been found to be substituted at O-3 by α-D-Glc*p*. This residue has been found to be terminal[14] or to be further elongated at O-4 by globotriose (α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→)[20] or truncated analogues. HepII can also be substituted by the globotetraose-like epitope β-D-GalNAc*p*-(1→3)-α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→4)-α-D-Glc*p*-(1→ and/or truncated versions thereof.[15,23] Recently, we found α-Neu5Ac-(2→3)-β-D-Gal*p*-(1→4)-α-D-Glc*p*-(1→ to be linked to HepII.[23]

In every strain investigated to date, the proximal heptose (HepI) is substituted by β-D Glc*p* at O-4 (see Scheme 1). This residue (GlcI) is the only hexose in *H. influenzae* LPS that can appear branched, *i.e*. can have oligosaccharide elongations at O-4 and O-6 in the same glycoform. Oligosaccharide extensions from O-6 of GlcI occur only in strains having heptoses (L-α-D-Hep or D-α-D-Hep) in the outer-core. Thus extensions having β-D-Gal*p*-(1→4)-D-α-D-Hep*p*-(1→,β-D-Glc*p*-(1→4)-D-α-D-Hep*p*-(1→ or β-D-Gal*p*NAc-(1→3)-α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→6)-L-α-D-Hep*p*-(1→ and truncated versions thereof have been found at O-6 of GlcI.[17,18,32] Further extended variants of these oligosaccharides have also been detected but not fully characterized due to their low expression. Structures identified as having elongations from O-4 of GlcI are β-D-Gal*p*NAc-(1→3)-α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→, β-D-Gal*p*NAc-(1→3)-α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→, α-Neu5Ac-(2→3)-β-D-Gal*p*-(1→4)-β-D-Glc*p*NAc-(1→3)-β-D-Gal*p*-(1→ and their truncated variants.[17,22,33] Sialyllactose (α-Neu5Ac-(2→3)-β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→) and its disialylated counterpart were indicated as extensions from HepI.[22] In addition, tetrasaccharide units bearing either terminal (*P*Etn→6)-α-D-Gal*p*NAc or sialic acid linked to O-4 of GlcI have been reported.[34]

Both inner- and outer-core glycosyl residues of *H. influenzae* LPS can carry phosphate-containing substituents. In every strain investigated to date, Kdo has

Table 1. Structural motifs that have been identified to extend from HepI, HepII and HepIII of the inner-core region of *H. influenzae* LPS

Extension[a,b]	Heptose
β-GalNAc(1→3)-α-Gal(1→4)-β-Gal(1→4)-β-Glc(1→	HepI, HepII[c], HepIII
β-GalNAc(1→3)-α-Gal(1→4)-β-Gal(1→4)-β-Glc(1→4)-β[c]-Glc(1→	HepI, HepII
β-GalNAc(1→3)-α-Gal(1→4)-β-Gal(1→4)-β-Glc(1→4)[-β-Gal(1→4)-D-α-D-Hep(1→6]-β-Glc(1→	HepI
β-Glc(1→4)-D-α-D-Hep(1→6)-β-Glc(1→	HepI
β-GalNAc(1→3)-α-Gal(1→4)-β-Gal(1→6)-L-α-D-Hep(1→6)-β-Glc(1→	HepI
α-Neu5Ac(2→3)-β-Gal(1→4)-β-GlcNAc(1→3)-β-Gal(1→4)-β-Glc(1→	HepI
α-Neu5Ac(2→3)-β-Gal(1→4)-β[c]-Glc(1→	HepI, HepII, HepIII
α-Neu5Ac(2→8)-α-Neu5Ac(2→3)-β-Gal(1→4)-β-Glc(1→	HepI, HepIII
β-Gal(1→	HepIII

[a]All sugars are pyranosidic; all hexoses have D configuration.
[b]Shown are most extended chains. Truncated versions are possible.
[c]Is α-Glc(1→ when linked to HepII.

been found to be substituted at O-4 by *PP*Etn, while HepII carries *P*Etn at the O-6 position. HepIII has been found substituted by *P* or *P*Etn in some strains.[11,17,22] Substitution by *P*Cho is a common feature of *H. influenzae* LPS. A majority of *H. influenzae* strains including strain Rd[12] and a majority of NT*Hi* strains have been shown to carry *P*Cho at *O*-6 of GlcI; in other strains, including *H. influenzae* type b strains, RM7004 and Eagan, *P*Cho is found at *O*-6 of a terminal β-D-Gal*p* residue at HepIII (Scheme 1: R_4 = *P*Cho→6-β-D-Gal*p*);[13,28] in a third structural variant, *P*Cho has been found at *O*-6 of the α-D-Glc*p* residue at HepII (R_3 = *P*Cho→6-α-D-Glc*p*).[14,23] Moreover, certain strains of *H. influenzae* can elaborate LPS glycoforms that carry two *P*Cho substituents.[28,35]

Ester-linked glycine is a prominent substituent in the inner core region of *H. influenzae* LPS.[36] Recent studies by our group would suggest that all strains are capable of expressing minor amounts of this amino acid on their LPS. HepIII is most frequently substituted by glycine although it can also be found on HepII or Kdo. In certain strains, two sites for glycine substitution have been identified; however, due to the low abundance of this substituent, the exact linkage positions have not as yet been determined. Both outer- and inner-core residues can be *O*-acetylated. Thus, HepIII has been found to carry *O*-acetyl substituents at the O-2[14] or O-3 position.[16] *O*-Acetyl groups are also found in some strains attached at the hexose linked to HepIII via O-6 or O-4 linkages[28,37] or at GlcI.[13,18] We have also observed acetylation of O-2 of HepI.[21]

Biosynthesis of the oligosaccharide portion of NTHi LPS

The availability of the complete genome sequence of *H. influenzae* strain Rd[38] has facilitated a comprehensive study of LPS biosynthetic loci in the homologous strain RM118[24] and in the type b strains Eagan (RM153) and RM7004[25]. Gene functions have been identified that are responsible for most of the steps in the biosynthesis of the oligosaccharide portion of their LPS molecules. The recent completion and release of further genome sequences for NTHi strains[39,40] offers additional opportunity to identify novel LPS related genes. The genes required for oligosaccharide initiation from each of the three heptose residues in the inner core have been identified as *lgtF* (GlcI to HepI), *lic2C* (α-D-Glc*p* to HepII) and *lpsA* (HepIII).[41] Homologous genes having similar functions have been identified in a wide range of nontypeable strains. The *lgtF* and *lpsA* genes are invariably present in each strain examined while *lic2C* is found in less than half the strains. Glucose or galactose can be added to HepIII through either a β-(1→2) or 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.[42] In that study, it was found that the nature of a single key amino acid at position 151 governed the addition of glucose or galactose and that the 3′ end of the gene directs the anomeric linkage [β-(1→2) or β-(1→3)] of the added hexose. Thus, allelic variation is another important mechanism contributing to the heterogeneity of LPS glycoforms between different strains.

The biosynthesis of subsequent oligosaccharide extensions strongly depends on the expression of phase variable genes. Phase variation enables the organism to adapt to changes in host environment and thus to evade immune responses.[43] LPS phase variation is mediated by polymerase slippage in multiple tandem tetranucleotide repeats in six characterized chromosomal loci: *lic1*, *lic2*, *lic3*, *lgtC*, *lex2* and *oafA*.[44–48] Changes in the number of repeat units within a tract provide an on-off switch for translation of the relevant reading frame. The *lex2* locus has been shown to encode a glucosyltransferase (Lex2B) adding a second β-D-Glc*p* to the glucose (GlcI) linked to HepI in the inner-core and is important for further oligosaccharide extension.[47] Preliminary results from our laboratory indicate that, in certain strains, an allelic variant of Lex2B adds galactose to *H. influenzae* LPS (Hood DW, unpublished observations). It has been demonstrated that expression of *P*Cho substituents in *H. influenzae* LPS is subject to phase variation mediated by the *lic1* locus.[49] The *lic1* locus comprises four genes (*lic1A* to *lic1D*) of which *lic1D* encodes a diphosphonucleoside choline transferase. *P*Cho has been found in different molecular environments depending on the strain. The observed donor specificity of *lic1D* is associated with DNA sequence polymorphism in *lic1D*.[50] Some strains expressing two *P*Cho residues have duplicate copies of *lic1* directing the addition of *P*Cho in the lipopolysaccharide of *H. influenzae*.[35] Genes comprising the *lic2* locus have been shown to be required for chain extension from HepII[41] and, together with *lgtC*, the *lic2A* gene is 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 HepIII. The globoside epitope was also identified as an extension attached directly to HepI. In this extension, only *lgtC* was shown to be involved in the biosynthesis of galabiose.[22] As mentioned above, an allelic variant of Lex2B acts as a galactosyltransferase in this case. The *lic3* locus encodes an α-2,3-sialyltransferase that is responsible for addition of sialic acid (*N*-acetylneuraminic acid or Neu5Ac) to terminal lactose elongating from HepIII.[51] In addition, the genes *lsgB* and *siaA* are involved in sialylation of the

lacto-*N*-neotetraose units, extending from HepI.[34] The phase variable gene *oafA* is involved in *O*-acetylation of *H. influenzae* LPS. Comparative structural analyses of LPS from wild-type and *oafA* mutant strains have demonstrated a role for this gene in the addition of *O*-acetyl groups to the distal heptose of the inner-core moiety.[48] The gene *lpt6* is responsible for the addition of *P*Etn to HepII in strain Rd.[52] Following the recent completion and release of further genome sequences for NT*Hi* strains,[39,40] a candidate gene for a novel heptosyltransferase (*losB*) was identified through homology to a gene characterized in the related species *Haemophilus ducreyi*[53] that has recently been shown to direct the incorporation of a fourth heptose in the LPS of this species.[54] From the genome of NTHi R2486, two heptosyltransferase candidate genes, *losB1* and *losB2*, were identified which were predicted to direct the addition of the external heptose in the inner-core moiety of NTHi strain R2486. When *losB1* was inactivated in strain R2486, D,D-Hep was not expressed but minor quantities of L,D-Hep were found in its place in the outer-core region. Lipopolysaccharide from the R2846 *losB1*/*losB2* mutant showed no external heptose, leading us to conclude that the *losB2* gene is directing the addition of L,D-Hep to the same position (Lundström SL, in press).

To determine the structural genetic blueprint for LPS synthesis in *H. influenzae*, structural analysis of various mutant strains is fundamental. In addition, genetic manipulation of key genes is a powerful tool to aid structural analysis. In order to facilitate structural analysis, construction of defined mutants in known biosynthetic genes limits some oligosaccharide extensions, thus simplifying analyses of the remaining glycoforms in their LPS. *LpsA* mutant strains have been investigated since they do not exhibit chain elongation from HepIII in their LPS, but would be otherwise structurally identical to the LPS of the parent strain.[22,23] *lic2A* and *lgtC* mutants include glycoforms sequentially truncated in globoside extensions from HepIII [*i.e.* α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→] aiding in the structural determination.[22]

Structural features of particular importance

Sialic acid (Neu5Ac) in *H. influenzae* LPS is a major virulence factor and helps the bacteria to resist the killing effect of normal human serum.[31] Sialic acid capped oligosaccharide extensions appear from all heptoses of the inner-core (see above). In a survey of 24 clinical isolates, representative of the genetic diversity of the natural population of NTHi strains, Neu5Ac was shown to be present in each strain.[55] The level of sialylation in *H. influenzae* LPS is generally very low but, when completely absent, the corresponding strain is non-virulent. This was demonstrated when chinchillas were infected with genetically diverse NTHi strains and developed acute otitis media that lasted for several weeks while

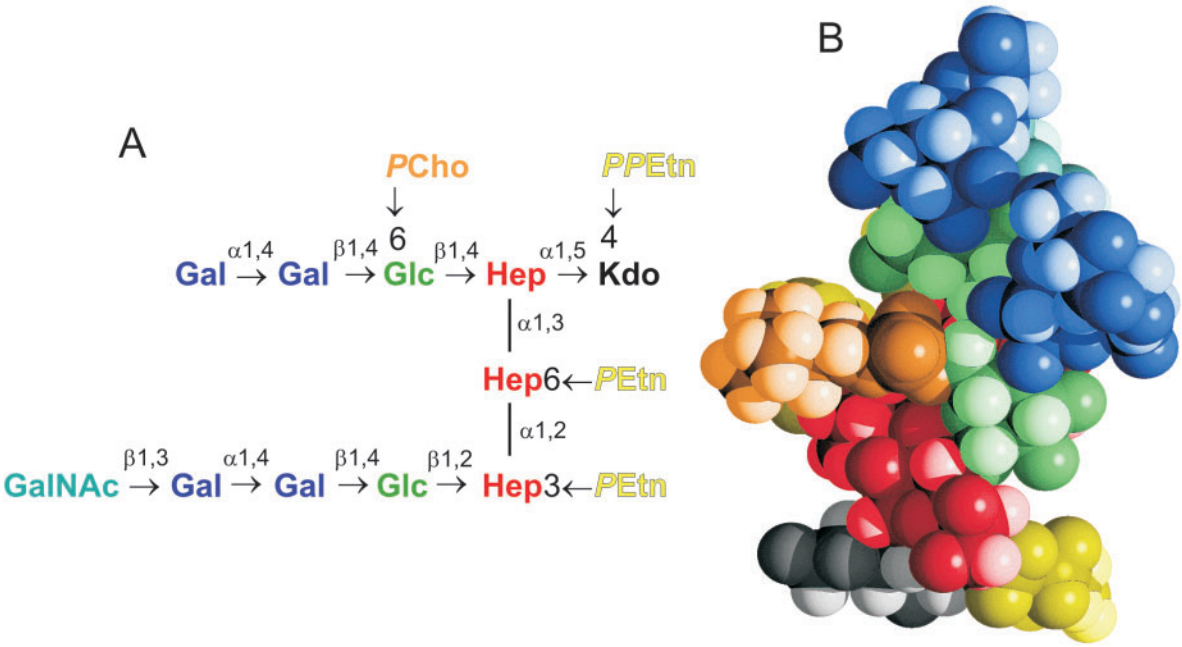


Fig. 1. LPS from non-typeable *H. influenzae* strain 1124 comprises globotetraose (β-D-Gal*p*NAc-(1→3)-α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→4)-β-D-Glc*p*) and its truncated versions globoside (α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1⟶4)-β-D-Glc*p*) and lactose (β-D-Gal*p*-(1→4)-β-D-Glc*p*) linked to both HepI and HepIII at the same time. (A) Structure (without lipid A) of the major HexNAcHex6 LPS glycoform of NTHi strain 1124. (B) Space-filling molecular model of its minimum energy conformer calculated by a Monte Carlo method.[74] Heptoses are depicted as red, Kdo is black, glucoses are green, galactoses are blue, *N*-acetyl-galactoseamine is blue-green, *P*Cho is orange, *PP*Etn and *P*Etn are yellow.

infection was absent or attenuated in chinchillas that were inoculated with sialic deficient mutants (*lic3A* or *siaB*). Sialic acid has also been shown to be important in the pathogenesis of otitis media in the gerbil middle ear infection model.[56]

In *H. influenzae* LPS, *P*Cho plays a role in persistence of the bacterium on the mucosal surface of the nasopharynx, at least in part, by mediating bacterial adherence to, and invasion of, the host epithelia,[57] which it has been proposed occurs through the receptor for platelet activating factor (rPAF) on bronchial epithelial cells.[58] *P*Cho expressed on the LPS of the organism is the target for the serum component C-reactive protein (CRP) which, when bound, mediates killing of the bacteria via the activation of complement.[43,57] However, CRP-mediated killing depends on the molecular environment of *P*Cho. Only strains expressing *P*Cho on chain extensions from HepIII rather than HepI were shown to be sensitive to CRP-mediated killing.[50] *P*Cho expression on the LPS of *H. influenzae* has also been associated with increased resistance to host antimicrobial peptide killing.[59]

The host-mimicking terminal digalactoside (α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→) has been shown to give resistance to killing by naturally acquired antibody and complement present in human serum.[43] Expression of two digalactoside extensions, rather than one (Fig. 1), in isogenic strains has been shown to be associated with increased virulence in an *in vivo* model of *H. influenzae* infection.[60]

O-Acetylation of *H. influenzae* LPS occurs frequently although not in all strains. Inactivation of the *oafA* gene led to removal of *O*-acetyl groups from the inner core of the LPS in two NTHi strains and to an increase in the killing effect of normal human serum on the bacterium when compared to wild-type.[48]

The biological importance of glycine and *P*/*P*Etn substitution in *H. influenzae* LPS is still unknown. Their role in virulence is of interest and needs to be established.

Strategies to unravel structural features of importance

As described above, short-chain *H. influenzae* LPS is comprised of heterogeneous populations of low-molecular-mass hydrophilic oligosaccharide components that are covalently linked to a hydrophobic lipid A moiety (see Scheme 1). LPS molecules are poorly soluble in aqueous media as well as common organic solvents. Thus, structural profiling of *H. influenzae* LPS involves, in most cases, initial de-lipidation to obtain water-soluble oligosaccharides that are suitable for subsequent analyses by nuclear magnetic resonance (NMR) and mass spectrometric (MS) methods. One method of choice is *O*-deacylation of LPS by treatment with anhydrous hydrazine under mild conditions.[61] This procedure provides material (LPS-OH) devoid of ester-linked fatty acids in the lipid A; however, any other ester-linked substituents are also removed. Mild acid hydrolysis of LPS with dilute aqueous acetic acid, on the other hand, affords insoluble lipid A and core oligosaccharide in which *O*-acyl substituents remain intact. However, oligosaccharide is devoid of any other acid labile component, *e.g.* Neu5Ac. The basic procedures for handling LPS have been reviewed.[62]

NMR spectroscopy is an indispensable tool in analytical carbohydrate chemistry. It provides information on the identity and anomeric configuration of sugar residues as well as their sequence and linkages. This information is absolutely necessary for the determination of novel structures. In general, oligosaccharide is more suitable for NMR analyses than LPS-OH. Spectra of LPS-OH in D_2O often show broad undefined peaks, even at higher temperature, due to molecular aggregation or micelle formation. This can be circumvented by the addition of deuterated EDTA and SDS to the sample solution.[37] NMR applications involve the use of homo- and heteronuclear ^{1}H, ^{13}C and ^{31}P correlation spectroscopy. One-dimensional ^{1}H-NMR is useful for profiling small amounts of LPS-OH or oligosaccharide samples. It provides valuable information on the triheptosyl inner-core region, attachment of *P*Cho units, sialic acid residues (from LPS-OH samples) and *O*-acyl substituents (from oligosaccharide samples). Two-dimensional NMR techniques provide more detailed information but demand also larger sample amounts. The inherent low sensitivity of NMR is a major disadvantage of the technique when compared to other methods since milligram amounts of material are generally needed for analysis. Moreover, due to significant overlap of signals from all glycoforms present in a heterogeneous mixture, individual components cannot always be unambiguously characterized.

Recent advances in MS have facilitated in-depth analysis of LPS. Ionization methods involving electrospray (ESI) and matrix-assisted laser desorption (MALDI) in conjunction with quadrupole, ion trap, time-of-flight and Fourier transform mass spectrometers have provided detailed information on both core oligosaccharide and lipid A structures.[63,64] We have extensively used ESI-MS which has proved to be particularly suitable for profiling the structural diversity of short-chain LPS of which the general aspects have been reviewed.[62,65] Electrospray-MS is a soft ionization method that can easily be combined with advanced separation techniques, including HPLC or capillary electrophoresis (CE).[66,67] It is readily applied to LPS-OH and oligosaccharide in either the negative or positive ion mode. Tandem mass spectrometry experiments (MS/MS and MS^n) are of particular value since they can provide

evidence for the structural features important in virulence (as mentioned above) on minute amounts of material. Recently, the method was applied to a single-colony quantity of intact LPS.[68] CE-ESI-MS is especially suitable for the determination of closely related LPS glycoform and isoform families by exploiting differences in their unique molecular conformations and ionic charge distributions by electrophoretic separation. On-line CE-MS also provides an additional avenue to improve detection limits. Aspects of CE-ESI-MS have been reviewed in 2000 and 2003 by Thibault *et al*.[69,70] and recently by Li *et al*.[67]

Profiling glycoform populations in LPS-OH by ESI-MS

Negative ion ESI-MS has been extensively employed to profile glycoform populations in LPS-OH and oligosaccharide derived from NTHi. The ESI-MS spectra of LPS-OH typically reveal multiply charged ions (doubly, triply, *etc*.) corresponding to the various molecular species present in the mixture. The composition of these molecular species can be proposed using average mass units of known components (Hex, 162.14; HexNAc, 203.19; Hep, 192.17; Kdo, 220.18; *P*, 79.98; *P*Etn, 123.05; *P*Cho, 165.13; *An*Kdo-ol, 222.20; Gly, 57.05; Ac, 42.04; lipid A-OH, 953.02, *etc*.) for calculation. To deduce their structure by employing MS/MS techniques is often ambiguous. However, MS/MS on sodiated adducts can give information on important structural elements (see below). Figure 2 shows the ESI-MS spectrum for LPS-OH derived from NTHi strain 1268 for which the major LPS structures were recently identified.[23] The MS data indicated the presence of heterogeneous mixtures 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). Two major glycoforms are observed as triply-charged ion at *m/z* 813.2 and 854.2, corresponding to respective compositions *P*Cho·Hex_2·Hep_3·*P*$Etn_{1–2}$·*P*·Kdo·lipid A-OH. Less prominent glycoforms are indicated for which the triply charged ion corresponding to the most extended one observed in this experiment, *P*Cho·HexNAc·Hex_5·Hep_3·*P*Etn·*P*·Kdo·lipid A-OH, is at *m/z* 1042.8.

Due to their very low abundance and/or extensive overlap with those corresponding to major, non-sialylated glycoforms, ions corresponding to sialylated glycoforms are rarely identified in the full ESI-MS spectra of LPS-OH samples. However, their presence can be confirmed in precursor ion monitoring MS/MS experiments (negative ion mode) by scanning for loss of *m/z* 290 (Neu5Ac) and/or *m/z* 581 (Neu5Ac2). Using this approach, the presence of important sialylation sites in *H. influenzae* LPS were recently detected.[22,23] Figure 3A shows the precursor ion scan spectrum (negative ion mode) of LPS-OH derived from NTHi strain R2846 *losB1* (unpublished data). In this strain, the heptosyl-

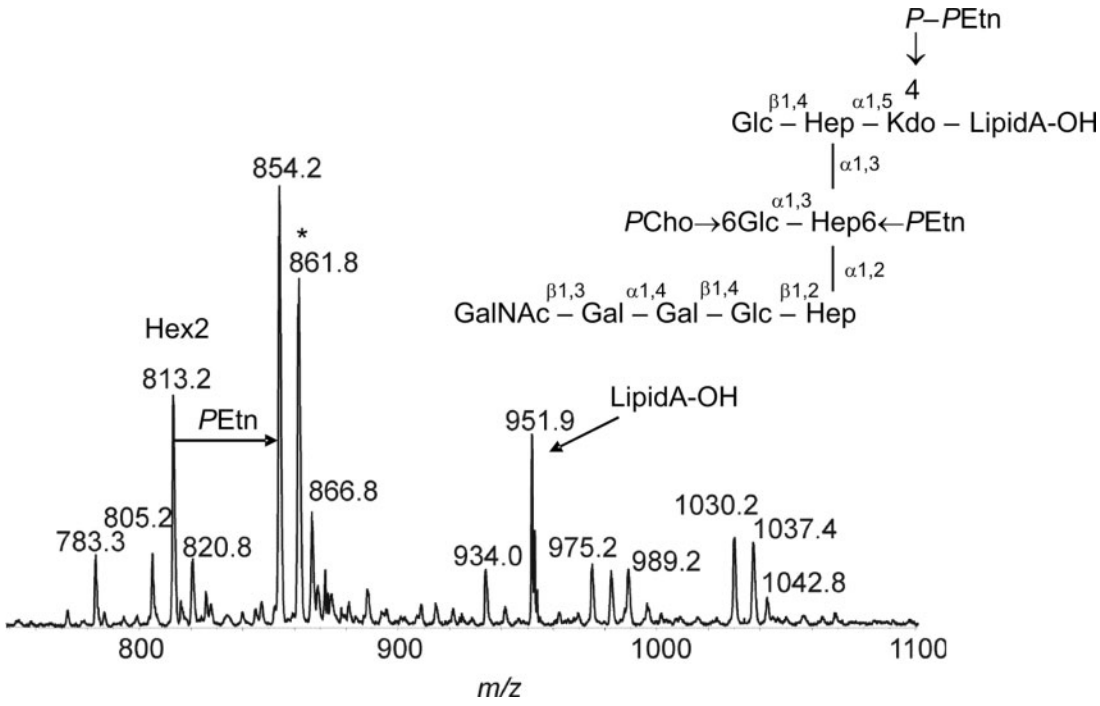


Fig. 2. Negative ion CE-ESI-MS of the triply charged molecular ion region of LPS-OH from NTHi strain 1268. Peaks corresponding to the major Hex2 glycoforms differing in the degree of substitution by *P*Etn are indicated. The structure of the HexNAcHex5 glycoform (*m/z* 1042.8) is shown in inset. *Indicates sodium adduct. The experiment was carried out with a Crystal model 310 CE instrument (ATI, Unicam, Boston, MA, USA) coupled to an API 3000 mass spectrometer (Perkin-Elmer/Sciex, Concord, ON, Canada) via a MicroIonspray interface.[23]

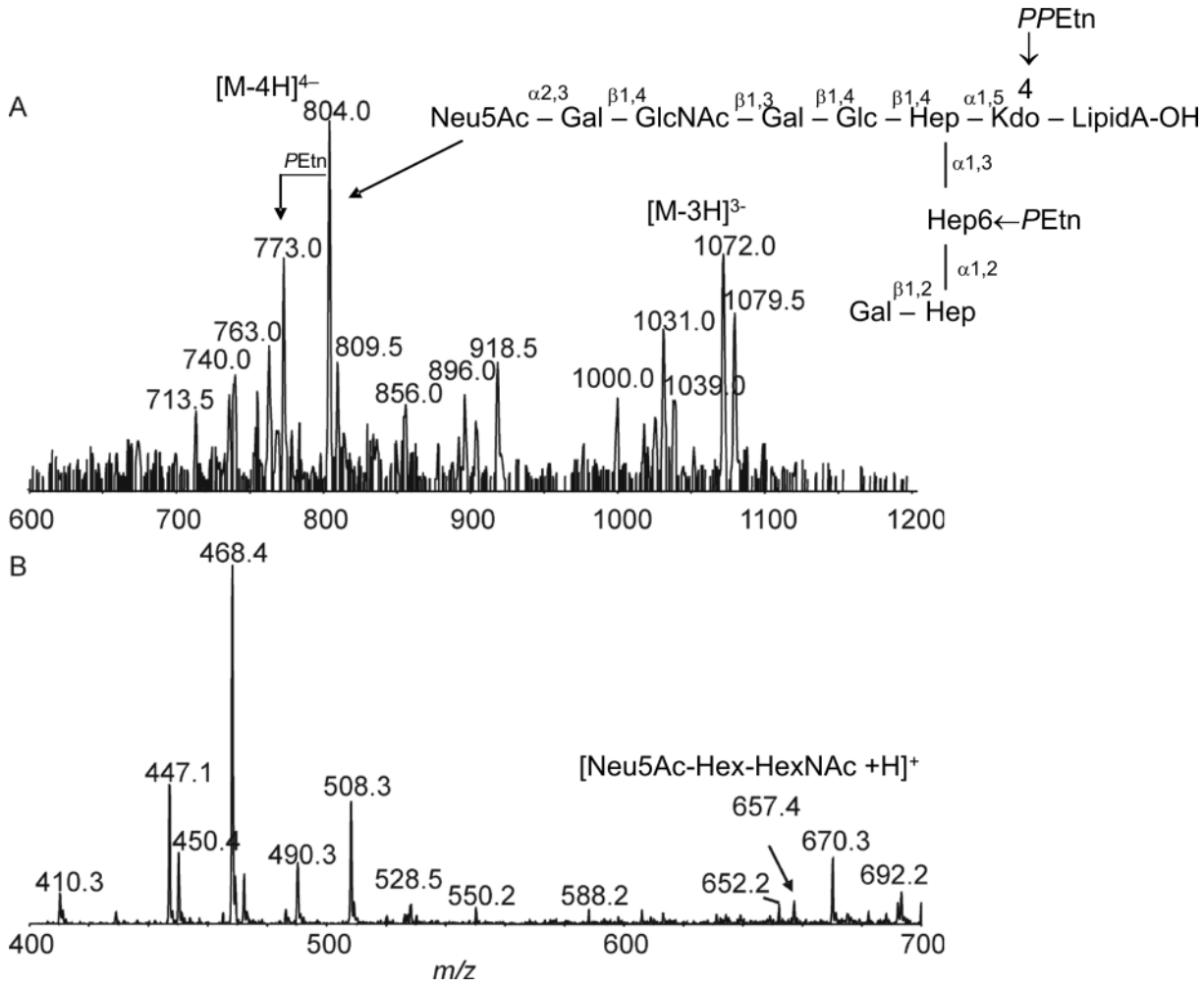


Fig. 3. CE-ESI-MS/MS spectra of LPS-OH derived from NT*Hi* strain R2846 *losB1*. (A) Precursor ion spectrum (negative ion mode) using *m/z* 290 as the fragment ion for identification of sialylated components in R2846 *losB1*. The indicated triply and doubly charged ions correspond to glycoforms with the composition Neu5Ac·HexNAc·Hex_4·Hep_3·*P*$Etn_{1–2}$·*P*·Kdo·lipid A-OH for which the structure is shown in inset. (B) ESI-MS/MS spectrum (positive mode) of the ion at *m/z* 1081 corresponding to the sodiated adduct of Neu5Ac·HexNAc·Hex_4·Hep_3·*P*Etn_2·*P*·Kdo·lipid A-OH. Indicated is the ion at *m/z* 657.4 corresponding to [Neu5Ac Hex HexNAc + H]$^+$. The experimental conditions are described elsewhere.[23,71]

transferase LosB1 was inactivated. The triply and doubly charged ions at *m/z* 773.0/1031.0 and 804.0/1072.0 correspond to glycoforms with the composition Neu5Ac·HexNAc·Hex_4·Hep_3·*P*$Etn_{1–2}$·*P*·Kdo·lipid A-OH indicating a structure in which sialyllacto-*N*-tetraose substitutes HepI and Gal substitutes HepIII of the inner-core triheptosyl moiety. More evidence for this can be obtained by MS^2 experiments. Li *et al.*[71] recently reported on structural characterization of sialylated glycoforms of *H. influenzae* by electrospray mass spectrometry using fragmentation of protonated and sodiated ions. The MS^2 spectra of protonated ions provided unambiguous evidence for the presence and sequence of sialylated lactosamine present in lacto-*N*-neotetraose oligosaccharide extensions but not for sialyl-lactose structures whilst fragmentation of sodiated adducts, $[M+Na]^+$, afforded information diagnostic of mono- and disialylated lactose extensions. The method was applied to the *H. influenzae* genome strain, Rd, in which glycoforms containing both sialyl-lactose and sialyl-lacto-*N*-neotetraose were detected from diagnostic B-ions at *m/z* 638.2 ($[Neu5AcHex2 + Na]^+$) and 657.2 ($[Neu5AcHexHexNAc + H]^+$). Figure 3B shows the MS^2 spectrum (positive mode) of *m/z* 1081 (corresponds to *m/z* 1079.5 in Fig. 3A). Although minor, the diagnostic ion at *m/z* 657.4 confirms the presence of sialyl-lacto-*N*-neotetraose in NTHi strain R2486.

Profiling the substitution pattern in oligosaccharide by ESI-MS

Electrospray-MS (negative or positive mode) on oligosaccharide samples derived from LPS by mild acid hydrolyses reveals the same glycoform populations as observed for LPS-OH, however, without acid labile

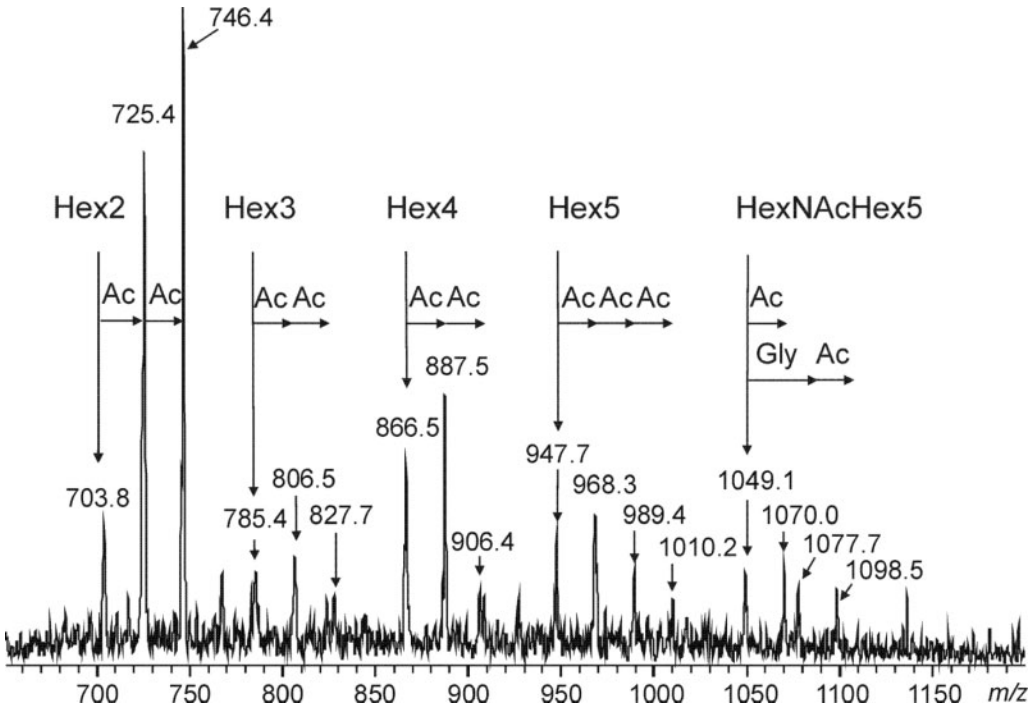


Fig. 4. Electrospray-mass spectrum (negative mode) of oligosaccharide derived from NTHi strain 1200. Indicated are the doubly charged ions of the Hex2 to HexNAcHex5 glycoforms and their acylated counterparts. The experiment was performed on a Finnigan LCQ ion-trap mass spectrometer (Finnigan-MAT, San Jose, CA, USA).

substituents. Instead, the heterogeneity in the oligosaccharide part due to various substituents, in particular acetates and glycine, is observed and their number and identity can be proposed from the MS data. Figure 4 shows the ESI-MS spectrum (negative mode) of oligosaccharide derived from NTHi strain 1200. LPS structures identified for NTHi 1200 are the same as for NTHi strain 1268 except for that strain 1200 is acetylated.[23] Thus, it was found that each glycoform carries up to three acetates in addition to one glycine residue. The position of two acetates could be obtained by MS^3 experiments. In general, information on the location of *P*Cho, *P*Etn, Gly and Ac can be provided by MS/MS experiments in the positive ion mode.[21,28,36] In the corresponding spectra, the occurrence of marker ions corresponding to specific building blocks is of particular interest. Figure 5 shows results obtained for a glycoform (composition *P*Cho·Ac·Gly·Hex_3·Hep_3·*P*Etn·*An*Kdo-ol; $[M+2H]^{2+}$ 837.0) from the allelic variant strain 486lpsAEa (unpublished data). In this strain, the LpsA from NTHi strain 486 is exchanged with that from *H. influenzae* strain Eagan.[42] Thus, only the glycose extension from HepIII differs from the 486 wild-type strain. For strain 486, the position of one acetyl group was assigned to O-2 of HepIII by NMR methods. The position of other acetyl groups and glycine was not established.[14] In Figure 5, ions at *m/z* 292, 328, 370 and 643 correspond to HepAcGly, *P*ChoHex, *P*ChoAcHex and *P*ChoHexHep*P*Etn, respectively. The ion at *m/z* 454 confirms the position of Ac at HepIII but also evidences Gly to be linked to that residue. The ion at *m/z* 370 indicates another position of Ac, at the hexose linked to HepII. The other marker ions confirm the position of phosphorylation sites in NTHi strain 486.

Profiling of oligosaccharide glycoforms in by MS^n

Analysis of permethylated oligosaccharides by MS^n has allowed us to identify isomeric glycoforms in very heterogeneous LPS preparations.[17–19,21–23,28,30,66,72] In general, permethylation increases the MS response by several orders of magnitude and several glycoforms can be observed in the MS spectra that are not detected in underivatized samples. Sequence information is readily obtained from permethylated samples since methyl tagging allows the distinction between fragment ions generated by cleavage of a single glycosidic bond and inner fragments resulting from the rupture of two glycosidic linkages. In contrast, the fragment ions arising from these cleavage pathways are isobaric in the underivatized compounds rendering them indistinguishable from each other. Figure 6 shows the application to oligosaccharide derived from NTHi strain 1200.[23] Strain 1200 contains virtually the same glycoforms as observed in strain 1268; however, traces of three other higher molecular

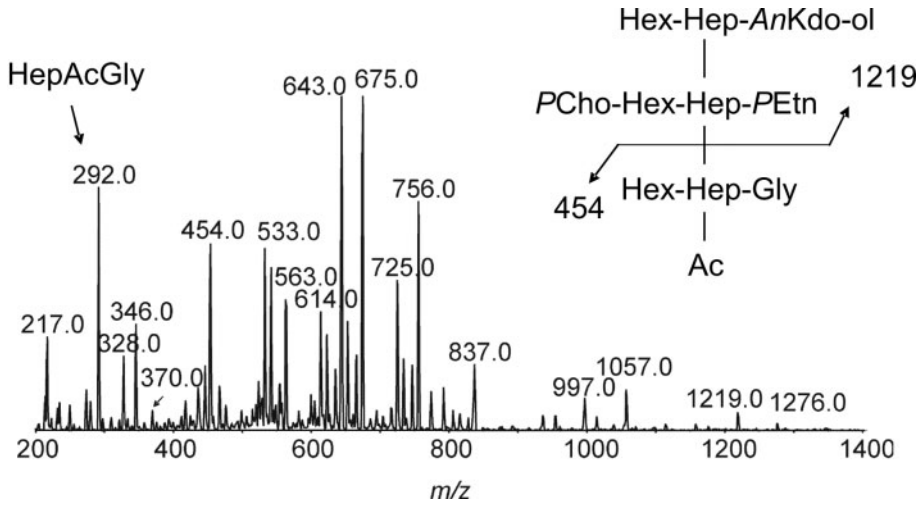


Fig. 5. CE-ESI-MS^2 (positive mode) analyses of the oligosaccharide from NTHi allelic variant strain 486lpsAEa. Shown is the product ion spectrum of $[M+2H]^{2+}$ at *m/z* 837.0 corresponding to *P*Cho·Ac·Gly·Hex_3·Hep_3·*P*Etn·*An*Kdo-ol. The proposed carbohydrate backbone and selected fragments are indicated. Experimental conditions are described elsewhere.[21]

mass forms – $HexNAc_1$·Hex_6·Hep_3·*An*Kdo-ol and $HexNAc_{1-2}$·Hex_7·Hep_3·*An*Kdo-ol – at *m/z* 2528.9, 2732.9 and 2976.9, respectively, are observed (Fig. 6A). The MS^2 spectrum of *m/z* 2976.9 is shown in Figure 6B. The fragment ions *m/z* 2718 and 1860, indicate the loss of t-HexNAc and t-HexNAc-Hex-Hex-Hex-HepIII. The ion at *m/z* 1860 was further fragmented (Fig. 6C) which gave the fragment ions at *m/z* 1600 and 1130 corresponding to

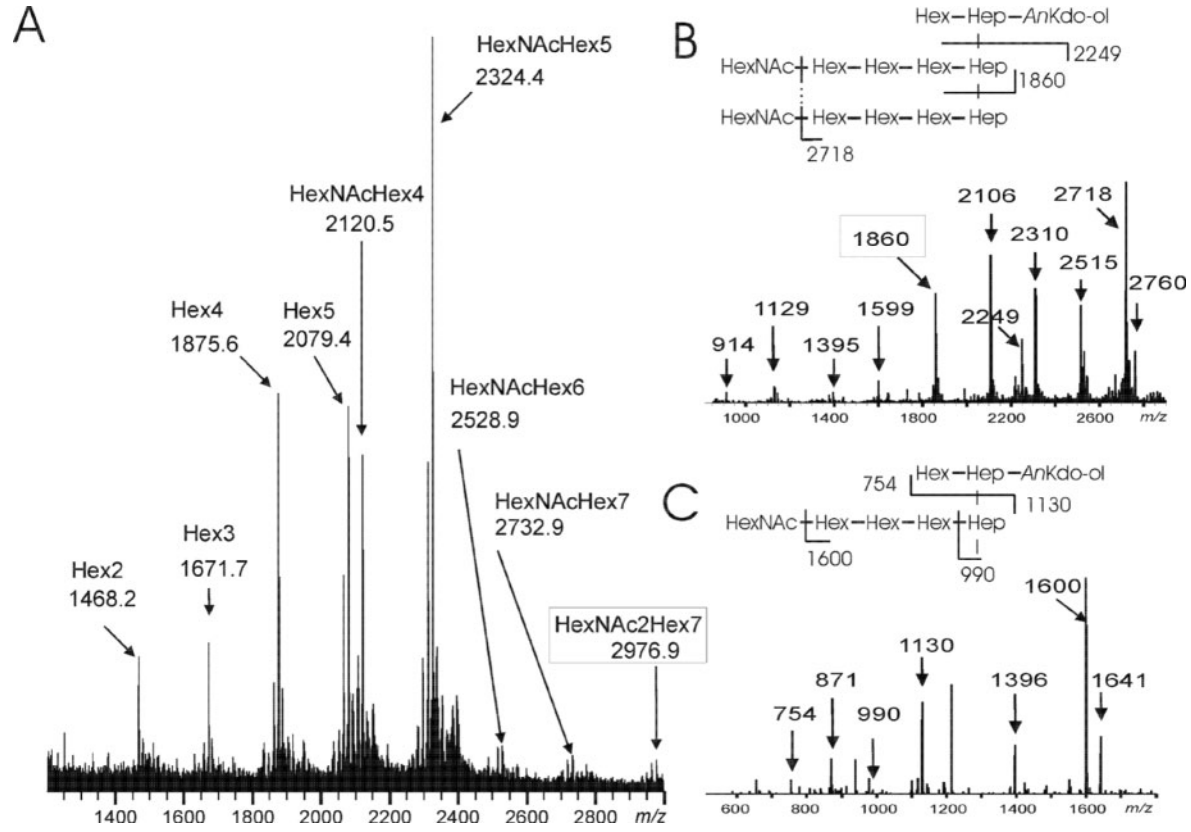


Fig. 6. Electrospray-MS^n analysis of permethylated oligosaccharide of strain 1200. (A) Full scan spectrum (positive mode) on permethylated dephosphorylated strain 1200 oligosaccharide. (B) Product ion spectrum of $[M+Na]^+$ *m/z* 2976.9 corresponding to the HexNAc2Hex7 glycoform. Proposed key fragments are indicated in the structure. (C) MS^3 of the ion at *m/z* 1860 from MS^2 of *m/z* 2976.9. Proposed key fragments are indicated in the structure. The experimental conditions are described elsewhere.[23]

the loss of t-HexNAc and t-Hex-HepI-Kdo. This indicates the Hex7HexNAc2 isomer to be substituted by one hexose at HepI and a HexNAcHexHexHex unit substituting both HepII and HepIII, respectively. By analogy, 13 different isomeric glycoforms were identified in oligosaccharide from NTHi strain 1200.[23] In another strain, NTHi strain 1233, this approach allowed us to detect 36 variant LPS glycoforms in one strain – the most so far.[18]

Profiling of glycoforms in intact LPS

Electrospray-MS analyses of intact unmodified LPS have not been exploited routinely, probably because of the insolubility of LPS in aqueous/organic solvents, although spectra have been reported for *H. influenzae* LPS.[37]

Recently, an electrophoresis-assisted open-tubular LC/MS method was applied to intact LPS from *H. influenzae* strain RM 118 (Rd).[73] In that study, we were able to obtain structural information on both oligosaccharide and the lipid A moiety including the sialylation, glycylation, and the distribution of fatty acid residues on the disaccharide backbone of lipid A. The fragmentation patterns of sodiated and protonated LPS molecules were investigated for determining the location of sialic acid. The MS/MS spectra of sodiated ions provided unambiguous evidence of both sialylated lactose and sialylated lacto-*N*-neotetraose. The structure of lipid A was characterized by MS/MS on intact LPS molecules without any prior chemical modifications. In the same way, information on the oligosaccharide could be obtained by MS/MS by focusing on ions originating from core oligosaccharide.

Conclusions

A characteristic feature of *H. influenzae* LPS is the extensive inter- and intra-strain heterogeneity of glycoform structure. The structural characterization of LPS is important for the understanding its role in disease pathogenesis. Of particular interest are terminal structures known to be responsible for the potential of the bacterium to cause disease. We know that sialic acid, digalactoside and *P*Cho are important virulence factors. Furthermore, acylation and phosphorylation of the core oligosaccharide may contribute to virulence. Identification of these structural features during the course of disease will permit further delineation of their role in pathogenesis. To date, this has only been done for sialic acid in the chinchilla model for experimental otitis media.

The analysis of complex LPS mixtures is further complicated due to the relatively small amounts derived from bacterial sources. This is especially true when samples are obtained as disease isolates from *in vivo* sources. To this end, mass spectrometry methods have provided the means to characterize clinical isolates or phase variants and have been applied to single colony analysis. Moreover, the advances made in MS technology together with the increased knowledge of how LPS molecules behave during MS fragmentation now permits the characterization of LPS glycoform populations in complex mixtures and biological matrices. Future developments in instrumentation and methodology will, no doubt, further enable the study of the contribution of LPS virulence determinants to disease development *in vivo*.

Acknowledgements

The authors thank Dr James C. Richards for kind criticism of the manuscript. EKHS is supported by the Swedish Research Council.

References

1. Owings MF, Kozak LJ. Ambulatory and inpatient procedures in the United States, 1996. *Vital Health Stat* 1998; **13**: 1–119.
2. Wang EE, Einarson TR, Kellner JD *et al.* Antibiotic prescribing for Canadian preschool children: evidence of overprescribing for viral respiratory infections. *Clin Infect Dis* 1999; **29**: 155–160.
3. Hood DW, Moxon ER. Lipopolysaccharide phase variation in *Haemophilus* and *Neisseria*. New York: Marcel Dekker, 1999.
4. Schweda EK, Richards JC, Hood DW *et al.* Expression and structural diversity of the lipopolysaccharide of *Haemophilus influenzae*: implication in virulence. *Int J Med Microbiol* 2007; **297**: 297–306.
5. Preston A, Mandrell RE, Gibson BW *et al.* The lipooligosaccharides of pathogenic Gram-negative bacteria. *Crit Rev Microbiol* 1996; **22**: 139–180.
6. Swords WE, Jones PA, Apicella MA. The lipo-oligosaccharides of *Haemophilus influenzae*: an interesting array of characters. *J Endotoxin Res* 2003; **9**: 131–144.
7. Phillips NJ, Apicella MA, Griffiss JM *et al.* Structural characterization of the cell surface lipooligosaccharides from a nontypable strain of *Haemophilus influenzae*. *Biochemistry* 1992; **31**: 4515–4526.
8. Phillips NJ, Apicella MA, Griffiss JM *et al.* Structural studies of the lipooligosaccharides from *Haemophilus influenzae* type b strain A2. *Biochemistry* 1993; **32**: 2003–2012.
9. Schweda EK, Hegedus OE, Borrelli S *et al.* Structural studies of the saccharide part of the cell envelope lipopolysaccharide from *Haemophilus influenzae* strain AH1-3 (lic3$^+$). *Carbohydr Res* 1993; **246**: 319–330.
10. Schweda EK, Jansson PE, Moxon ER *et al.* Structural studies of the saccharide part of the cell envelope lipooligosaccharide from *Haemophilus influenzae* strain galEgalK. *Carbohydr Res* 1995; **272**: 213–224.
11. Masoud H, Moxon ER, Martin A *et al.* Structure of the variable and conserved lipopolysaccharide oligosaccharide epitopes expressed by *Haemophilus influenzae* serotype b strain Eagan. *Biochemistry* 1997; **36**: 2091–2103.
12. Risberg A, Masoud H, Martin A *et al.* Structural analysis of the lipopolysaccharide oligosaccharide epitopes expressed by a capsule-deficient strain of *Haemophilus influenzae* Rd. *Eur J Biochem* 1999; **261**: 171–180.

13. Schweda EK, Brisson JR, Alvelius G *et al.* Characterization of the phosphocholine-substituted oligosaccharide in lipopolysaccharides of type b *Haemophilus influenzae*. *Eur J Biochem* 2000; **267**: 3902–3913.
14. Månsson M, Bauer SH, Hood DW *et al.* A new structural type for *Haemophilus influenzae* lipopolysaccharide. Structural analysis of the lipopolysaccharide from nontypeable *Haemophilus influenzae* strain 486. *Eur J Biochem* 2001; **268**: 2148–2159.
15. Cox AD, Masoud H, Thibault P *et al.* Structural analysis of the lipopolysaccharide from the nontypable *Haemophilus influenzae* strain SB 33. *Eur J Biochem* 2001; **268**: 5278–5286.
16. Månsson M, Hood DW, Li J *et al.* Structural analysis of the lipopolysaccharide from nontypeable *Haemophilus influenzae* strain 1003. *Eur J Biochem* 2002; **269**: 808–818.
17. Månsson M, Hood DW, Moxon ER *et al.* Structural characterization of a novel branching pattern in the lipopolysaccharide from nontypeable *Haemophilus influenzae*. *Eur J Biochem* 2003; **270**: 2979–2991.
18. Månsson M, Hood DW, Moxon ER *et al.* Structural diversity in lipopolysaccharide expression in nontypeable *Haemophilus influenzae*. Identification of L-*glycero*-D-*manno*-heptose in the outer-core region in three clinical isolates. *Eur J Biochem* 2003; **270**: 610–624.
19. Yildirim HH, Hood DW, Moxon ER *et al.* Structural analysis of lipopolysaccharides from *Haemophilus influenzae* serotype f. Structural diversity observed in three strains. *Eur J Biochem* 2003; **270**: 3153–3167.
20. Masoud H, Martin A, Thibault P *et al.* Structure of extended lipopolysaccharide glycoforms containing two globotriose units in *Haemophilus influenzae* serotype b strain RM7004. *Biochemistry* 2003; **42**: 4463–4475.
21. Yildirim HH, Li J, Richards JC *et al.* Complex *O*-acetylation in non-typeable *Haemophilus influenzae* lipopolysaccharide: evidence for a novel site of *O*-acetylation. *Carbohydr Res* 2005; **340**: 2598–2611.
22. Yildirim HH, Li J, Richards JC *et al.* An alternate pattern for globoside oligosaccharide expression in *Haemophilus influenzae* lipopolysaccharide: structural diversity in nontypeable strain 1124. *Biochemistry* 2005; **44**: 5207–5224.
23. Lundström SL, Twelkmeyer B, Sagemark MK *et al.* Novel globoside-like oligosaccharide expression patterns in nontypeable *Haemophilus influenzae* lipopolysaccharide. *FEBS Lett* 2007; **274**: 4886–4903.
24. Hood DW, Deadman ME, Allen T *et al.* Use of the complete genome sequence information of *Haemophilus influenzae* strain Rd to investigate lipopolysaccharide biosynthesis. *Mol Microbiol* 1996; **22**: 951–965.
25. Hood DW, Cox AD, Wakarchuk WW *et al.* Genetic basis for expression of the major globotetraose-containing lipopolysaccharide from *H. influenzae* strain Rd (RM118). *Glycobiology* 2001; **11**: 957–967.
26. Helander IM, Lindner B, Brade H *et al.* Chemical structure of the lipopolysaccharide of *Haemophilus influenzae* strain I-69 Rd-/b$^+$. Description of a novel deep-rough chemotype. *Eur J Biochem* 1988; **177**: 483–492.
27. Mikhail I, Yildirim HH, Lindahl EC *et al.* Structural characterization of lipid A from nontypeable and type f *Haemophilus influenzae*: variability of fatty acid substitution. *Anal Biochem* 2005; **340**: 303–316.
28. Landerholm MK, Li J, Richards JC *et al.* Characterization of novel structural features in the lipopolysaccharide of nondisease associated nontypeable *Haemophilus influenzae*. *Eur J Biochem* 2004; **271**: 941–953.
29. Schweda EK, Li J, Moxon ER *et al.* Structural analysis of lipopolysaccharide oligosaccharide epitopes expressed by non-typeable *Haemophilus influenzae* strain 176. *Carbohydr Res* 2002; **337**: 409–420.
30. Schweda EK, Landerholm MK, Li J *et al.* Structural profiling of lipopolysaccharide glycoforms expressed by non-typeable *Haemophilus influenzae*: phenotypic similarities between NTHi strain 162 and the genome strain Rd. *Carbohydr Res* 2003; **338**: 2731–2744.
31. Hood DW, Makepeace K, Deadman ME *et al.* Sialic acid in the lipopolysaccharide of *Haemophilus influenzae*: strain distribution, influence on serum resistance and structural characterization. *Mol Microbiol* 1999; **33**: 679–692.
32. Rahman MM, Gu XX, Tsai CM *et al.* The structural heterogeneity of the lipooligosaccharide (LOS) expressed by pathogenic non-typeable *Haemophilus influenzae* strain NTHi 9274. *Glycobiology* 1999; **9**: 1371–1380.
33. Cox AD, Hood DW, Martin A *et al.* Identification and structural characterization of a sialylated lacto-N-neotetraose structure in the lipopolysaccharide of *Haemophilus influenzae*. *Eur J Biochem* 2002; **219**: 4009–4019.
34. Hood DW, Randle G, Cox AD *et al.* Biosynthesis of cryptic lipopolysaccharide glycoforms in *Haemophilus influenzae* involves a mechanism similar to that required for O-antigen synthesis. *J Bacteriol* 2004; **186**: 7429–7439.
35. Fox KL, Li J, Schweda EK *et al*. Duplicate copies of *lic1* direct the addition of multiple phosphocholine residues in the lipopolysaccharide of *Haemophilus influenzae. Infect Immun* 2008; **76**: 588–600.
36. Li J, Bauer SH, Månsson M *et al*. Glycine is a common substituent of the inner core in *Haemophilus influenzae* lipopolysaccharide. *Glycobiology* 2001; **11**: 1009–1015.
37. Risberg A, Alvelius G, Schweda EK. Structural analysis of the lipopolysaccharide oligosaccharide epitopes expressed by *Haemophilus influenzae* strain RM.118-26. *Eur J Biochem* 1999; **265**: 1067–1074.
38. Fleischmann RD, Adams MD, White O *et al.* Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* 1995; **269**: 496–512.
39. Erwin AL, Nelson KL, Mhlanga-Mutangadura T *et al.* Characterization of genetic and phenotypic diversity of invasive nontypeable *Haemophilus influenzae. Infect Immun* 2005; **73**: 5853–5863.
40. Harrison A, Dyer DW, Gillaspy A *et al.* Genomic sequence of an otitis media isolate of nontypeable *Haemophilus influenzae*: comparative study with *H. influenzae* serotype d, strain KW20. *J Bacteriol* 2005; **187**: 4627–4636.
41. Hood DW, Dead2; **269**: 4009–4man ME, Cox AD *et al.* Three genes, *lgtF*, *lic2C* and *lpsA*, have a primary role in determining the pattern of oligosaccharide extension from the inner core of *Haemophilus influenzae* LPS. *Microbiology* 2004; **150**: 2089–2097.
42. Deadman ME, Lundstrom SL, Schweda EK *et al.* Specific amino acids of the glycosyltransferase LpsA direct the addition of glucose or galactose to the terminal inner core heptose of *Haemophilus influenzae* lipopolysaccharide via alternative linkages. *J Biol Chem* 2006; **281**: 29455–29467.
43. Weiser JN, Pan N. Adaptation of *Haemophilus influenzae* to acquired and innate humoral immunity based on phase variation of lipopolysaccharide. *Mol Microbiol* 1998; **30**: 767–775.
44. Weiser JN, Maskell DJ, Butler PD *et al.* Characterization of repetitive sequences controlling phase variation of *Haemophilus influenzae* lipopolysaccharide. *J Bacteriol* 1990; **172**: 3304–3309.
45. Jarosik GP, Hansen EJ. Identification of a new locus involved in expression of *Haemophilus influenzae* type b lipooligosaccharide. *Infect Immun* 1994; **62**: 4861–4867.
46. Hood DW, Deadman ME, Jennings MP *et al.* DNA repeats

identify novel virulence genes in *Haemophilus influenzae*. *Proc Natl Acad Sci USA* 1996; **93**: 11121–11125.
47. Griffin R, Cox AD, Makepeace K *et al.* The role of *lex2* in lipopolysaccharide biosynthesis in *Haemophilus influenzae* strains RM7004 and RM153. *Microbiology* 2003; **149**: 3165–3175.
48. Fox KL, Yildirim HH, Deadman ME *et al.* Novel lipopolysaccharide biosynthetic genes containing tetranucleotide repeats in *Haemophilus influenzae*, identification of a gene for adding O-acetyl groups. *Mol Microbiol* 2005; **58**: 207–216.
49. Weiser JN, Shchepetov M, Chong ST. Decoration of lipopolysaccharide with phosphorylcholine: a phase-variable characteristic of *Haemophilus influenzae*. *Infect Immun* 1997; **65**: 943–950.
50. Lysenko E, Richards JC, Cox AD *et al.* The position of phosphorylcholine on the lipopolysaccharide of *Haemophilus influenzae* affects binding and sensitivity to C-reactive protein-mediated killing. *Mol Microbiol* 2000; **35**: 234–245.
51. Hood DW, Cox AD, Gilbert M *et al.* Identification of a lipopolysaccharide alpha-2,3-sialyltransferase from *Haemophilus influenzae*. *Mol Microbiol* 2001; **39**: 341–350.
52. Wright JC, Hood DW, Randle GA *et al. lpt6*, a gene required for addition of phosphoethanolamine to inner-core lipopolysaccharide of *Neisseria meningitidis* and *Haemophilus influenzae*. *J Bacteriol* 2004; **186**: 6970–6982.
53. Tullius MV, Phillips NJ, Scheffler NK *et al.* The *lbgAB* gene cluster of *Haemophilus ducreyi* encodes a beta-1,4-galactosyltransferase and an alpha-1,6-DD-heptosyltransferase involved in lipooligosaccharide biosynthesis. *Infect Immun* 2002; **70**: 2853–2861.
54. Post DM, Munson Jr RS, Baker B *et al.* Identification of genes involved in the expression of atypical lipooligosaccharide structures from a second class of *Haemophilus ducreyi*. *Infect Immun* 2007; **75**: 113–121.
55. Bauer SH, Månsson M, Hood DW *et al.* A rapid and sensitive procedure for determination of 5-*N*-acetyl neuraminic acid in lipopolysaccharides of *Haemophilus influenzae*: a survey of 24 non-typeable *H. influenzae* strains. *Carbohydr Res* 2001; **335**: 251–260.
56. Swords WE, Moore ML, Godzicki L *et al.* Sialylation of lipooligosaccharides promotes biofilm formation by nontypeable *Haemophilus influenzae*. *Infect Immun* 2004; **72**: 106–113.
57. Weiser JN, Pan N, McGowan KL *et al.* Phosphorylcholine on the lipopolysaccharide of *Haemophilus influenzae* contributes to persistence in the respiratory tract and sensitivity to serum killing mediated by C-reactive protein. *J Exp Med* 1998; **187**: 631–640.
58. Swords WE, Buscher BA, Ver Steeg Ii K *et al.* Non-typeable *Haemophilus influenzae* adhere to and invade human bronchial epithelial cells via an interaction of lipooligosaccharide with the PAF receptor. *Mol Microbiol* 2000; **37**: 13–27.
59. Lysenko ES, Gould J, Bals R *et al.* Bacterial phosphorylcholine decreases susceptibility to the antimicrobial peptide LL-37/hCAP18 expressed in the upper respiratory tract. *Infect Immun* 2000; **68**: 1664–1671.
60. Griffin R, Bayliss CD, Herbert MA *et al.* Digalactoside expression in the lipopolysaccharide of *Haemophilus influenzae* and its role in intravascular survival. *Infect Immun* 2005; **73**: 7022–7026.
61. Holst O, Brade L, Kosma P *et al.* Structure, serological specificity, and synthesis of artificial glycoconjugates representing the genus-specific lipopolysaccharide epitope of *Chlamydia* spp. *J Bacteriol* 1991; **173**: 1862–1866.
62. Schweda EK, Richards JC. Structural profiling of short-chain lipopolysaccharides from *Haemophilus influenzae*. *Methods Mol Med* 2003; **71**: 161–183.
63. Schilling B, McLendon MK, Phillips NJ *et al.* Characterization of lipid A acylation patterns in *Francisella tularensis*, *Francisella novicida*, and *Francisella philomiragia* using multiple-stage mass spectrometry and matrix-assisted laser desorption/ionization on an intermediate vacuum source linear ion trap. *Anal Chem* 2007; **79**: 1034–1042.
64. Vinogradov E, Lindner B, Seltmann G *et al.* Lipopolysaccharides from *Serratia marcescens* possess one or two 4-amino-4-deoxy-L-arabinopyranose 1-phosphate residues in the lipid A and D-glycero-D-talo-oct-2-ulopyranosonic acid in the inner core region. *Chemistry* 2006; **12**: 6692–6700.
65. Melaugh WJJE, Autiola S, Phillips NJ, Gibson BW. *Mass Spectrometry in the Biological Sciences*. Totowa, NS: Humana, 1996.
66. Gulin S, Pupo E, Schweda EK *et al.* Linking mass spectrometry and slab-polyacrylamide gel electrophoresis by passive elution of lipopolysaccharides from reverse-stained gels: analysis of gel-purified lipopolysaccharides from *Haemophilus influenzae* strain Rd. *Anal Chem* 2003; **75**: 4918–4924.
67. Li J, Richards JC. Application of capillary electrophoresis mass spectrometry to the characterization of bacterial lipopolysaccharides. *Mass Spectrom Rev* 2007; **26**: 35–50.
68. Dzieciatkowska M, Brochu D, van Belkum A *et al.* Mass spectrometric analysis of intact lipooligosaccharide: direct evidence for O-acetylated sialic acids and discovery of O-linked glycine expressed by *Campylobacter jejuni*. *Biochemistry* 2007; **46**: 14704–14714.
69. Thibault P, Martin A, Gilbert M *et al.* Analysis of bacterial glycolipids by capillary electrophoresis-electrospray mass spectrometry: *Haemophilus influenzae* and *Neisseria meningitidis* lipopolysaccharides. *Methods Mol Biol* 2003; **213**: 241–259.
70. Thibault P, Richards JC. Applications of combined capillary electrophoresis-electrospray mass spectrometry in the characterization of short-chain lipopolysaccharides. *Haemophilus influenzae*. *Methods Mol Biol* 2000; **145**: 327–344.
71. Li J, Dzieciatkowska M, Hood DW *et al.* Structural characterization of sialylated glycoforms of *H. influenzae* by electrospray mass spectrometry: fragmentation of protonated and sodiated O-deacylated lipopolysaccharides. *Rapid Commun Mass Spectrom* 2007; **21**: 952–960.
72. Reinhold VN, Reinhold BB, Costello CE. Carbohydrate molecular weight profiling, sequence, linkage, and branching data: ES-MS and CID. *Anal Chem* 1995; **67**: 1772–1784.
73. Dzieciatkowska M, Schweda EK, Moxon ER *et al.* Characterization of intact lipopolysaccharides from the *Haemophilus influenzae* strain RM 118 using electrophoresis-assisted open-tubular liquid chromatography-mass spectrometry. *Electrophoresis* 2008; **29**: 2171–2181.
74. Peters T, Meyer B, Stuike-Prill R *et al.* A Monte Carlo method for conformational analysis of saccharides. *Carbohydr Res* 1993; **238**: 49–73.