From Department of Laboratory Medicine
Clinical Research Center

Karolinska Institutet, Stockholm, Sweden

Structural studies of lipopolysaccharides expressed by non-typeable *Haemophilus influenzae* and *Haemophilus parainfluenzae* strains

Varvara Vitiazeva

Karolinska Institutet

Stockholm 2012

ABSTRACT

The present thesis describes lipopolysaccharide (LPS) structures expressed by non-typeable *Haemophilus influenzae* and *Haemophilus parainfluenzae* strains. LPS is a major surface component of Gram-negative bacteria. Structural studies of LPS are very important for understanding the adaptive mechanisms which help bacteria to survive in the host environment.

Non-typeable *Haemophilus influenzae* (NT*Hi*) is a common human commensal of the nasopharynx. It is also pathogenic and causes both acute and chronic diseases, such as otitis media, sinusitis, pneumonia and bronchitis. *H. influenzae* expresses rough type LPS (lacking O-antigen), which is implicated as a major virulence factor. 25 NT*Hi* otitis media isolates were selected for structural studies of LPS. These clinical isolates represent the structural diversity of LPS in the natural population.

Structural studies of *H. influenzae* LPS have resulted in a molecular model consisting of a conserved (*P*Etn)-substituted triheptosyl inner-core moiety (HepI–HepII-HepIII) in which each of the heptose residues can provide a point for elongation by oligosaccharide chains (outer-core region).

NT*Hi* strains 1158/1159 and 1232, described in this thesis, were selected from this collection of clinical isolates. These strains express additional D,D-Hep residue in the outer-core region of LPS.

Haemophilus parainfluenzae is a part of normal human flora. Previous studies have indicated that *H. parainfluenzae* expresses LPS structures that are very similar to those expressed by *H. influenzae*. On the other hand some *H. parainfluenzae* strains express O-antigen containing LPS. The structures of the O-antigen from *H. parainfluenzae* strains 20 and 16 are described in this thesis.

The structural investigations of LPS of *H. influenzae* and the comparison with LPS expressed by *H. parainfluenzae* will increase the knowledge of biological properties of LPS and its role in bacterial virulence.

LIST OF PUBLICATIONS

Paper I. **The structural diversity of lipopolysaccharide expressed by non-typeable *Haemophilus influenzae* strains 1158 and 1159.**
Vitiazeva, V., Li, J., Hood, D. W., Richard Moxon, E., and Schweda, E. K. (2012) *Carbohydrate research 357*, 98-110.

Paper II. **The structural studies of a novel branching pattern in the lipopolysaccharide expressed by non-typeable *Haemophilus influenzae* strain 1232.**
Vitiazeva, V., Månsson, M., Li, J., Hood, D. W., Richard Moxon, E., and Schweda, E. K. *Manuscript.*

Paper III. **Structural studies of the lipopolysaccharide from *Haemophilus parainfluenzae* strain 20.**
Vitiazeva, V., Twelkmeyer, B., Young, R., Hood, D. W., and Schweda, E. K. (2011) *Carbohydrate research 346*, 2228-2236.

Paper IV. **Structural studies of the O-antigen from *Haemophilus parainfluenzae* strain 16.**
Vitiazeva, V., Young, R., Hood, D. W., and Schweda, E. K. *Manuscript.*

Related publications not included in thesis

Genes required for the synthesis of heptose-containing oligosaccharide outer core extensions in *Haemophilus influenzae* lipopolysaccharide.
Hood, D. W., Deadman, M. E., Engskog, M. K., **Vitiazeva, V**., Makepeace, K., Schweda, E. K., and Moxon, R. (2010) *Microbiology 156*, 3421-3431

Duplicate copies of lic1 direct the addition of multiple phosphocholine residues in the lipopolysaccharide of *Haemophilus influenza*.
Fox, K. L., Li, J., Schweda, E. K., **Vitiazeva, V**., Makepeace, K., Jennings, M. P., Moxon, E. R., and Hood, D. W. (2008) *Infection and immunity 76*, 588-600

A *Haemophilus influenzae* strain associated with Fisher syndrome expresses a novel disialylated ganglioside mimic.
Houliston, R. S., Koga, M., Li, J., Jarrell, H. C., Richards, J. C., **Vitiazeva, V**., Schweda, E. K., Yuki, N., and Gilbert, M. (2007) *Biochemistry 46*, 8164-8171.

CONTENTS

CONTENTS

LIST OF ABBREVIATIONS

1D	One-dimensional
2D	Two-dimensional
Ac	Acetate
CE	Capillary electrophoresis
COSY	Correlation spectroscopy
DMSO	Dimethyl sulfoxide
ESI	Electrospray ionization
Gal	Galactose
GalNAc	2-acetamido-2-deoxygalactose
GC	Gas chromatography
Glc	Glucose
GlcNAc	2-acetamido-2-deoxyglucose
Gly	Glycine
Hep	Heptose
D,D-Hep	D-*glycero*-D-*manno*-heptose
L,D-Hep	L-*glycero*-D-*manno*-heptose
Hex	Hexose
HexNAc	N-acetylhexoseamine
Hi	*Haemophilus influenzae*
HMBC	Heteronuclear Multiple-Bond Correlation
HMQC	Heteronuclear Multiple Quantum Coherence
HPLC	High performance liquid chromatography
Kdo	3-deoxy-D-*manno*-oct-2-ulsonic-acid
*An*Kdo-*ol*	Reduced anhydro Kdo
LPS	Lipopolysaccharide
LPS-OH	O-deacylated LPS

LIST OF ABBREVIATIONS

MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MS^n	Multiple step tandem mass spectrometry
Neu5Ac	N-acetylneuraminic acid
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser Enhancement
NOESY	NOE spectroscopy
NT*Hi*	Non-typeable *Haemophilus influenzae*
OS	Oligosaccharide
PAD	Pulsed amperometric detection
P	Phosphate
*P*Cho	Phosphocholine
PCP	Phenol: chloroform: light petroleum
*P*Etn	Phosphoethanolamine
PMAA	Partially methylated alditol acetate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SIM	Selective ion monitoring
TIC	Total ion chromatogram
TFA	Trifluoroacetic acid
TOCSY	Total Correlation Spectroscopy

1 INTRODUCTION

This thesis presents results from structural studies of lipopolysaccharides from non-typeable *Haemophilus influenzae* (NT*Hi*) and *Haemophilus parainfluenzae*. Both *H. influenzae* and *H. parainfluenzae* are Gram-negative bacteria that colonize the upper respiratory tract of humans. *H. parainfluenzae* is a part of normal flora, but in rare instances it causes infections such septicaemia, endocarditis, pneumonia, periodontal disease and biliary tract infections (*1-4*). Non-typeable *H. influenzae* is an important cause of respiratory tract infections in children and adults. NT*Hi* is the second most common cause of acute otitis media in children after *Streptococcus pneumoniae* and is responsible for up to 35 % of all cases (*5*). It is a common cause of sinusitis, pneumonia and bronchitis (*6, 7*). It is now well established that the cell surface components of bacteria play extremely important roles in colonization and persistence to the host environment.

1.1. The bacterial cell envelope.

Bacteria can be classified into two groups: Gram-positive and Gram-negative on the basis of a method called Gram staining, developed by Christian Gram in 1884. The Gram staining uses structural differences in bacterial cell surfaces. The bacterial cell envelope has a very complex structure (Fig. 1). It does not only control selective passage of nutrients from outside and waste products from inside, but also serves as a protection of bacteria from very hostile environments (*8-11*). Both Gram-negative and Gram-positive, can be surrounded by a capsule, composed of a polysaccharide, or by S-layer, composed of a single protein (*8*).
Polysaccharides found on the bacterial surfaces are involved in different processes such as cell-cell recognition, differentiation and antigenic expression (*12, 13*).

1.2. The Gram-negative cell envelope.

The cell envelope of Gram-negative bacteria consists of a cytoplasmic membrane covered by a peptidoglycan layer and an outer membrane (Fig. 1A) (*14*). The peptidoglycan layer can be called the skeleton of bacteria. It is composed of a disaccharide repeating unit:

$$\rightarrow 4)\text{-}\beta\text{-D-GlcNAc-}(1\rightarrow 4)\text{-}\beta\text{-D-MurNAc-}(1\rightarrow.$$
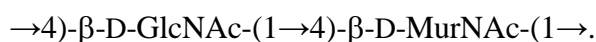
The polysaccharide is cross-linked by peptide chains. Due to the presence of peptidoglycan bacteria do not lyse even in distilled water.
The cytoplasmic membrane is a phospholipid bilayer. The inner leaflet of the outer membrane is also composed of phospholipids. However the outer leaflet is formed by glycolipids (lipopolysaccharides). Thus lipopolysaccharide (LPS) is a major surface component of almost all Gram-negative bacteria.

The LPS layer is very important for viability of bacteria in the host environment and is also responsible for inflammation and toxic symptoms (*15-17*).
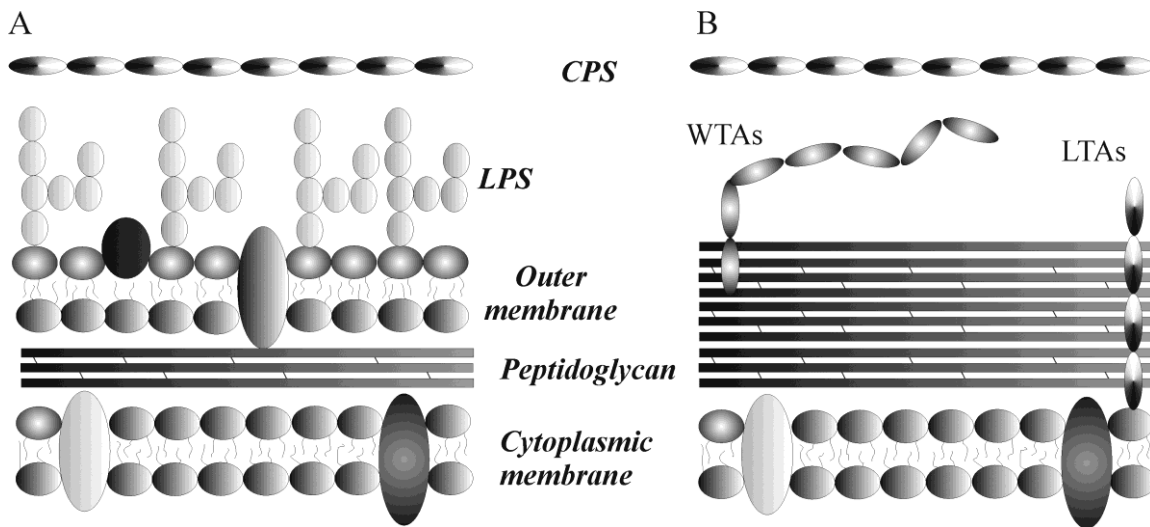
Fig.1 The cell envelopes of Gram-negative (A) and Gram-positive (B) bacteria.
*[Abbreviations: **CPS**-capsular polysaccharide; **LPS**-lipopolysaccharide; **WTAs**-wall teichoic acid*
***LTAs**- lipoteichoic acid*]

1.2.1.   The structure of lipopolysaccharide.

LPS are heat-stable amphiphilic molecules, composed of two regions: a lipophilic region (lipid A) and a hydrophilic region (poly- or oligosaccharide part). The carbohydrate region can be divided into a terminal O-specific chain (O-antigen) and a core region, which is covalently linked to the lipid A (Fig. 2). The O-antigen usually consists of up to 50 repeating oligosaccharide units, which in turn are formed of 2-8 monosaccharides (*17, 18*).
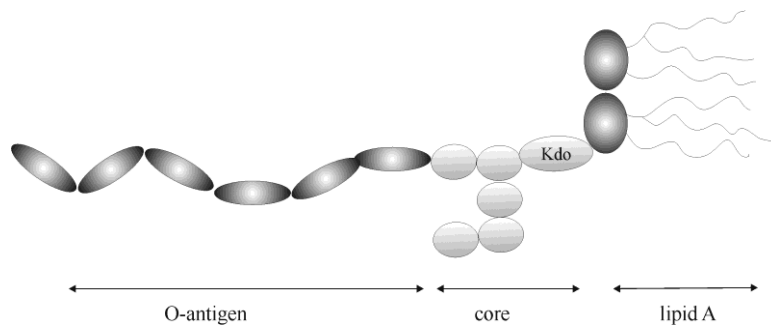
Fig. 2. Schematic representation of lipopolysaccharide.

On the other hand, many pathogenic Gram-negative bacteria such as *N. meningitides*, *N. gonorrhoeae*, *H. influenzae*, *B. pertussis* and *C. trachomatis,* which occupy mucosal surfaces of the respiratory and urogenital tracts, lack the O-antigen in LPS structures (*17, 19*). Such LPS is sometimes referred to as lipooligosaccharide (LOS). Depending on the presence of the O-antigen, Gram-negative bacteria are divided into smooth (S)- and rough (R)-forms.

O-Antigen.

The O-specific polysaccharide is characterized by a very high variation even within the same species. The synthesis of the O-antigen is controlled by genes of the *rfb* locus. The O-specific polysaccharide is synthesized and added *en bloc*. Mutant strains that have any defect in the *rfb* locus synthesize LPS lacking the O-antigen. These mutants grow and multiply *in vitro* studies. However such mutants of pathogenic *Salmonellae*, for example, cannot persist and survive in tissues or body fluids (*15*).

Core region.

The core region of LPS can generally be subdivided into an inner- and an outer-core region. The inner-core region is usually composed of heptoses and 2-keto-3-deoxyoctulosonic acid(s) (Kdo). The Kdo residue is linked to the lipid A via ketosidic bond, which is very sensitive to mild acidic conditions. The inner-core region together with the lipid A moiety correspond to the most conserved part of LPS. The inner core is very often decorated by noncarbohydrate substituents such as free phosphate groups (*P*), phosphoethanolamine (*P*Etn), pyrophosphoethanolamine (*PP*Etn), phosphocholine (*P*Cho), acetate (Ac) and glycine (Gly).

The outer-core possesses more structural diversity, but is still more conserved in structure than the O-antigen.

Lipid A.

The lipid A is the biologically active part of the LPS molecule, which is recognized by host innate immunity (*20*). Depending on the amount of released LPS along with other different factors such as individual sensitivity of the mammalian organism, LPS can either stimulate resistance of the immune system against infection or lead to septic shock (*15, 21, 22*).

Structural studies of the lipid A from different bacteria resulted in the structure containing β-(1-6)- linked D-glucosamine disaccharide (D-Glc*p*N or D-Glc*p*N3N), which carry 3-hydroxy fatty acids at 2, 2´and 3, 3´positions. The 3-hydroxyl group of these fatty acids can be further acylated. As it is shown in Fig. 3, position 1 and position 4´ can be substituted by phosphate groups.
Differences in the structures of the lipid A depend on the nature of sugar residues, phosphorylation pattern, as well as the nature, length and number of fatty acids (*17*).
The number and length of fatty acids affect the toxicity of the lipid A. The most toxic lipid A contains six fatty acids such as: C_{12}, $C_{12}OH$, C_{14} and $C_{14}OH$ (*15, 18*).

1.3. The Gram-positive cell envelope.

Gram-positive bacteria quite often colonize in the same environment as Gram-negative cells. However Gram-positive bacterial cells lack the outer membrane (Fig. 1B). To protect Gram-positive cells from turgor pressure exerted on the plasma membrane, these bacteria are surrounded by a peptidoglycan layer that is much thicker than in Gram-negative bacteria (8). The peptidoglycan layer is penetrated by long carbohydrate polymers called teichoic and lipoteichoic acids. These polymers are composed of repeating carbohydrate units that are linked together by a glycerol phosphate or ribitol phosphate via a phosphodiester linkage. The lipoteichoic acid is a glycolipid and is anchored in the outer layer of the cytoplasmic membrane. The teichoic acid is covalently attached by the phosphodiester bond to the peptidoglycan.

1.3.1. Teichoic acid and lipoteichoic acids of Streptococcus pneumoniae.

Streptococcus pneumoniae express teichoic and lipoteichoic acids having identical chain structures: →6)-β-D-Glc*p*-(1→3)-α-D-Fuc*p*NAc4N-(1→ 4)-α-D-Gal*p*NAc-(1→ 4)-β-D-Gal*p*NAc -(1→ 1)-D-ribitol-5-*P*-(O→, in which D-Fuc*p*NAc4N is 2-acetamido-4-amino-2,4,6-trideoxy-D-galactose (23). This structure is substituted by one or two phosphocholine (*P*Cho) residues (24). The large numbers of pneumococcal proteins need to bind to the *P*Cho residues for their activation. The surface-exposed *P*Cho residues play a very important role in pneumococcal infection (25). However *P*Cho also interacts with C-reactive protein (CRP), an acute-phase protein of mammalian blood serum (26).

1.4. Haemophilus influenzae.

Depending on the presence of the capsular polysaccharide, *H. influenzae* can be subdivided into encapsulated (type a-f) and non- encapsulated (non-typeable) forms. Encapsulated type b strains cause invasive bacteraemic diseases, such as meningitis, epiglottitis, cellulitis and pneumonia. Introduction of vaccines against serotype b *H. influenzae* have dramatically reduced the incidence of diseases caused by this type of *H. influenzae (27)*.

25 NT*Hi* isolates obtained from Finnish children with otitis media have been chosen for structural investigations of LPS. These isolates span a *H. influenzae* species-level ribotyping dendrogram comprised of more than 400 non-typeable and encapsulated strains (28) and represent the diversity of LPS in the natural population of NT*Hi*.
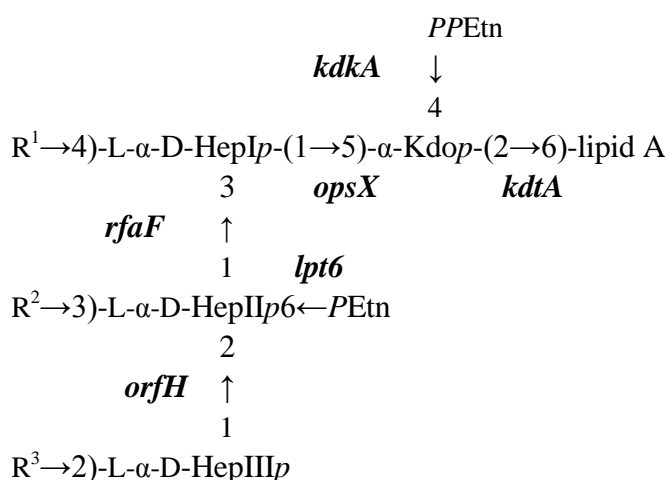
In contrast to encapsulated strains, which have relatively clonal populations, the non-typeable *H. influenzae* (NT*Hi*) show extensive genetic diversity (29-31).

1.4.1.   Lipopolysaccharide structure of *H. influenzae*.

H. influenzae express rough (R)-type LPS which is composed of the lipid A and the core. The core OS region of LPS plays an important role in infections caused by NT*Hi* *(32)*.

The inner-core region.

The inner-core is the most conserved part of NT*Hi* LPS, which is composed of a *P*Etn-substituted triheptosyl moiety linked via one phosphorylated Kdo to the lipid A moiety (Scheme. 1) *(33)*. Structural variations of the inner-core region depend on the presence of non-carbohydrate substituents such as acetate group (Ac), phosphoethanolamine (*P*Etn) and glycine (Gly) *(34-36)*.

$$PP\text{Etn}$$
$$\textbf{\textit{kdkA}} \quad \downarrow$$
$$4$$
$$R^1{\rightarrow}4)\text{-L-}\alpha\text{-D-HepI}p\text{-}(1{\rightarrow}5)\text{-}\alpha\text{-Kdo}p\text{-}(2{\rightarrow}6)\text{-lipid A}$$
$$3 \qquad \textbf{\textit{opsX}} \qquad\qquad \textbf{\textit{kdtA}}$$
$$\textbf{\textit{rfaF}} \qquad \uparrow$$
$$1 \qquad \textbf{\textit{lpt6}}$$
$$R^2{\rightarrow}3)\text{-L-}\alpha\text{-D-HepII}p6{\leftarrow}P\text{Etn}$$
$$2$$
$$\textbf{\textit{orfH}} \quad \uparrow$$
$$1$$
$$R^3{\rightarrow}2)\text{-L-}\alpha\text{-D-HepIII}p$$

Scheme 1. *The conserved inner-core part of LPS from H. influenzae and genes involved in the biosynthesis of inner-core*.
(R^1, R^2, R^3 - H or sugar residues, representing the outer-core region.)

The outer-core region.

The outer-core region of *H. influenzae* LPS is extremely diverse. Each heptose from the triheptosyl inner-core can be an attachment point for an oligosaccharide chain. The complete genome sequence of *H. influenzae* strain Rd has facilitated the study of lipopolysaccharide genes *(37)*. Following the completion of further genome sequences for NT*Hi* strains *(38, 39)*, all of the major genes responsible for synthesis of oligosaccharide part of LPS were identified by sequence similarity comparisons along with structural studies of LPS from wild-type and mutant strains (Schemes 1, 2B) *(40-45)*.

The elongations from the triheptosyl moiety can differ between strains (inter-strain variation) as well as within a single strain (intra-strain variation). The heterogeneity within the same strain appears as differences in lengths of oligosaccharide extensions from the triheptosyl moiety. This intra-strain variation depends on uncompleted syntheses of LPS molecules, on enzyme competition and sterical hinderance as well as genetic mechanism called phase variation *(46)*. Phase variation is also found in other mucosal pathogens such as *Neisseria* *(47, 48)*.

Phase variation is a high frequency on-off switching of gene expression. Several chromosomal *loci* from *H. influenzae* were found to contain a number of tetranucleatide repeats within the open reading frame. Spontaneous variation of these repeats during replication leads to the gain and loss of phase variable structures. (*49-51*).

Table 1. Terminal extensions from LPS of *H. influenzae* that mimic human structures.

Host-mimicking structure	Trivial name
PCho→	Phosphocholine
β-Gal-(1→4)-β-Glc-(1→	Lactose
α-Neu5Ac-(2→3)-β-Gal-(1→4)-β-Glc-(1→	Sialyllactose
α-Gal-(1→4)-β-Gal-(1→	Digalactoside
β-GalNAc-(1→3)-α-Gal-(1→4)-β-Gal-(1→4)-β-Glc-(1→	Globotetraose
α-Gal-(1→4)-β-Gal-(1→4)-β-Glc-(1→	Globotriose
α-Neu5Ac-(2→3)-β-Gal-(1→4)-β-GlcNAc-(1→3)-β-Gal-(1→	Sialyllacto-*N*-neotetraose

NT*Hi* is a highly adapted human pathogen (*52, 53*). Extensive structural studies of the lipopolysaccharides indicated that *H. influenzae* express structures that are immunochemically identical to groups of different human glycosphingolipids and glycolipid antigens (Table 1). Remarkably, expression of almost all host-mimicking structures is controlled by phase variable genes (Scheme 2B).

Six genetic loci, *lic1, lic2, lic3, lgtC, lex2* and *oafA,* have been identified to be responsible for phase variation.

The expression of the digalactoside structure (α-Gal-(1→4)-β-Gal-(1→)) is phase variable from every possible position (*41, 54*). *Lex2* or *lic2A* genes are responsible for addition of the β-Gal residue to Glc. The further addition of α-Gal to β-Gal is controlled by a phase variable gene *lgtC (41)*. Notably, *lex2 (55, 56)* is not only phase variable but also exists in two allelic variants. The presence of the phase-variable and host-mimicking digalactoside epitope on LPS increases resistance to naturally acquired bactericidal antibody in humans (*57*). Most likely the host mimicking allows the bacterium to evade the immune defense system by covering its surface with structures that do not cause the production of host antibody (*58*).

Lic3A and *lic3B* are both phase variable genes. It has been found that *Lic3A* is responsible for sialylation of β-D-Gal linked to β-Glc-HepIII (Scheme 2) (*59*). *Lic3B* is responsible for synthesis of disialyllactose. Interestingly Lic3B is bifunctional and can be involved in syntheses of both sialyl- and disialyllactose (Scheme 2) (*60*). Furthermore the *lic3B* gene is responsible for sialylation of β-D-Gal linked to the external heptose (*61*). Sialic acid can also be linked to β-D-Gal that is either linked to β-Glc-HepI (*62*) or α-Glc-HepII (*63*). Since all tested NT*Hi* strains contain *lic3A* genes, it can be suggested that *lic3A* and/or *lic3B* genes are responsible for addition of sialic acid to these epitopes.

A

β-Glc (β-Gal)

↓1,4

L(D),D-Hep   1,4   R

1,4 1,6↓

1,4   β-Glc→β-Glc →L,D-HepI-(1→

1,6(7*)   1,4   3

1,4   β-Glc   ↑

1,4   ↑ 1,4   1   6,1

1,3   1,4   α-Glc →L,D-HepII←PEtn

β-GalNAc →α-Gal →β-Gal   1,4   2

1,4   ↑

α-Neu5Ac→α-Neu5Ac   1,4   1,2(3)   1

β-Glc →L,D-HepIII

1,2(3)

β-Gal

R: α-Neu5Ac-(2→3)-β-Gal-(1→4)-β-GlcNAc-(1→3)-β-Gal-(1→

PEtn→6)-α-GalNAc-(1→6)-β-Gal-(1→4)-β-GlcNAc-(1→3)-β-Gal-(1→

(7*) – Substituted position of D,D-Hep is identified in this study.

B

β-Glc (β-Gal)

↓*losA*

L(D),D-Hep ←   PCho

lex2* ↓*losB*

?   **β-Glc→β-Glc →L,D-HepI-(1→

lex2 *   *lgtF*   3

**β-Glc   PCho ↑

lgtD   *lgtC* *   *lic2B* ↓   1   *lpt6*

β-GalNAc→α-Gal→β-Gal   ** α-Glc →L,D-HepII←PEtn

lic3A*   *lic2C*   2

α-Neu5Ac→α-Neu5Ac *lic3B**   *lic2A **   ↑

lic3B*   *lpsA*   1

β-Glc →L,D-HepIII

lpsA

β-Gal

lic1* ↑

PCho

Scheme 2. Schematic representation of oligosaccharide elongations from

HepI-HepII-HepIII (A) and genes involved in biosynthesis of the outer core (B).

All sugars are D-pyranosides.

*-phase variable genes;

substitution of **Glc by β-Gal is controlled by *lic2A* gene

7

The sialyllacto-*N*-neotetraose epitope was found in some strains linked to Glc attached to HepI (*64*). Interestingly, biosynthesis of this unit is different from the biosynthesis of the rest of the LPS molecule. It is synthesized and added *en bloc*, by a mechanism related to synthesis of the O-antigen (*65*). Two different sialyltransferases LsgB and SiaA are involved in sialylation of this epitope (*66*).

Almost all clinical isolates studied to date can incorporate sialic acid in their LPS (*67*). The presence of sialylated glycoforms significantly increases the resistance to the killing effects of normal human serum. Sialylated glycoforms are an essential requirement for inflammation of the middle ear in chinchillas (*68-70*).

The *lic 1* locus is associated with incorporation of phosphocholine (*P*Cho) to LPS and comprises 4 genes (*lic1A-lic1D*) (*71*). Comprehensive structural studies of LPS from NT*Hi* strains indicated four possible positions for phosphocholine. The external heptose and the first hexose residue, which is linked to any heptose (HepI, HepII or HepIII) could be substituted by *P*Cho (Scheme 2B). The position of *P*Cho depends on the sequence of *lic1D*, encoding a diphosphonucleoside choline transferase (*72*). Some *H. influenzae* strains express lipopolysaccharide structures containing two *P*Cho residues. Genome sequence analysis of these strains indicated that they contain two distinct copies of the *lic1* operon where the *lic1D* gene from each operon is responsible for position of *P*Cho (*73*).

*P*Cho plays a very important role in colonization of the bacterium on the mucosal surface of the nasopharynx (*57, 74-76*). In addition expression of *P*Cho on LPS of *H. influenzae* has also been associated with increased resistance to host antimicrobial peptide killing (*77*). On the other hand *P*Cho is a target for the serum component C-reactive protein (CRP), which, when bound, mediates killing of the bacteria via activation of complement. However, the sensitivity to CRP depends on the position of *P*Cho in LPS (*72*).

Some NT*Hi* strains express highly acetylated LPS. It has been found that the addition of acetate to HepIII is controlled by the phase-variable gene *oafA* (*78*).
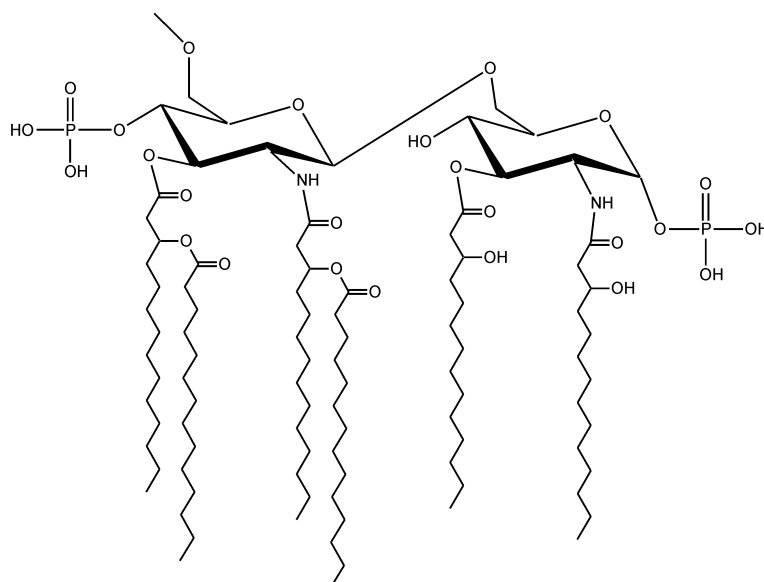
1.4.1.2. The lipid A structure of H. influenzae.

The structure of the lipid A was first established by Helander I. *et al* (*79*). In 2005 Mikhail I. *et al* characterized in detail the lipid A part from 22 NT*Hi* strains and two type f strains by ESI-MS^n (*80*).

The major structure of the lipid A is composed of two 2-amino-2-deoxy-D-glucopyranose residues with phosphates at C1 and C4´.

The C2/C2´and C3/C3´positions were found to be substituted by 3-hydroxytetradecanoic acids. Moreover the fatty acids at C3´and C2´were further esterified by tetradecanoic acids (Fig. 3).

8

Fig. 3. The lipid A structure of H. influenzae.

1.5. Haemophilus parainfluenzae.

Almost all people carry *H. parainfluenzae* and, by multi-locus sequence typing and partial 16S rRNA sequences (Derek Hood, unpublished data), it has been found to be a highly diverse population of organisms. *H. parainfluenzae* is closely related to *H. influenzae*. Despite their relatedness and similar presence in the nasopharynx, *H. parainfluenzae* strains very rarely cause diseases.

The significant difference in LPS structures is that some strains of *H. parainfluenzae* express smooth (S)-type LPS (*81*). Unlike LPS of *H. influenzae,* the structural information of the lipopolysaccharide from *H. parainfluenzae* is very limited.
Before our investigations, only one paper was published on the structures of LPS from *H. parainfluenzae* (strains 4201 and 4282), strains that are lacking O-antigen (*82*).
Interestingly, LPS expressed by these strains are similar to those of *H. influenzae* and are composed of the triheptosyl inner-core moiety *(33)*,

L-α-D-HepIII*p*-(1→2)-[*P*Etn→6]-L-α-D-HepII*p*-(1→3)-L-α-D-HepI*p*-(1→5)-α-Kdo*p*- (Scheme 1).

Furthermore the structures expressed in the outer-core region were found to be similar to those in LPS of *H. influenzae*. Specifically, in strain 4201 HepI is substituted by

β-D-Glc*p*-(1→4)-D-α-D-Hep*p*-(1→6)- β-D-Glc*p*-(1→

and strain 4282 expresses

β-D-Gal*p*-(1→6)-β-D-Glc*p*-(1→4)-D-α-D-Hep*p*-(1→6)- β-D-Glc*p*-(1→

linked to HepI.

2 METHODS

The wild type and mutant strains used in this thesis were provided by our colleagues from Oxford University.

The lipopolysaccharide can be obtained from bacteria by extraction. Two methods are used, the hot phenol-water extraction (*83*) and the PCP-method (phenol-chloroform-light petroleum) (*84*). The choice of method depends on the structure of LPS. The LPS from smooth strains is more hydrophilic than from rough strains. Therefore LPS containing the O-antigen is extracted by phenol-water method. The rough strains express shorter lipopolysaccharides, which can be extracted by PCP-method.

Usually, LPS very poorly dissolves in water, making it difficult to be analyzed in its native form. To resolve this problem LPS can be degraded to the oligosaccharide (OS) and the lipid A materials by the mild acid hydrolysis, or by O-deacylation with hydrazine to O-deacylated LPS (LPS-OH). It should be kept in mind that the mild acid hydrolysis also leads to hydrolysis of all acid sensitive linkages like the ketosidic linkage in sialic acid (Neu5Ac), phosphodiester linkages (*PP*Etn) and phosphoglycosidic linkages. On the other hand O-deacylation with hydrazine removes not only ester-linked fatty acids but also all acetate groups and ester linked glycine. Hence analyses of OS and LPS-OH alone do not give all structural information about carbohydrate part of LPS but together can complement one another.

In order to elucidate the structure of carbohydrate polymers it is necessary to determine the identity, the absolute configuration, the ring size and the linkage positions of all monosaccharide residues, as well as their sequence and anomeric configurations. In addition LPS is often decorated by non-carbohydrate substituents, for which linkage positions must be determined. This information can be analyzed by different mass spectrometry methods and by nuclear magnetic resonance (NMR) spectroscopy (Fig.4).

Sometimes LPS containing O-antigen can be analyzed directly by NMR spectroscopy due to the repeating OS unit structure.
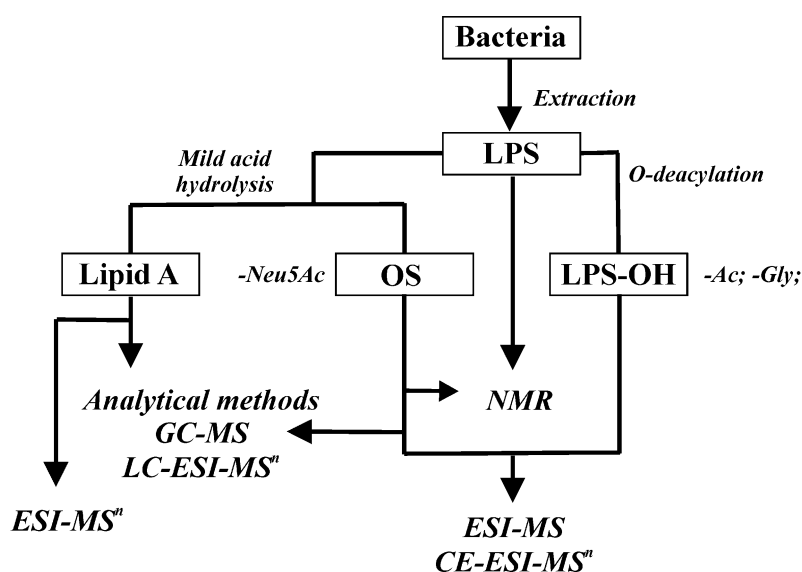
Fig. 4. Summary of methods used for the structural elucidation of LPS.

2.1 Preparation, purification and degradation of LPS, OS, LPS-OH and lipid A materials.

2.1.1 Bacterial cultivation.

H. influenzae strains 1158/1159 and 1232 were grown in brain-heart infusion broth supplemented with haemin (10µg·ml^{-1}) and NAD (2µg·ml^{-1}). *H. parainfluenzae* strains 20 and 16 were isolated as commensals from the throats of two children in Oxfordshire, UK. Bacteria were grown on solid brain-heart infusion (BHI) medium (agar 1% w/v) supplemented with 10% Levinthals reagent.

2.1.2 Extraction of lipopolysaccharides from bacteria.

PCP extraction.

H. influenzae express rough type LPS. Hence the PCP extraction method (*84*) is more preferable for extraction of LPS from NT*Hi* strains 1158, 1159 and 1232.
The lyophilized bacteria were dissolved in the phenol:chloroform:light petroleum mixture (2:5:8) and stirred at room temperature. After 24h the mixture was centrifuged (7500rpm; 30min), and the pellet was dissolved one more time in PCP solution. The supernatants from day one and day two were filtered through filter paper, pooled together and evaporated on rotary evaporator until only phenol phase was left. LPS was precipitated by adding a mixture of diethyl ether and acetone (1:5). The obtained LPS was washed with acetone. In the final step, LPS was purified by ultracentrifugation to remove all impurities as RNA and proteins.

Hot phenol-water extraction.

Since SDS-PAGE indicatied *H. parainfluenzae* strains *Hp*20 and *Hp*16 express (S)-type LPS, the hot phenol-water extraction method was chosen as more appropriate method for these strains.
The lyophilized bacteria were dissolved in a phenol:water (1:1) mixture and stirred at 68°C for 2h. The mixture was cooled and centrifuged (7500rpm; 40min) at 4°C. The water phase layer was removed and saved. The phenol phase layer with added water was stirred one more time at 68°C for 2h, and centrifuged. Both water phase layers were pooled together and dialyzed against tap water, followed by dialyzing by distilled water and lyophilized. The extracted LPS can be further purified by treatment with RNAse, DNAse and proteinase K, followed with dialysis against distilled water.

2.1.3 O-deacylation with hydrazine.

The O-deacylation with hydrazine is used to remove ester-linked fatty acids from the lipid A moiety (*85*).
To obtain the O-deacylated LPS (LPS-OH), the lipopolysaccharide was treated with anhydrous hydrazine at 40°C for 1h. After cooling on ice, the excess of hydrazine was destroyed by dropwise addition of cold acetone (hydrazine:acetone-1:4). The precipitated LPS-OH after washing with acetone was dissolved in water and lyophilized.

2.1.4 Delipidation by mild acid hydrolysis.

The ketosidic linkage between the Kdo and the lipid A can be selectively cleaved by mild acid hydrolysis (delipidation). During the delipidation the Kdo is changed to several anhydro-Kdo (*An*Kdo-*ol*) forms due to the β-elimination of a phosphate group at C-4 (*86*) (Fig. 5). The simultaneous reduction by borane-*N*-methyl-morpholine complex reduces the amount of *An*Kdo-*ol* forms. Thus, reduced core oligosaccharide (OS) samples and the lipid A were obtained after mild acid hydrolysis of LPS with 1-2% aqueous acetic acid at 100°C for 2h in the presence of borane-*N*-methyl-morpholine complex. The insoluble lipid A was separated from the mixture by centrifugation (7500rpm, 35min). The water-soluble part (OS) was purified by gel filtration on Bio Gel G-15 or Bio Gel P-4 columns. The lipid A was purified by partition using chloroform:methanol:water (2:1:1). The lower chloroform phase was evaporated to dryness.
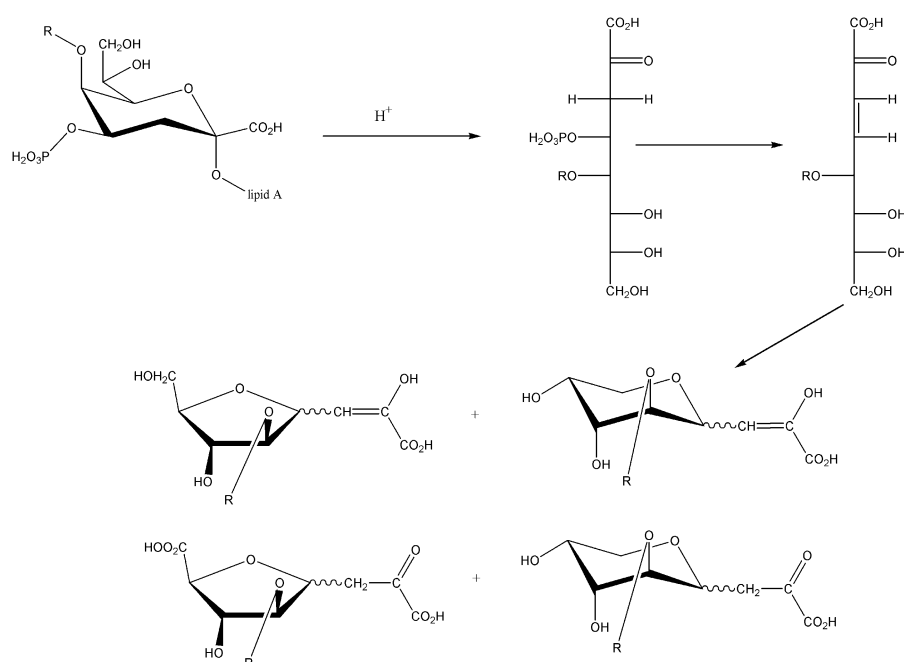
Fig. 5. Mild acid hydrolysis of LPS and formation of anhydro-forms of Kdo.

2.1.5 Dephosphorylation.

Dephosphorylation of OS materials was performed with 48% aqueous HF (4 °C, 48h) in order to remove all phosphate containing substituents.

2.1.6 O-Deacylation.

In order to reduce the heterogeneity due to the different amount of acetate groups and glycine in the same glycoform, the OS material was deacytylated by 1M NH_3 for 24h at room temperature.

2.2 Analytical methods with mass spectrometry.

The structural characterization of carbohydrates is a challenging task.
Monosaccharaide components from complex oligosaccharides typically differ from each other in their stereochemistry, and the positions of interglycosidic linkages. In addition, the oligosaccharides can be decorated by non-carbohydrate structures. Mass spectrometry is a very powerful and useful tool for the structural analysis of lipopolysaccharides (*87*).

Gas-liquid chromatography-mass spectrometry (GC-MS) is mostly used for analysis of volatile derivatives, such as alditol acetates (*88*), permethylated alditol acetates (*89*) and methyl esters which are identified by their retention time in the GC chromatogram and characteristic electron ionization (EI) spectra (Fig. *6*).
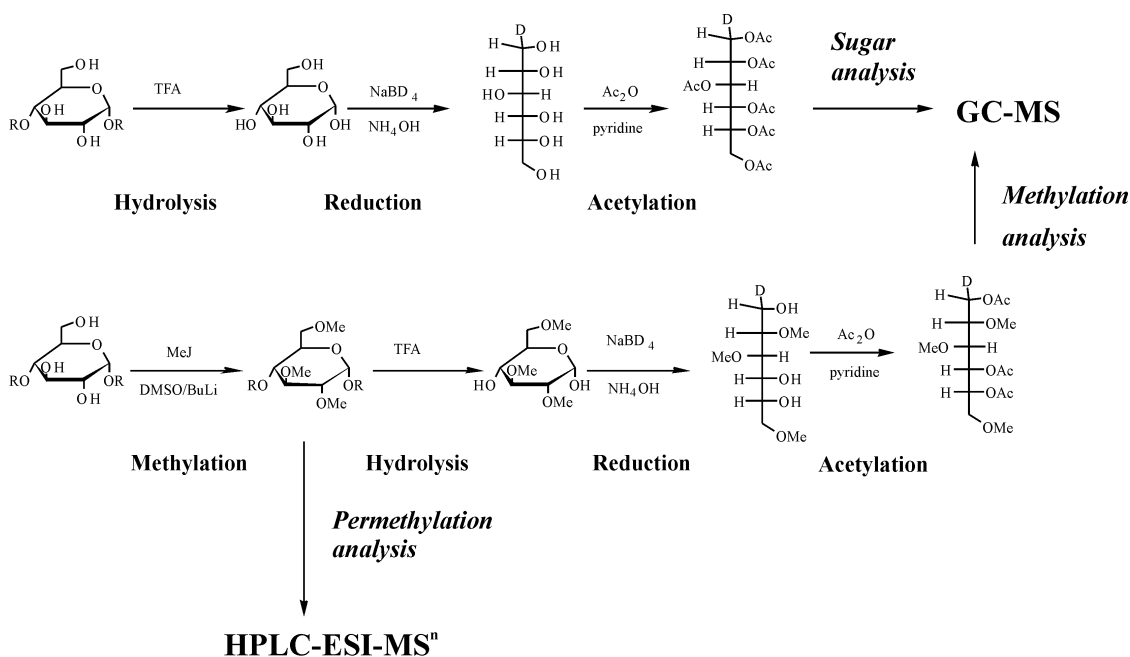
Fig. 6. Degradation methods used for the structural analysis of the oligosaccharide part of LPS.

Structural studies of carbohydrates were revolutionized by the development of mass spectrometry with mild ionization sources, such as fast atom bombardment (FAB), matrix-assisted laser desorption ionization (MALDI), and **electrospray ionization (ESI)**, which together with tandem spectrometric methods provide very powerful means for determination of carbohydrate sequence, in derivatized or native forms (*90-94*). ESI is the most effective method for transforming carbohydrate molecules from solution to gas-phase ions.

Different separation techniques such as **high-performance liquid chromatography (HPLC)** and **capillary electrophoresis (CE)** are often coupled to mass spectrometry to analyze complex mixtures of saccharides (*70, 95-98*).

In order to obtain the molecular mass of oligosaccharides and their distribution in the oligosaccharides mixture, OS and LPS-OH samples can be analyzed by ESI-MS in native form in positive or negative mode (*99*). It has been shown that HPLC using graphitized carbon columns (GCC) coupled to ESI-MS/MS is a powerful method, which can be used for structural characterization of complex mixtures of carbohydrates, without prior derivatization (*100, 101*).

The sensitivity of ESI-MS can also be improved by adding Li+, Na+ or K+ ions, especially Li+ ions (*102*). However, the sensitivity of ESI on underivatized carbohydrates is much lower than on peptideds and proteins. The improvement of electrospray ionization by nano-ESI technique (*103*) increases the sensitivity, due to increasing the surface activity of formed droplets. Underivatized carbohydrates can be measured by nano-ESI with the same level of sensitivity as proteins. Consequently the sequence and branching information of oligosaccharide can be obtained by nano-ESI -MS^n without prior derivatization of carbohydrates (*104*).

The main fragmentation ions observed in MS/MS spectra are formed by cleavages of glycosidic bonds. The nomenclature of a **fragmentation mechanism** was introduced by Domon, B., and Costello, C. E. and as shown in Fig. 7 allows detailed sequence information to be obtained (*105*).

Derivatization of carbohydrates by permethylation prior ESI-MS/MS increases the detection sensitivity of their ions by several orders. Permethylation (*106-108*) in conjunction with ionization by sodium adduction simplifies structural elucidation of carbohydrates.

In addition, permethylation permits to couple HPLC to ESI-MS/MS using reversed-phase column where selected ion monitoring can be used for selecting critical *m/z* values. Since α- and β- anomers can be separated by reversed-phase HPLC, the oligosaccharide should be reduced before permethylation (*106*).
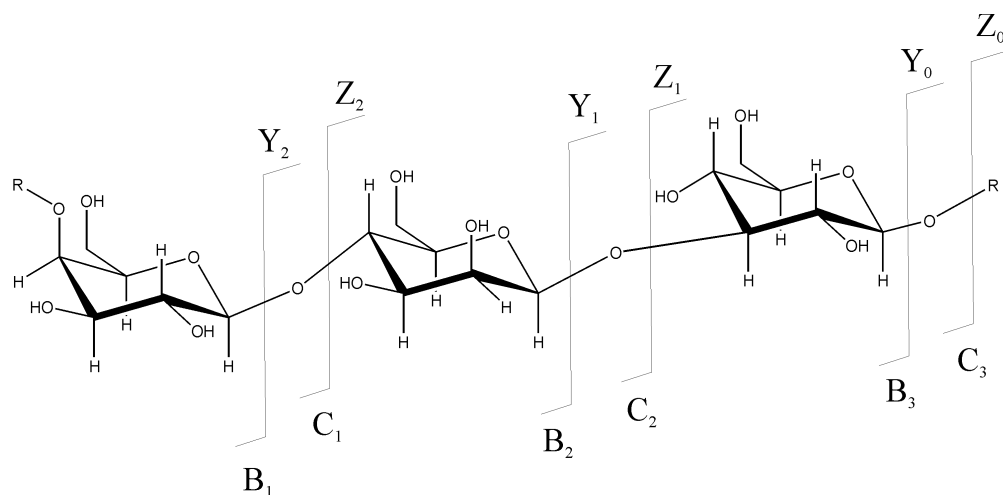
Fig. 7. The nomenclature of fragmentation ions obtained from cleavages of glycosidic bonds.

14

2.3 Analytical methods used in these studies.

Chemical modifications of the carbohydrates (lipopolysaccharides) followed by mass spectrometric methods provide important structural information such as:

1. The identity of monosaccharide residues (*sugar analysis*)

2. The D or L configuration of sugar residues (*absolute configuration analysis*).

3. The ring form i.e. pyranose or furanose form of monosaccharides (*methylation analysis*).

4. The linkage position, the position to which other glycosyl residues are linked (*methylation analysis*).

5. The molecular mass and sequence information (*permethylation analysis*).

6. The identity of fatty acids from to the lipid A part (*Fatty acid analysis*).

7. Molecular mass and relative distribution of glycoforms (*ESI-MS*).

8. Sequence information and the information about location of non-carbohydrate substituents such as *P*Cho, *P*Etn, Ac and Gly (*CE-ESI-MS on OS*).

9. Molecular mass and relative distribution of sialylated glycoforms by precursor ion monitoring by scanning for the loss of sialic acid and disialic acid (*CE-ESI-MS on LPS-OH*).

2.3.1 Sugar analysis.

The monosaccharide residues were identified by GC-MS as theirs corresponding alditol acetates (Fig. 6). The preparation of the sample includes:

> - Hydrolysis of glycosidic linkages* (2M, 0.5M TFA).
> - Subsequent reduction of monosaccharide residues by $NaBH_4$ in 1M NH_3 for 16h at 21°C.
> - Acetylation of hydroxyl groups with acetic anhydride/pyridine (1:1) for 20 min at 120°C.
> - Extraction with EtOAc:water (1:1).
> - Analysis by GC-MS.

***the choice of the hydrolysis condition depends on the sugar residues.**

2.3.2 Methylation analysis.

In order to identify the linkage positions of monosaccharaides to which other glycosyl residues are linked, OS was modified to partially methylated alditol acetates and analyzed by GC-MS (Fig. 6). Since phosphorylated sugars are not detected by GC-MS, the OS was dephosphorylated prior methylation.
The preparation includes:

> - Dissolving the OS in anhydrous DMSO.
> - Preparation of DMSO anion (BuLi, 40°C, 1h).
> - Methylation (MeI, 16h, 21°C).

- ➤ Evaporation of excess amount of MeI by vacuum.
- ➤ Purification on SepPackC18 column:
  - ▪ Preconditioned with 10mL ethanol, 4mL water.
  - ▪ Addition of an equal amount of water to sample.
  - ▪ Washing of applied sample with 10mL water, 6mL 10% acetonitrile in water.
  - ▪ Elution of sample with 4ml acetonitrile.
- ➤ Reduction and acetylation as in sugar analysis ($NaBD_4$ must be used instead of $NaBH_4$).
- ➤ Analysis by GC-MS.

2.3.3   Absolute configuration analysis.

The absolute configuration of sugar residues was obtained by modifications of sugar enantiomers to diastereomers (*109*). It was done by Fischer glycosylation with secondary alcohols, usually (+)-2-butanol.
Briefly,
- ➤ OS samples are hydrolyzed by 2M TFA at $120^{o}C$ for 2h.
- ➤ Re-N-acetylation step is done, when necessary (the presence of N-acetylhexoseamine sugar), by reaction with Ac_2O at $21^{o}C$ for 4h.
- ➤ Glycosylation ((+)-2-butanol, acetyl chloride, at $80^{o}C$, for 16h).
- ➤ Acetylation (acetic anhydride/pyridine (1:1) for 20 min at $120^{o}C$).
- ➤ Analysis by GC-MS.

2.3.4   Fatty acid analysis.

Fatty acids of the lipid A were derivatized to methyl esters and analyzed by GC-MS. Preparation method:
- ➤ The lipid A is treated with 4M HCl at $100^{o}C$ for 4h.
- ➤ The sample is extracted with 6mL of mixture of chloroform:water (1:1).
- ➤ The chloroform phase is collected and dried by Na_2SO_4.
- ➤ The solution is filtrated and evaporated to dryness.
- ➤ The sample is subjected to methanolysis (MeOH, acetyl chlorid, $80^{o}C$, 16h)
- ➤ The methyl esters are purified by extraction with chloroform:NaCl solution (30mg/ml), and chloroform phase is collected and evaporated.
- ➤ The methyl esters are dissolved in EtOAc and analyzed by GC-MS.

2.3.5   Permethylation analysis (HPLC-ESI-MS[n]).

Electrospray ionization tandem mass spectrometry (ESI-MS[n]) on dephosphorylated and permethylated oligosaccharide samples is a very important method for determination of the sequence and branching information. Dephosphorylation decreases the heterogeneity of samples.

- ➤ Dephosphorylation and methylation were achieved as described above (**2.3.2**).
- ➤ HPLC-ESI-MS[n] on dephosphorylated and permethylated OS samples was carried out on a Waters 2690 system coupled to the Finnigan LCQ ion trap mass spectrometer) in the positive ion mode.
- ➤ A microbore C18-column (Phenomenex LUNA 5u C18) was used with an eluent gradient consisting of 1mM NaOAc and 1% HOAc in MeOH as

eluent A and 1mM NaOAc and 1% HOAc in water as eluent B. A gradient program was used with 50% A rising to 100% in 50min and thereafter 100% A for 20 min. The flow rate was 100µL/min.

2.3.6 ESI-MS.

ESI-MS (Finnigan LCQ ion trap mass spectrometer) on

- OS samples (positive mode) were done using a running solvent of 1% acetic acid in aceteonitrile/water (1:1, v/v) and the flow rate of 5µL/min.
- LPS-OH samples (negative mode) were done using a running solvent of 1M NH_4OH in water and the flow rate of 5µL/min.

2.3.7 CE-ESI-MS.

CE-ESI-MS[n] experiments in positive and negative mode on OS and LPS-OH samples provided information about molecular masses of glycoforms and their distribution. Furthermore CE-ESI-MS[n] on OS samples provide the information about location of non-carbohydrate substituents such as: *P*Cho, *P*Etn, Ac and Gly. The distribution of sialylated glycoforms is very low but they can be detected by precursor ion monitoring for scaning for the loss of sialic acid (*m/z* 290) and disialic acid (*m/z* 581). These CE-ESI-MS[n] experiments were performed by our colleges in Institute for Biological Sciences, Ottawa.

- CE-ESI-MS[n] experiments on OS and LPS-OH materials were carried out in negative and positive mode with a Crystal model 310 CE instrument coupled to an API 3000 mass spectrometer via a MicroIonspray interface as described previously (*68*).

2.4 NMR spectroscopy.

Nuclear Magnetic Resonance (NMR) spectroscopy has been used for structural studies of carbohydrates for relatively long time. Developments in instrumentation and pulse sequences have made NMR spectroscopy a very powerful and absolute necessary technique for structural elucidation of carbohydrates and carbohydrate containing structures (*110-115*).

The detailed structural analysis of carbohydrates by 1D and 2D NMR spectroscopy can give information about the identity of monosaccharide, their anomeric configuration, ring form, linkage positions, sequence information; even information about absolute configuration can be achieved. However the structural information from analytical and mass spectroscopic methods is often needed for simplifying the elucidation and for confirming the results form NMR spectra. The major weakness of NMR spectroscopy is its sensitivity. But one of the advantages is that NMR is a non-destructive method.

The spectra of carbohydrates are usually recorded in D_2O. The almost full exchange of all exchangeable protons (NH, OH) can be gained by lyophilization from D_2O prior analysis. When determination of chemical shifts of NH_2 and NH protons is needed for confirming positions of amino or acetamido groups the sample can be recorded in a mixture of H_2O/D_2O.

Due to thermostability of carbohydrates, they can be analyzed over a wide range of temperature, usually between 22-85°C. The chemical shifts of some protons can be temperature dependent, but the effect is usually small. On the other hand the change of the temperature leads to large shifts of the HDO signal so that sugar resonances under HDO can be observed.

2.4.1 One- dimensional spectra.

1D ^{1}H is a first step in structural studies by NMR. First of all, the 1D ^{1}H spectrum gives information about purity of the sample. The 1D ^{1}H spectrum of carbohydrate has very characteristic pattern and can be divided into several regions.
The first region is a region of anomeric protons at 4.3-5.9 p.p.m. which can give preliminary information about the number of sugar residues in the carbohydrate structure (Fig. 8A)
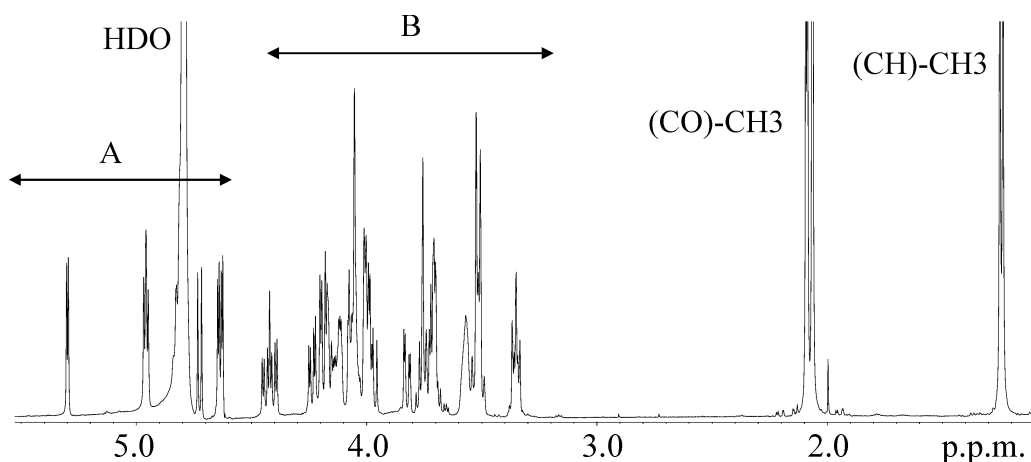
Fig. 8. The 1D ^{1}H spectrum of OS-2 derived from H. parainfluenzae strain 20.
A-region of anomeric protons; B-region of ring protons; (NH-CO)-CH3- methyl signals from acetoamido groups; (CH)-CH3- methyl protons from 6-deoxy sugar.

The well resolved anomeric signal appears as doublet and coupling constant gives information about anomeric configuration of the sugar. The coupling constant about 4Hz corresponds to α-anomeric configuration and the coupling constant about 8Hz corresponds to β-anomeric configuration (Fig. 9).
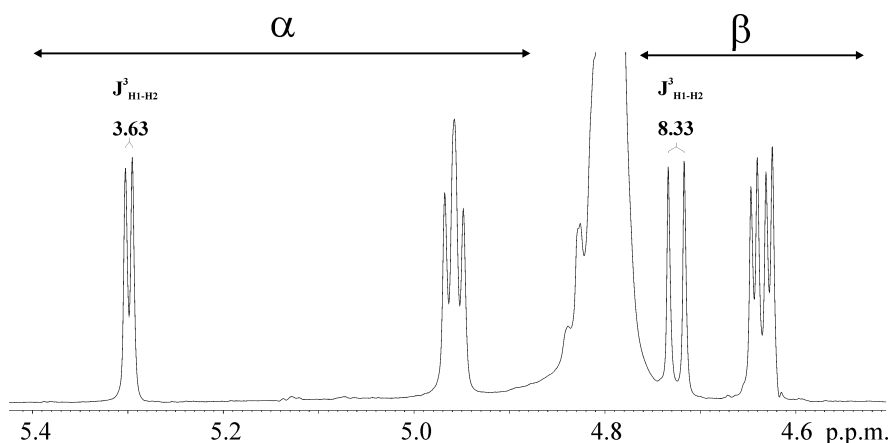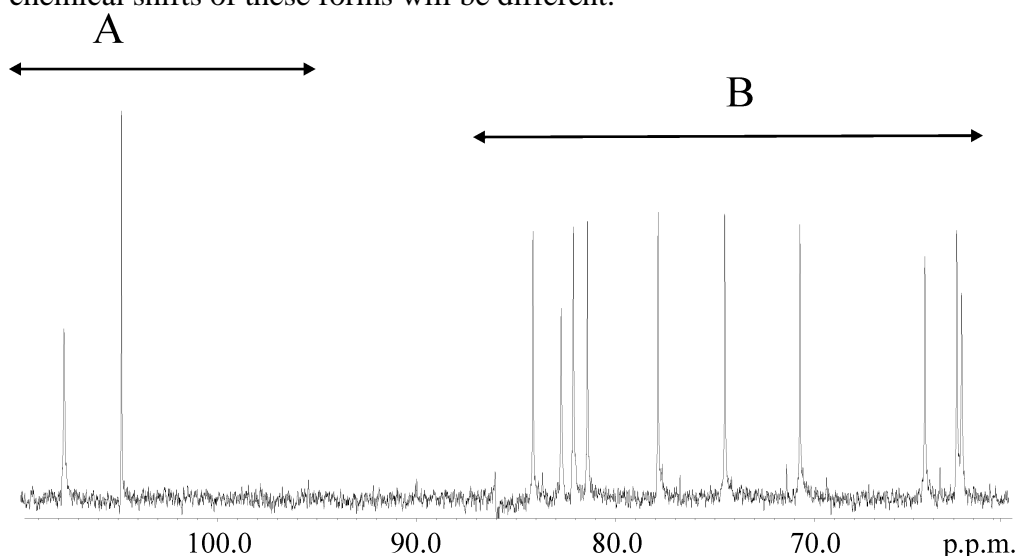

Fig. 9. Anomeric region of the 1D ^{1}H spectrum of OS-2 from LPS from Hp20.

18

These statements can be made for sugar residues having *gluco-* and *galacto-* configurations and adopting pyranose form. For pyranose sugar residues having *manno*-configuration $^{3}J_{H-H}$ is small for both α- and β-anomers, as well as for furanoses.

Methyl protons of 6-deoxy sugar appear as doublet between 1.1-1.3 p.p.m. and methyl singlets of acetamido groups at 2.0-2.2 p.p.m. The remaining majority of proton resonances appear in a very narrow region at 3.0-4.2 p.p.m. (Fig. 8B), making further interpretation by 1D NMR impossible.

^{13}C NMR spectroscopy is much less sensitive than ^{1}H due to the low natural abundance of the ^{13}C nucleus (1.1%). But 1D ^{13}C spectroscopy can be very useful for determination of carbon chemical shifts of carbohydrates due to its greater dispersion (Fig. 10). The anomeric carbon signals resonate in a region 90-112 p.p.m. giving more clearly information about the number of O-linked monosaccharaides. However, the monosaccharide at reducing end can adopt different forms and thereby the chemical shifts of these forms will be different.

Fig. 10. The 1D ^{13}C spectrum of LPS from H. parainfluenzae strain 16.

A-anomeric region, B-region of ring carbons.

The resonances between 52 and 57 p.p.m. indicate the presence of animo-substituted carbons. The presence of acetamido groups can be confirmed by methyl resonances from amino sugar residues between 21-24 p.p.m. The unsubstituted hydroxymethylene (C-6) gives signal between 57.7-64.7 p.p.m. However, the glycosylated C-6 resonates in region between 66-70 p.p.m. Resonances in the region 16-19 p.p.m. indicate the presence of 6-deoxysugars (H_3C-6). Resonances in the region between 80-85 p.p.m. usually, indicate the presence of furanoses. However, 1D spectrum cannot provide all information to complete structural characterization of an unknown carbohydrate.

2.4.2   Two-dimensional spectra.

Through 2D NMR the severe resolution problem related to determination of ring proton resonates (they resonate in narrow region at 3.0-4.2 p.p.m) has been generally overcome.

In general, structural studies of carbohydrates by NMR include: 2D COSY, TOCSY experiments, which are used to assign all protons belonging to closed spin systems; heteronuclear experiments such as ^{1}H-^{13}C HMQC or HSQC are used to obtain carbon chemical shifts and, the sequence information can be achieve by ^{1}H-^{1}H NOESY and/or ^{1}H-^{13}C HMBC experiments.

2.4.2.1.   COSY.

Correlation spectroscopy (COSY) is a homonuclear experiment which allows the identification of the proton chemical shifts through identification of scalar coupled spins.

The determination of proton chemical shifts of carbohydrates can be started from anomeric proton, which is coupled only to one proton and gives well-resolved cross-peaks to H-2. H-2 gives cross-peaks to H-3 and H-3 gives cross-peaks to H-4 etc. (Fig. 11).

However the assignment of all proton chemical shifts within closed systems can be very difficult or impossible due to overlapping signals or the lack of cross-peaks due to ^{3}J couplings constants that are too small.
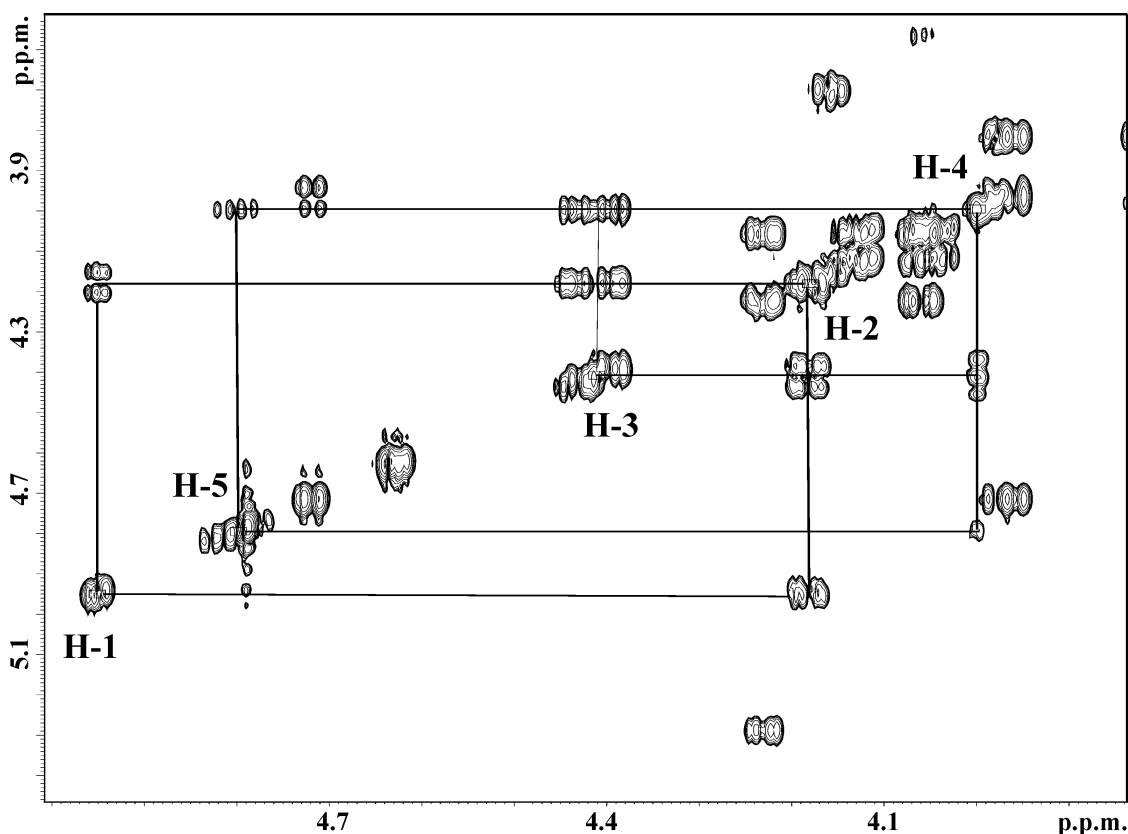
Fig. 11. The DQF-COSY (Double -quantum filtered COSY) spectrum of OS-2 from Hp20.

The DQF-COSY (Double -quantum filtrered COSY) experiment is preferred to COSY for two reasons. It gives a nicer spectrum with better balance of intensity between the cross-peaks and diagonal-peaks and the spectrum doesn't contain the signals from uncoupled spins.

2.4.2.2   TOCSY.

The identification of proton chemical shifts which belong to the same spin system can be achieved by a TOCSY (Total Correlation Spectroscopy) experiment.
In the TOCSY spectrum the chemical shift of one spin shows the correlation to all spins of the unbroken chain. The chain can be "broken" between two neighboring spins from one spin system if the coupling constant is too small. However it also points the configuration of sugar residues (*gluco-*, *galacto-* and *manno-*configuration) (see Fig. 24).

2.4.2.3   NOESY.

In the two-dimensional NOESY (Nuclear Overhauser Effect Spectroscopy) spectrum the cross-peaks appear between two spins that are close in space (less than 5Å).
Typically, the protons from the glycosidic bond are close in space. The analysis of cross-peaks in the anomeric region gives information about substitution positions and sequences of sugar residues, the information about the anomeric configurations (α, β) as well as confirming *galacto- gluco-* and *manno-*configurations (see Fig. 24).

2.4.2.4   HMQC (HSQC) and HMBC.

HMQC (Heteronuclear Multiple Quantum Coherence) or HSQC (Heteronuclear Single Quantum Coherence) experiments are used to identify the carbon chemical shift through one bond correlation between directly attached ^{1}H and ^{13}C observed by ^{1}H NMR (Fig. 12).

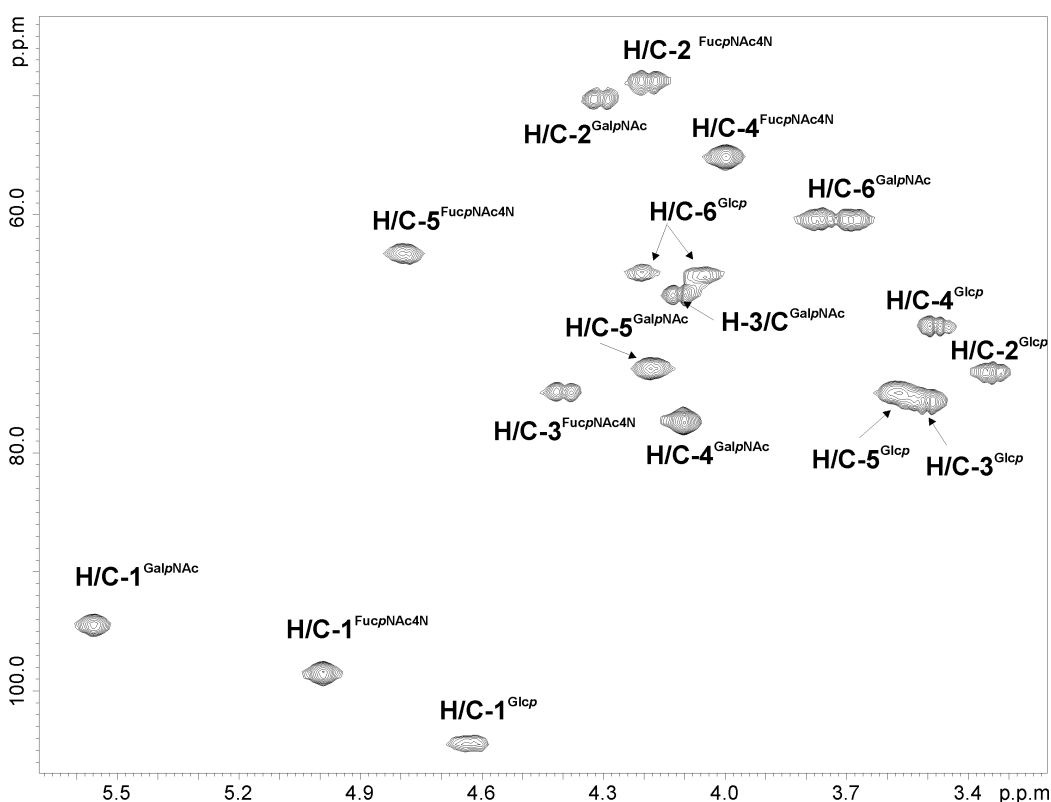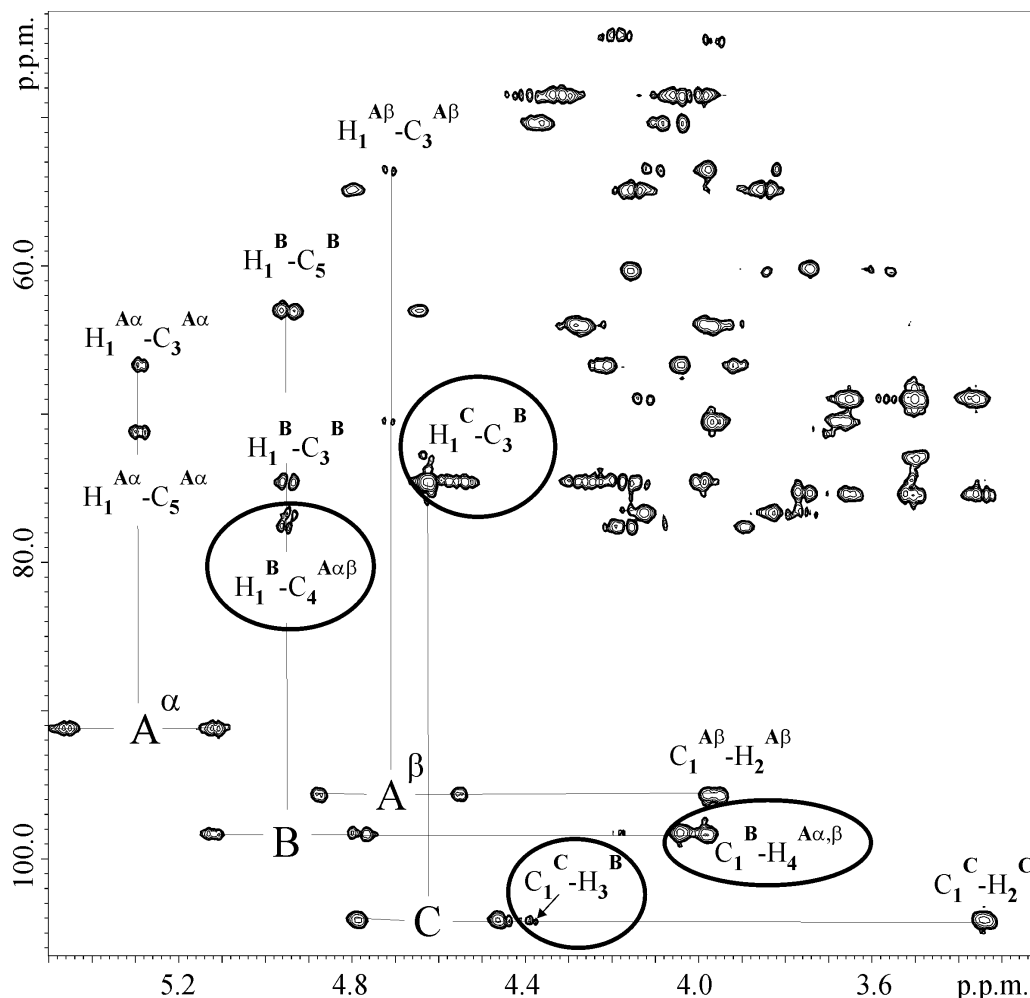***Fig. 12. ^{1}H-^{13}C HMQC spectrum on LPS from H. parainfluenzae strain 20.***

The HMBC (Heteronuclear Multiple-Bond Coherence) experiment is very useful for sequencing (Fig. 13). ^{1}H-^{13}C correlations over three bonds are particularly valuable for determining linkage positions of carbohydrates with low molecular masses.

Fig. 13. Selected region of the ^{1}H-^{13}C HMBC spectrum of OS-2 from

H. parainfluenzae strain 20.

*The inter-residue cross-peaks between C-1 of **C** and H-3 of **B**; C-1 of **B** and H-4 of **A**$^{\alpha,\beta}$, as well as H-1 of **C** and C-3 of **B**; H-1 of **B** and C-4 of **A**$^{\alpha,\beta}$ indicated that **C** was substituted by **B** at O-4 position , which was further substituted by **A** at O-3 position.*

P→6)β-D-Glc*p*-(1→3)- α-D-Fuc*p*NAc4N-(1→4)- α,β-D-Gal*p*NAc
**C**   **B**   **A**
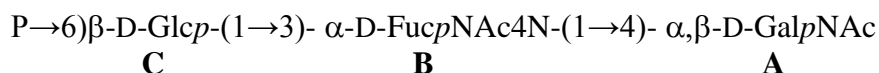
3 RESULTS AND DISCUSSION

Structural studies of LPS from non-typeable Haemophilus influenzae.

NT*Hi* strains 1158/1159 and 1232 were obtained from the collection of 25 NT*Hi* otitis media clinical isolates that express additional D,D-Hep residues in the outer-core region of LPS.

3.1 Paper I.
The structural diversity of lipopolysaccharide expressed by non-typeable *Haemophilus influenzae* strains 1158 and 1159.

The analysis of two NT*Hi* strains 1158 and 1159, which were obtained from the left and right middle ear of one patient with otitis media on the same day, indicated that both strains express almost identical lipopolysaccharide structures. The only differences which were found were the levels of acylation and phosphorylation. LPS were isolated by phenol/chloroform/light petroleum method and were degraded to O-deacylated LPS (LPS-OH) and core oligosaccharides (OS). Analyses were done by NMR spectroscopy on deacylated OS, capillary electrophoresis coupled to electrospray ionization mass spectrometry (CE-ESI-MS) on LPS-OH and OS materials, as well as HPLC-ESI-MS^n on permethylated and dephosphorylated OS (Fig 14). It was confirmed that LPS 1158 and 1159 contain the conserved triheptosyl inner-core which is attached via the phosphorylated Kdo unit to the lipid A moiety.


1716.0 Hep-Hex-Hep-*An*Kdo-*ol*
Hep
Hex-Hep
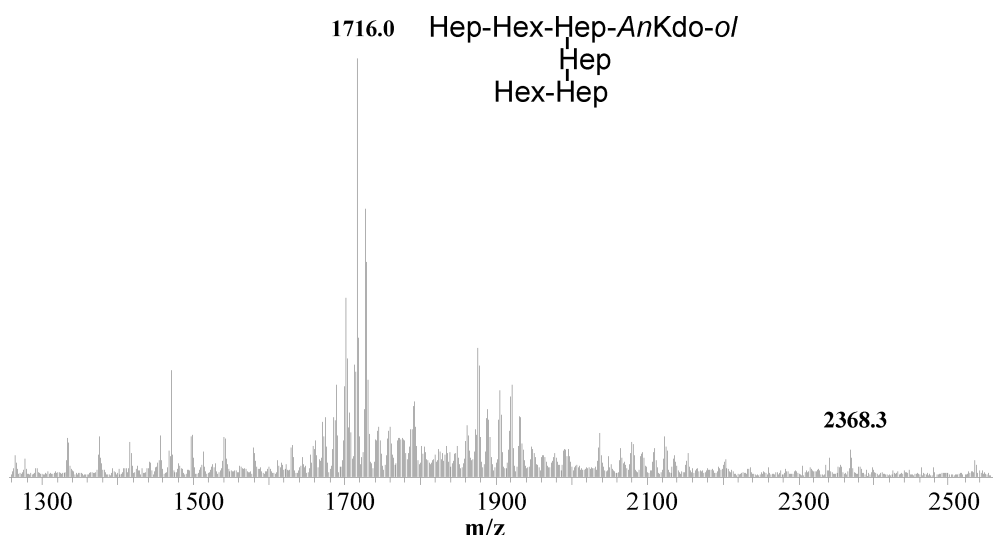
2368.3

Fig. 14. ESI-MS spectrum of dephosphorylated and permethylated OS 1158.

The ion at m/z *1716.0 corresponds to Hex$_2$Hep$_4$AnKdo-ol glycoform, and the ion at* m/z *2368.3 to HexNAc$_1$Hex$_4$Hep$_4$AnKdo-ol.*

The structural studies revealed that strains 1158 and 1159 express the additional D,D-Hep residue in the outer core region of the LPS. In the most abundant glycoform the D,D-Hep residue is terminal and links to O-6 of the β-D-Glc$p$ linked to HepI.

However HPLC-ESI-MS^n analysis also indicated glycoforms in which the external heptose was further substituted by HexNAc-Hex-Hex (Fig. 15).
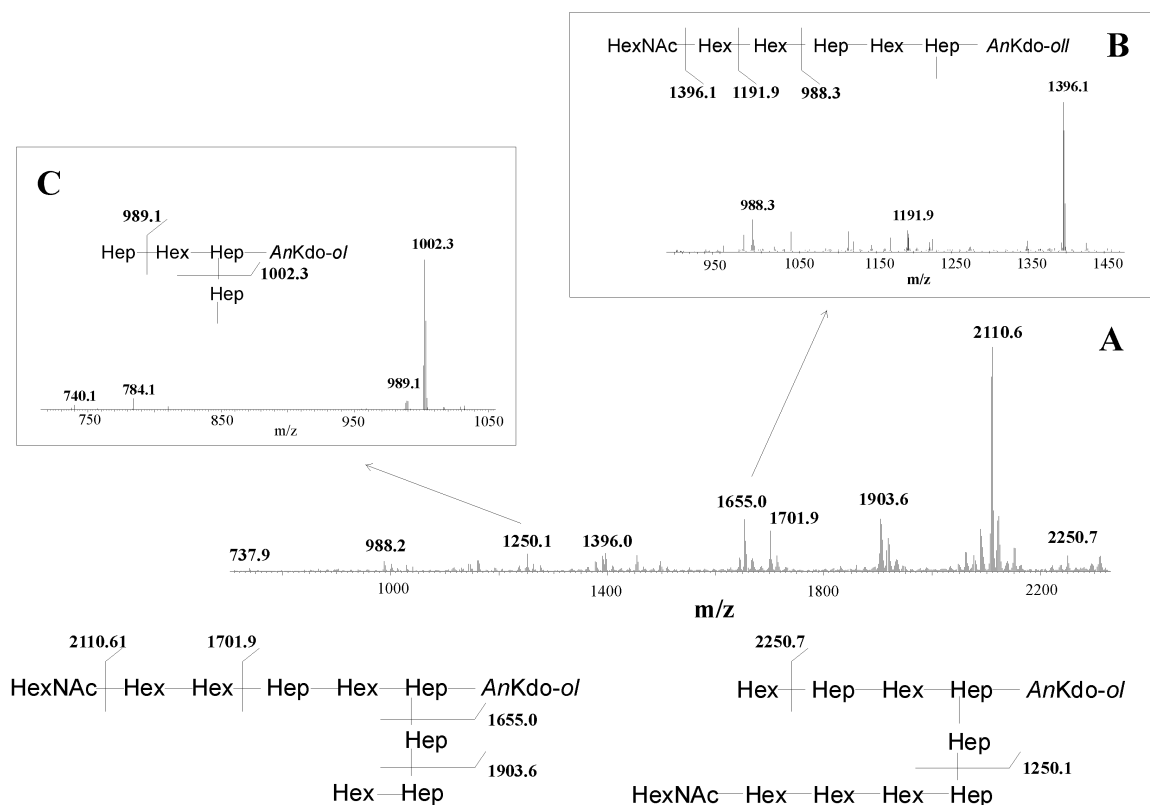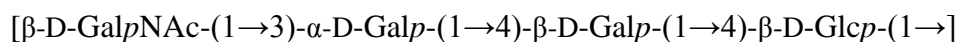

Fig. 15. HPLC-ESI-MS^n analysis on dephosphorylated and permethylated OS 1159.

(A) MS^2 spectrum of ion at m/z 2368.3, corresponding to HexNAc$_1$Hex$_4$Hep$_4$Ankdo-ol glycoform.
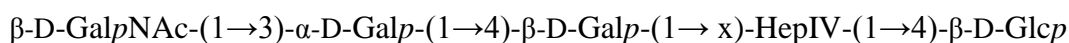*(B) MS^3 spectrum of ion at m/z 1655.0, corresponding to the loss of terminal hexose and 2-substituted heptose.*
(C) MS^3 spectrum of ion at m/z 1250.1, corresponding to the loss of HexNAc-Hex-Hex-Hex-Hep-.

The same structural element can be attached to β-D-Glc$p$ linked to HepIII (*33*). Although these structures could not be confirmed by NMR analysis, the data from methylation and HPLC-ESI-MS^n analyses clearly showed the expression of 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→]

from HepIII, as well as allowing us to propose the presence of

β-D-Gal$p$NAc-(1→3)-α-D-Gal$p$-(1→4)-β-D-Gal$p$-(1→ x)-HepIV-(1→4)-β-D-Glc$p$

from HepI.

24

The linkage position of the external heptose could not be determined due to the low abundance of these glycoforms. Interestingly, methylation analysis did not show any substituted D,D-Hep, but instead showed small amount of O-6 substituted L,D-Hep (Fig. 16).

Fig. 16. Methylation analysis of dephosphorylated OS-1158

β-D-GalNAc-(1→3)-α-D-Galp-(1→4)-β-D- Galp-(1→?)-HepIVp-(1→6)-β-D-Glcp-(1→4)-L-α-D-HepIp-(1→5)AnKdo-ol
$$\begin{array}{c} 3 \\ \uparrow \\ 1 \\ \text{L-α-D-HepIIp6←}P\text{Etn} \end{array}$$

α-Neu5Ac-(2→8)-α-Neu5Ac-(2   2
          ↓   ↑
          3   1
β-D-GalNAc-(1→3)-α-D-Galp-(1→4)-β-D- Galp-(1→4)-β-D-Glcp-(1→2)-L-α-D-HepIIIp

Scheme 3. Structure proposed for the Hep$_4$ glycoforms of NTHi strains 1158/1159.

Previously, two genes *losB1* and *losB2* were found, which are responsible for addition of D,D-Hep or L,D-Hep, respectively (*44*). These two genes were found alongside *losA1, losA2* genes encoding glycosyltransferases that are responsible for substitution of the external heptose (*116*). DNA sequence analyses indicated that strains 1158/1159 have *losA2/losB2* and, however instead of *losA1/losB1* they have *lic2B/lic2C* gene pairs. Nevertheless they still express the D,D-Hep residue in the outer-core region of LPS. The confirmation, that *losB2* is responsible for incorporation of D,D-Hep in strains 1158/1159 was done by structural analysis of 1158*losB2* mutant strain. HPLC-ESI-MS[n] analysis on permethylated and dephosphorylated OS 1158*losB2* indicated only glycoforms, containing three heptoses.
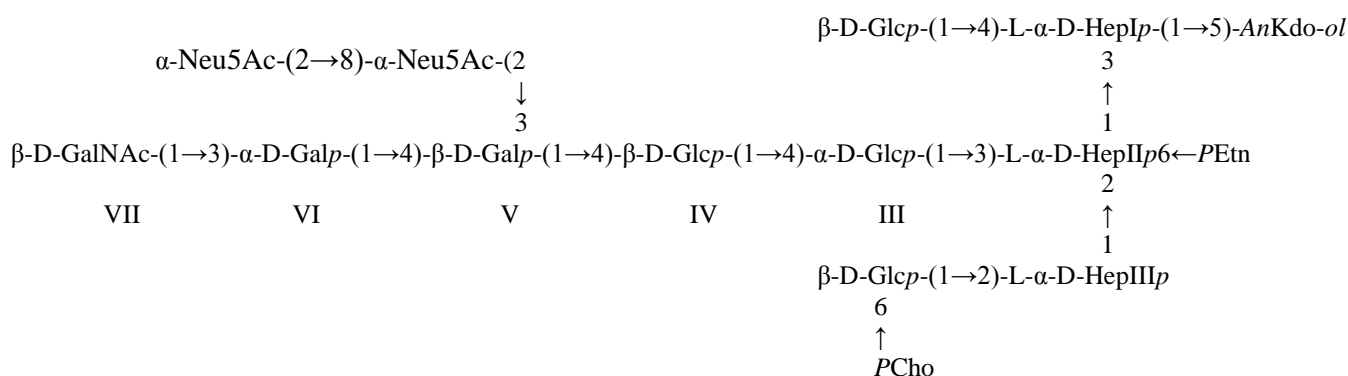
More likely that despite the high degree of homology between *losB2* from 1158 and *losB2* from R2846 and 1207, the function of 1158LosB2 is more closely related to the function of LosB1 in the other NT*Hi* strains (*116*).

Giving the small amount of 6-substituted D,D-Hep observed by methylation analysis, it is possible that LosB2 in 1158 may still possess low level activity to add L,D-Hep to β-D-Glc*p* linked to HepI . Then L,D-Hep (HepIV) would be elongated as shown in Scheme 3. This would need to be confirmed through further investigations.

The *lic2B/lic2C* gens are responsible for elongation from middle heptose (HepII) (*42*). It was found that α-D-Glc*p* linked to HepII can be also substituted by 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→] or truncated version of it. Interestingly, chain elongation from HepII was found only in Hep3 glycoforms, probably due to steric hinderance.
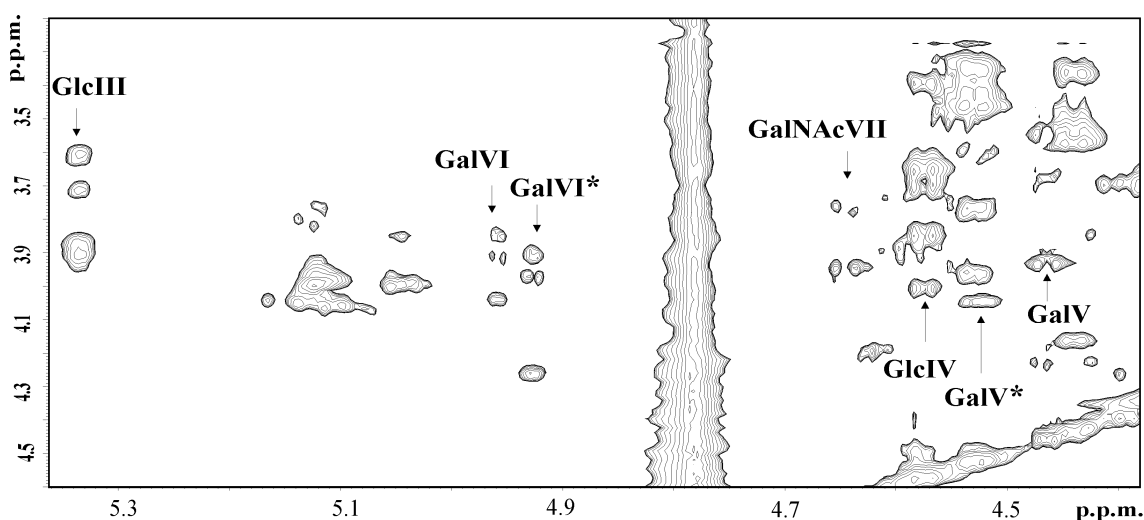
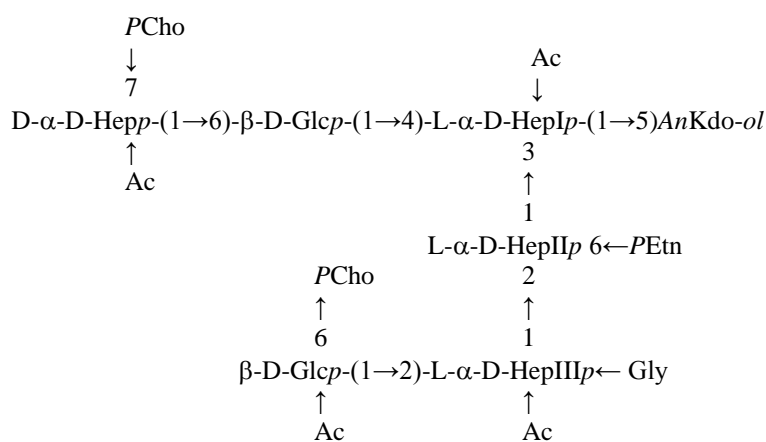*Fig. 17. Selected regions of the 2D TOCSY spectrum of deacylated OS 1158losB2*

```
                                                   β-D-Glcp-(1→4)-L-α-D-HepIp-(1→5)-AnKdo-ol
        α-Neu5Ac-(2→8)-α-Neu5Ac-(2                                     3
                             ↓                                         ↑
                             3                                         1
β-D-GalNAc-(1→3)-α-D-Galp-(1→4)-β-D-Galp-(1→4)-β-D-Glcp-(1→4)-α-D-Glcp-(1→3)-L-α-D-HepIIp6←PEtn
                                                                      2
        VII              VI            V              IV        III    ↑
                                                                      1
                                                   β-D-Glcp-(1→2)-L-α-D-HepIIIp
                                                                6
                                                                ↑
                                                              PCho
```

Scheme 4. Structure proposed for Hep₃ glycoforms of NTHi strains 1158/1159.

CE-ESI-MS[n] experiments on LPS-OH samples allowed us to suggest that a lactose unit [β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→] linked to HepIII (Scheme 3) or to α-D-Glc*p* attached to HepII could be substituted by sialic or disialic acid (Scheme 4).

26

The strains 1158/1159 can be decorated by one or two *P*Cho residues (*73*). These strains have two distinct copies of the *lic1* operon consisting of *lic1A, lic1B, lic1C* and *lic1D* genes, which are responsible for addition of *P*Cho (*73*). Location of *P*Cho depends on the sequence of *lic1D*. Hence *lic1D* is responsible for addition of *P*Cho at O-6 position to β-D-Glc*p* linked to HepIII, and lic1D* for addition to D,D-Hep*p* at O-7 position.

This novel linkage position of *P*Cho attached to D,D-Hep was determined by structural study of LPS from 1158*lpsA* mutant strain. Since the *lpsA* gene is responsible for chain elongation from HepIII, the major glycoform from LPS 1158*lpsA* can contain only one *P*Cho linked to D,D-Hep.

CE-ESI-MS^n experiments indicated also that LPS from NT*Hi* strains 1158/1159 can be decorated by up to four acetate groups and by glycine (Scheme 5)

```
        PCho
         ↓                                    Ac
         7                                     ↓
D-α-D-Hepp-(1→6)-β-D-Glcp-(1→4)-L-α-D-HepIp-(1→5)AnKdo-ol
         ↑                                     3
        Ac                                     ↑
                                               1
                                  L-α-D-HepIIp 6←PEtn
                     PCho          2
                      ↑            ↑
                      6            1
            β-D-Glcp-(1→2)-L-α-D-HepIIIp← Gly
                      ↑            ↑
                     Ac           Ac
```

Scheme 5. Structure proposed for the Hex$_2$ Hep$_4$ glycoform of NTHi strains 1158/1159.

3.2 PaperII.

The structural studies of a novel branching pattern in the lipopolysaccharide expressed by non-typeable *Haemophilus influenzae* strain 1232.

NT*Hi* strain 1232 was obtained from a patient with otitis media. The lipopolysaccharide, extracted by phenol/chloroform/light petroleum method was O-deacylated by anhydrous hydrazine. LPS was also degraded to OS material by mild acid hydrolysis, and further purified by gel filtration.

ESI-MS^n and CE-ESI-MS^n on OS and LPS-OH samples as well as the detailed structural analysis by NMR spectroscopy confirmed that LPS from NT*Hi* strain 1232 comprises the conserved *P*Etn substituted triheptosyl inner core which is attached via the phosphorylated Kdo unit to the lipid A moiety.

HPLC-ESI-MS experiment on dephosphorylated and permethylated OS as well as ESI-MS on OS fractions indicated that the most abundant glycoform contains four hexoses and four heptoses (Fig. 18). Further analysis by MS^2 and MS^3 and MS^4 experiments showed that the external heptose was di-substituted by two hexoses.
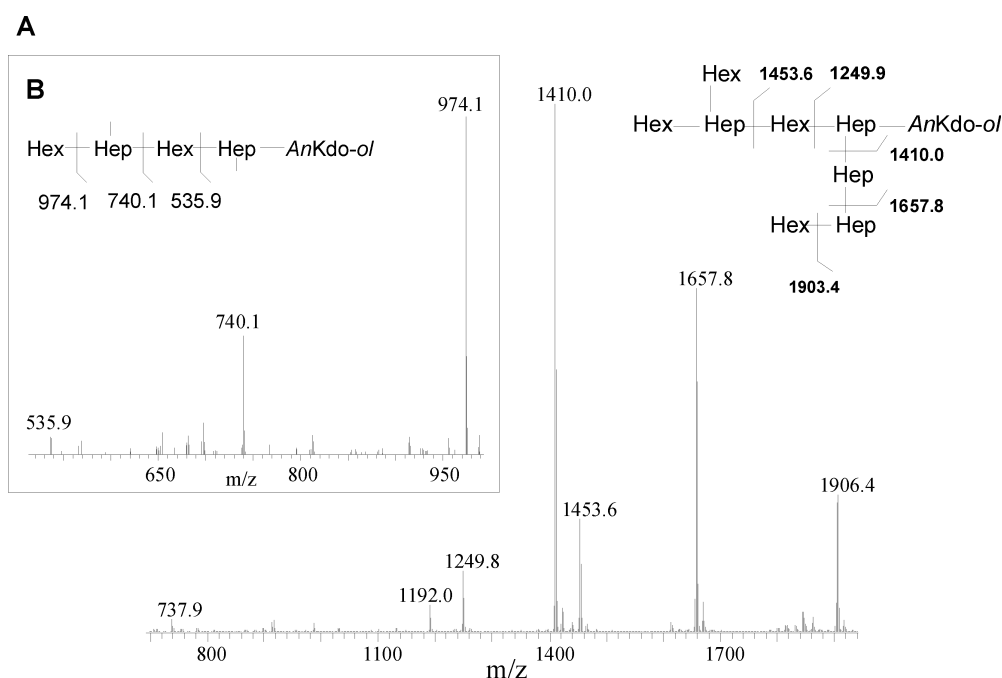
Fig. 18. HPLC-ESI-MS^n analysis of permethylated dephosphorylated OS-1232.
(A) MS^2 spectrum of ion at m/z *2124.0 corresponding to Hex$_4$Hep$_4$AnKdo-ol glycoform.*
(B) MS^4 spectrum of the fragment ion at m/z *1396.1 corresponding to the loss of terminal hexose from the ion at* m/z *1410.0 (MS^3 spectrum not shown).*

Thus NT*Hi* strain 1232 is the first reported strain that expresses di-substituted heptose in the outer-core region of LPS. Previously it has been shown that the D,D-Hep residue can be substituted by β-D-Glc*p* or β-D-Gal*p* at O-4 position (*35, 44*). Elucidation of the structure of LPS 1232 indicated that the D,D-Hep residue can be also substituted by β-D-Gal*p* at O-7 position (Fig. 19).

28
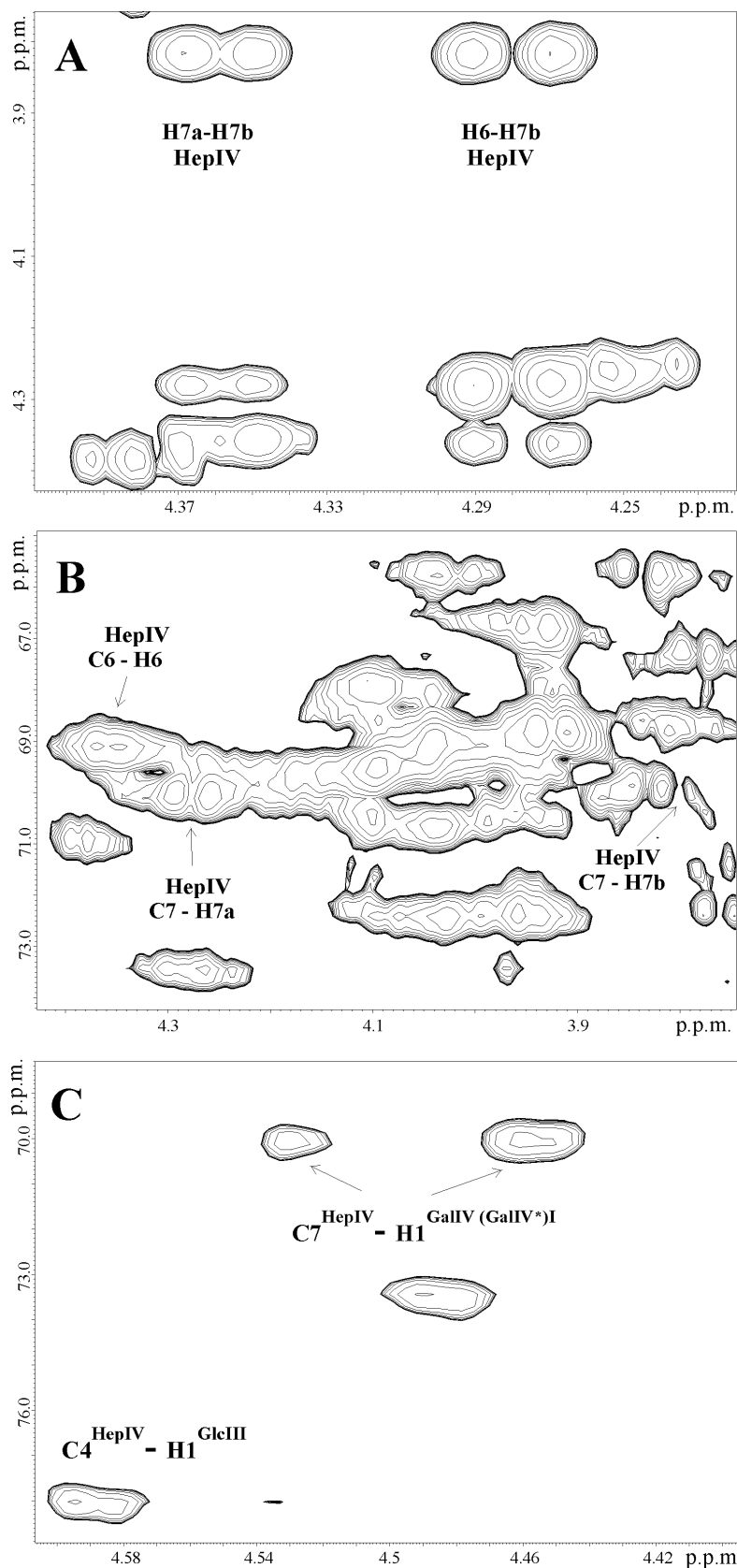
Fig. 19. Selected regions of 2D ^1H-^{13}C HMBC (A), ^1H-^{13}C HMQC (B) and TQF COSY (C) spectra. TQF-COSY spectrum of deacylated OS-1232(A) indicated that 3 signals at δ_H 4.28, δ_H 4.36 and δ_H 3.81 belong to –CH-CH$_2$- system. Carbon chemical shifts were obtained by ^1H-^{13}C HMQC experiment (B) and indicated that signal at δ_C 69.1/δ_H 4.36 belong to C6/H6, and signals at δ_C 69.1/δ_H 4.28, 3.81belong to C7/H7a,b. ^1H-^{13}C HMBC experiment (C), which showed the cross-peaks between H1 of terminal and 4-substituted GalIV comfirmed the chemical shift of C7 of HepIV*, and also indicated that D,D-HepIV was substituted at O-7 by β-D-Galp.

* methylation analysis of dephosphorylated OS-1232 showed the presence of 4,7-disubstituted D,D-Hep.

The *losB1* gene responsible for addition of D,D-Hep is adjacent to the glycosyltransferase gene *losA1*. This *losA1* is required for addition of β-D-Glc to D,D-Hep at O-4 position (*116*). The glycosyltransferase gene responsible for addition of β-D-Gal to D,D-Hep at O-7 position is still under investigation.

HPLC-ESI-MS^n on permethylated and dephosphorylated OS, ESI-MS and CE-ESI-MS experiments on OS and LPS-OH samples as well as data from methylation analysis, indicated the presence of sialyllacto-*N*-neotetraose [α-Neu5Ac-(2→3)-β-D-Gal*p*-(1→4) β-D-GlcNAc*p*-(1→3)-β-D-Gal*p*-(1→] epitope linked to GlcI at O-4 position.

Purification of OS material by gel filtration resulted in two fractions, OS-1 and OS-2. ESI-MS analyses as well as NMR experiments indicated that these fractions differed in the presence of *P*Cho. A ^{1}H-^{31}P HMQC experiment on the deacylated OS-1 sample revealed that β-D-Gal*p* linked to HepIII can be substituted by *P*Cho at O-6 position (Fig. 20).
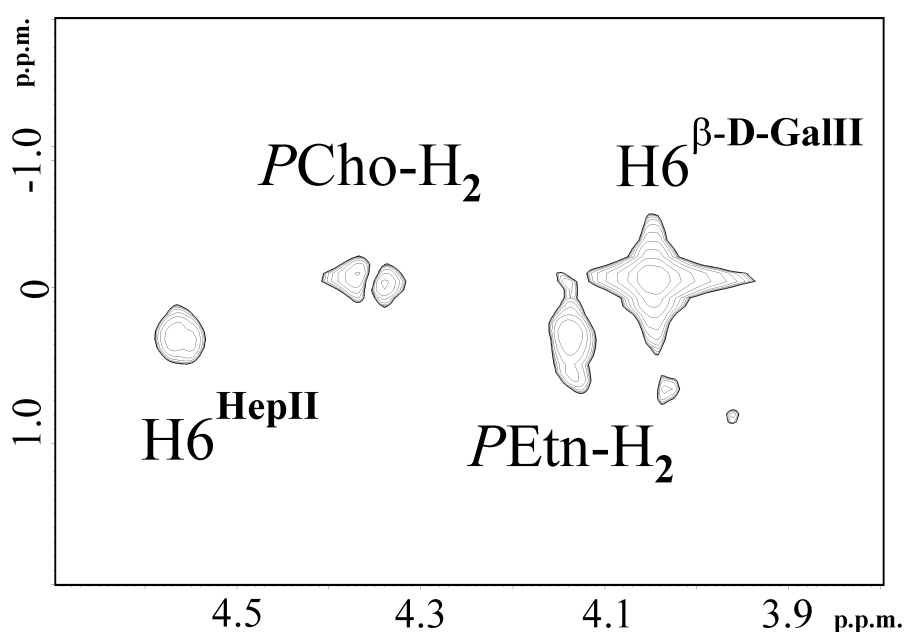
Fig. 20. The ^{1}H-^{31}P HMQC spectrum of deacylated OS-1.

ESI-MS experiments on OS samples indicated glycoforms that were substituted by ester-linked glycine, and acetate groups. Information on the location of acetate was provided by CE-ESI-MS^n on OS-1. One acetate group was found to be linked to β-D-Gal*p* attached to acetylated HepIII. Another acetate group is linked to D,D-Hep residue. It has been shown that O-acetyl groups help bacteria to increase their resistance to the killing effect of normal human sera (*78*).

Thus analysis of LPS from NT*Hi* strain 1232 allows us to propose the structure of the fully extended HexNAc$_1$Hex$_7$Hep$_4$ glycoform (Scheme 6).

α-D-Gal*p*-(1→4)-β-D-Gal*p*
1
↓
7
β-D-Glc*p*-(1→4)-D-α-D-HepIV*p*←Ac
1   *PP*Etn
↓   ↓
6   4
α-Neu5Ac-(2→3)-β-D-Gal*p*-(1→4)-β-D-GlcNAc*p*-(1→3)-β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→4)-L-α-D-HepI*p*-(1→5)-α-Kdo*p*-(2→6)-lipid A
3
↑
1
L-α-D-HepII*p* 6←*P*Etn
2
↑
1
*P*Cho→6 β-D-Gal*p*-(1→2)-L-α-D-HepIII*p*
↑   ↑
Ac   Ac
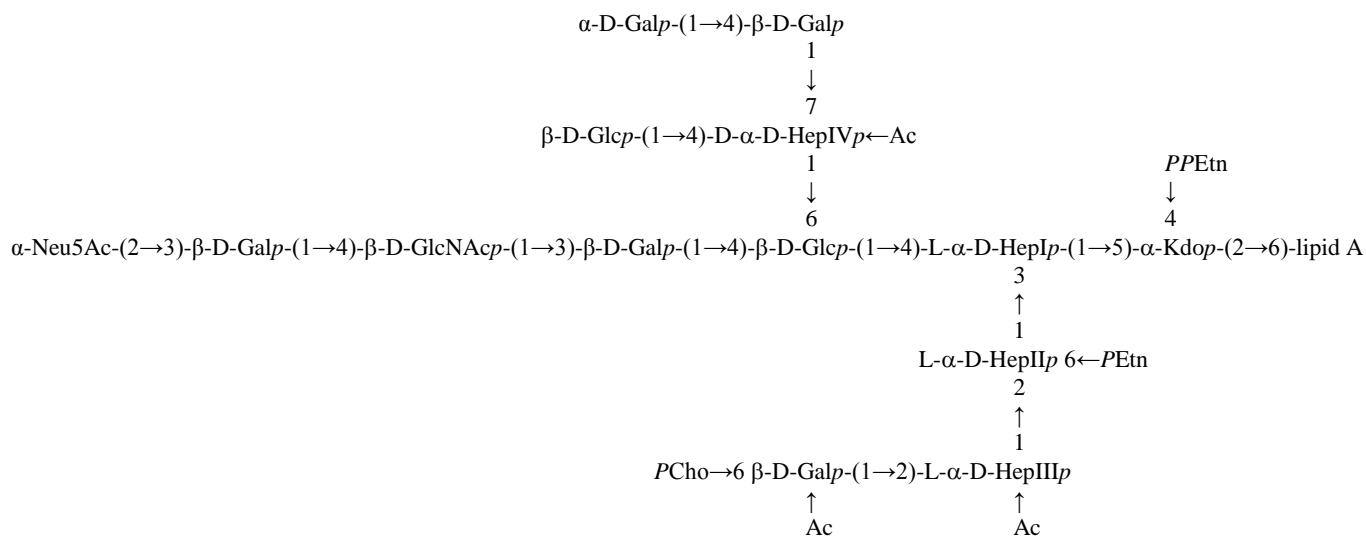
Scheme 6. The structure of the fully extended HexNAc$_1$Hex$_7$Hep$_4$ glycoform from NTHi strain 1232.

Structural studies of LPS from Haemophilus parainfluenzae.

SDS-PAGE analysis of LPS, which was done by our colleges in Oxford University, indicated that *H. parainfluenzae* strains 20 and 16 express O-antigen on their surfaces.

3.3   Paper III.

Structural studies of the lipopolysaccharide from *Haemophilus parainfluenzae* strain 20.

The lipopolysaccharide was isolated by hot phenol/water extraction followed with RNAse/DNAse and proteinase K treatment and then purified by dialysis.
OS material, obtained by the mild acid hydrolysis of LPS, after purification by gel filtration resulted in oligosaccharide samples OS-1 and OS-2. The ESI-MS spectrum of OS-2 showed one single charged ion corresponding to a phosphorylated trisaccharide (Fig. 21).
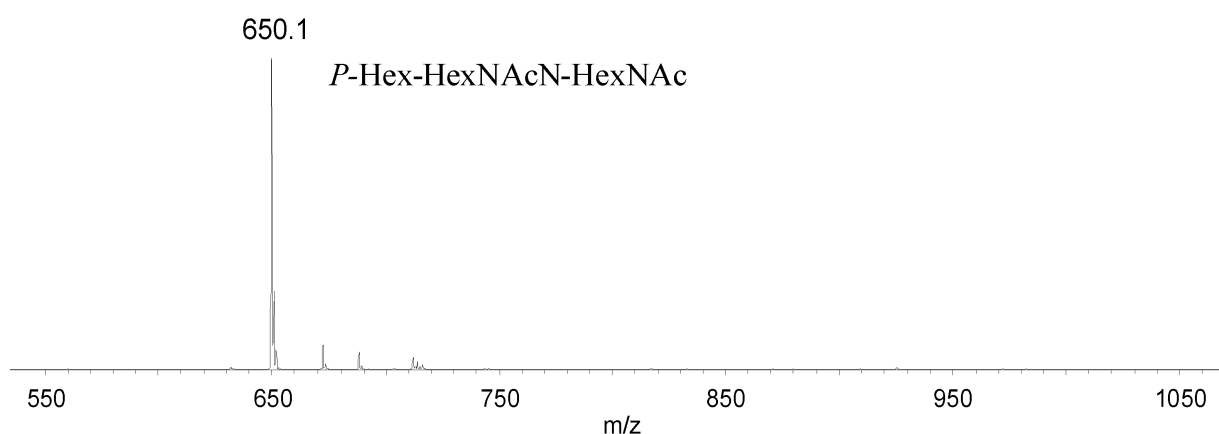
Fig. 21. The ESI-MS spectrum of OS-2 (positive mode).

The OS-2 fraction was subjected to NMR spectroscopy. The ^{1}H-^{13}C HMQC spectrum of OS-2 showed four anomeric signals, two of these signals belonged to the linked monosaccharides and two signals to α- and β-form of the monosaccharide at the reducing end (Fig. 22). The presence of the phosphate group was confirmed by ^{1}H-^{31}P HMQC experiment. Thus the OS-2 consists of trisaccharide with the structure:

P→6)-β-D-Glc*p*-(1→3)-α-D-Fuc*p*NAc4N-(1→4)-α,β-D-Gal*p*NAc-(1→

It was suggested that this structure represents the O-antigen repeated unit, which contains phosphate linkage that was degraded under mild acid hydrolysis.

32
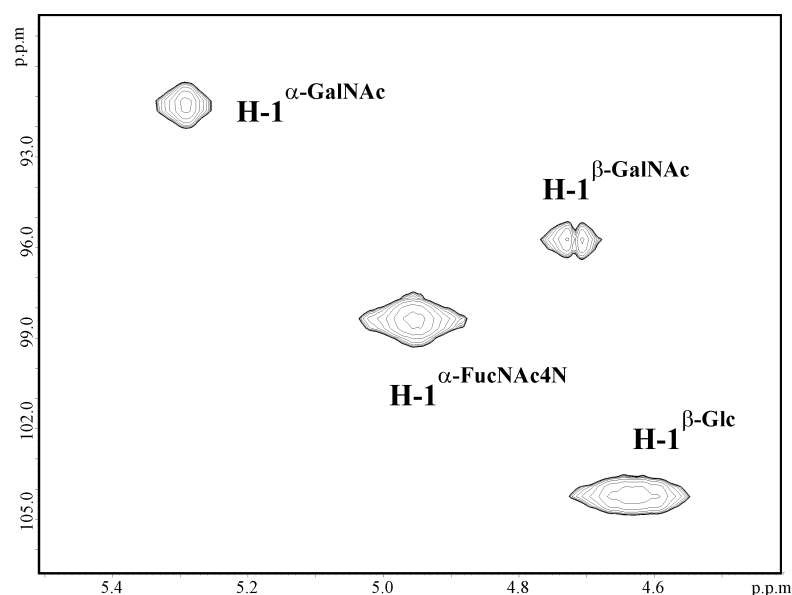
Fig. 22. Anomeric region of the ^1H- ^{13}C HMQC spectrum of OS-2

This was confirmed by 1D and 2D NMR spectroscopy on LPS. The 1D ^1H spectrum showed signals for three anomeric protons belonging to GalNAc, FucNAc4N and Glc (Fig. 23). Relatively small coupling constants have indicated that of Fuc*p*NAc4N and Gal*p*NAc residues have the α-anomeric configuration and a relatively large coupling constant have indicated Glc*p* residue as having the β-anomeric configuration. The phosphate group was found to be coupled to H-6,6′of β-Glc*p* and to the anomeric proton of α-Gal*p*NAc that also appears as a doublet of a doublet in 1D ^1H spectrum due to the coupling of H-1 to H-2 and H-1 to the phosphorus atom (Fig. 23A).
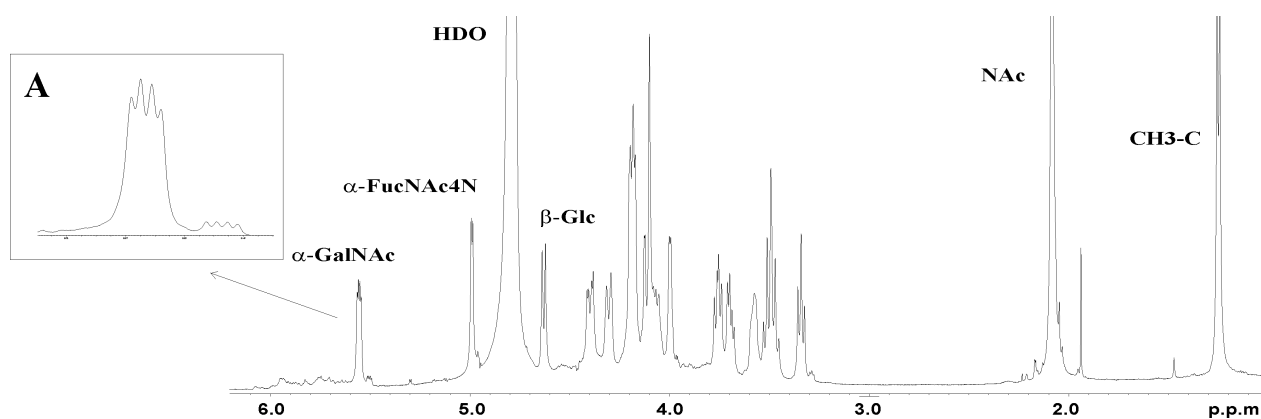
Fig. 23. The 1D ^1H spectrum of LPS in D_2O.

TOCSY and NOESY spectra were used to confirm *gluco*-configuration of Glc*p* and *galacto*-configuration of Gal*p*NAc and Fuc*p*NAc4N residues and the sequence was established by NOESY experiment (Fig. 24)
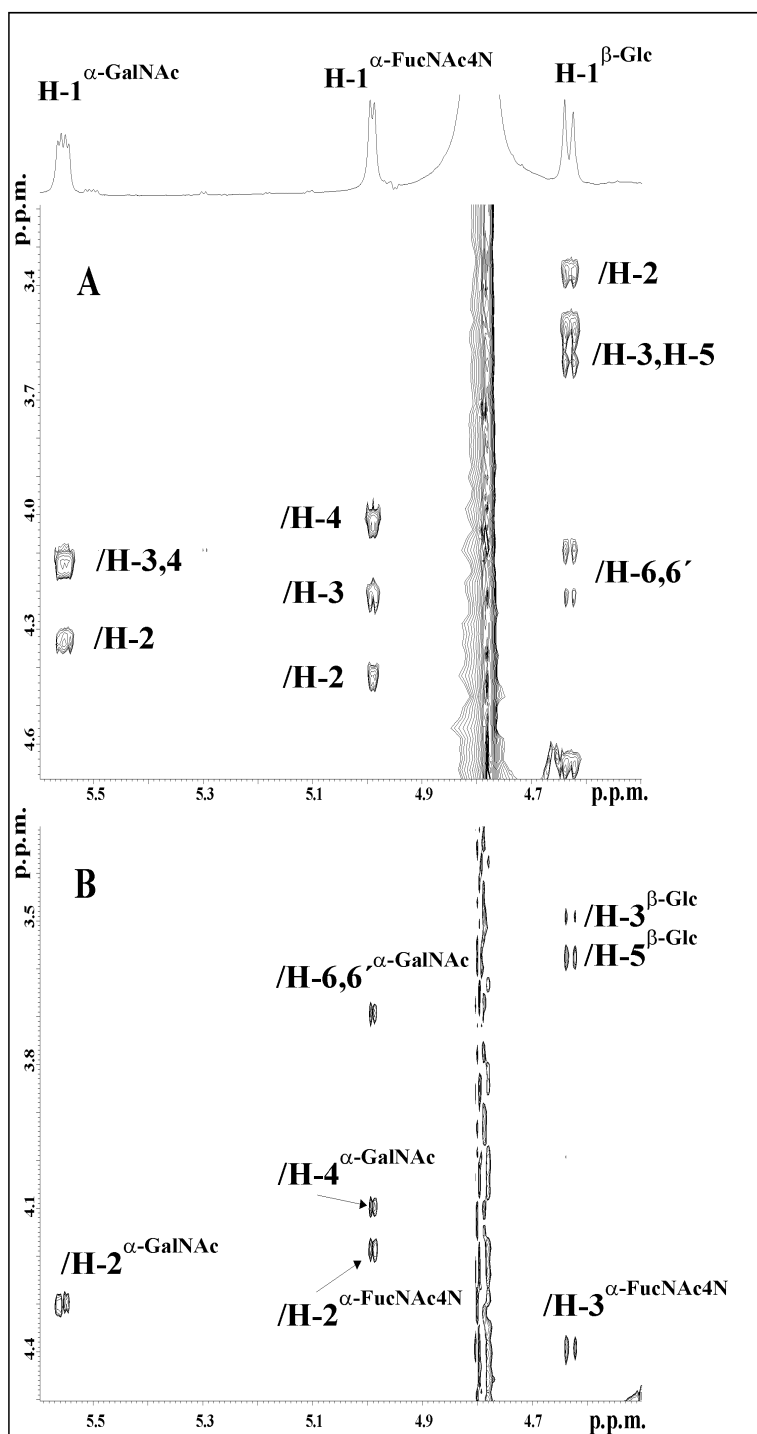
Fig. 24. Selected regions from 2D TOCSY (A) and NOESY (B) spectra of LPS.

NOE cross picks between the protons H-1 and H-2 indicated that FucpNAc4N and GalpNAc residues have α-amoneric configurations. NOE connectivities between the protons H-1 and H-3,5 indicated Glcp residue having the β-anomeric configuration.

The gluco-configuration of Glcp was confirmed by correlation of H-1 with H-2,3,4,5,6 in TOCSY spectrum and galacto-configuration of GalpNAc and FucpNAc4N residues by correlation of H-1 with H-2,3,4 in TOCSY spectrum and by NOE connectivities between the protons H-3, H-4 and H-5 (not shown).

The NOESY connectivities between the protons H-1 of α-FucpNAc4N and H-4/H-6,6´ of α-GalpNAc indicated that α-GalpNAc was substituted at O-4 position by α-FucpNAc4N, which was further substituted by β-Glcp residue at O-3 position (NOE connectivitise between the proton H-1 of β-Glcp and H-3 of α-FucpNAc4N) .

34

Absolute configurations of β-D-Glc and α-D-GalNAc residues were identified by GC-MS as the acetylated (R)-2-butyl ester. Absolute configuration of 2-acetamido-4-amino-2,4,6-tridioxygalactose cannot be obtained by this method due to full destruction of this acid-labile monosaccharide under hydrolysis (*117*). However analysis of effects of glycosylation on ^{13}C chemical shifts of β-D-Glc*p* allowed us to suggest the absolute configuration of 2-acetamido-4-amino-2,4,6-tridioxygalactose. A sufficiently large α-effect of +7,4p.p.m. on C1 of β-D-Glc*p* points to the D-configuration of α-Fuc*p*NAc4N residue (*118*).

α-Effects on C1 of β-D-Glc*p* from ^{13}C NMR simulation spectra (*119*)	
→ *P*→6)-β-D-Glc*p*-(1→3)-α-**D**-Fuc*p*NAc-(1→4)-α-D-Gal*p*NAc-(1→	+7.0
→ *P*→6)-β-D-Glc*p*-(1→3)-α-**L**-Fuc*p*NAc-(1→4)-α-D-Gal*p*NAc-(1→	+5.0

Based on these data, the following structure of the O-antigen repeating unit was concluded:

→4)-α-D-Gal*p*NAc-1→ *P*-(O→6)-β-D-Glc*p*-(1→3)-α-D-Fuc*p*NAc4N-(1→.
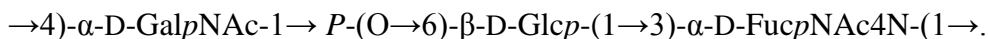
The O-antigen of *H. parainfluenzae* strain 20 contains the quite unusual sugar FucNAc4N or 2-acetamido-4-amino-2,4,6-tridioxy-D-galactose, that has been found mostly in the structures of teichoic and lipoteichoic acids of *Streptococcus pneumonia*. They have the following main structure →6)-β-D-Glc*p*-(1→3)-α-D-Fuc*p*NAc4N-(1→ 4)-α-D-Gal*p*NAc-(1→ 4)-β-D-Gal*p*NAc -(1→ 1)-D-ribitol-5-P-(O→ (*23*). This structure is substituted by one or two *P*Cho residues per repeating unit. However, *P*Cho was not found in the O-antigen structure from *H.parainfluenzae* strain 20.

ESI-MS on OS-1 indicated a low amount of glycoforms containing triheptosyl inner-core with compositions P_1•Hex$_2$•4NHexNAc$_1$•Hep$_3$•*P*Etn$_1$•*An*Kdo-*ol* and P_1•Hex$_2$•4NHexNAc$_1$•Hep$_3$•*P*Etn$_1$•*An*Kdo-*ol*.
In order to isolate these oligosaccharides, OS-1 was subjected to repetitive gel filtration and the resulting OS-1* fraction was analyzed by NMR spectroscopy. In this way it was confirmed that LPS from *H. parainfluenzae* and *H. influenzae* comprise the same *P*Etn-substituted triheptosyl inner-core moiety linked via the Kdo unit to the lipid A moiety (Scheme 7).

P→6)-β-D-Glc*p*-(1→3)-β-D-Fuc*p*NAc4N-(1→4)-β-D-Glc*p*-(1→4)-L-α-D-HepI*p*-(1→5)-*An*Kdo-*ol*
3
↑
1
L-α-D-HepII*p* 6←*P*Etn
2
↑
1
β-D-Gal*p*-(1→2)-L-α-D-HepIII*p*
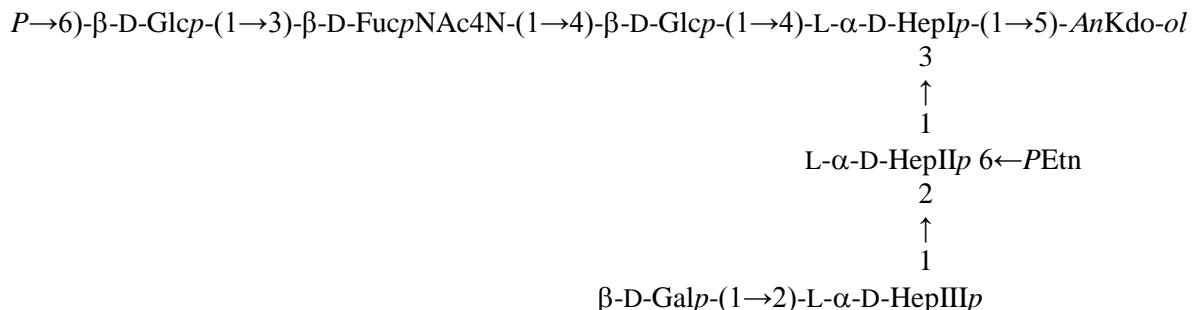
Scheme 7. Structure proposed for the P_1•Hex$_2$•4NHexNAc$_1$•Hep$_3$•PEtn$_2$•AnKdo-ol glycoform of H.parainfluenzae strain 20.

Interestingly, GlcI linked to HepI was substituted by a FucNAc4N residue having the β-configuration, however in the O-antigen it was in the α-configuration. This result allows us to conclude that FucNAc4N is the first monosaccharide in the O-antigen repeating unit.

It is known that O-units are synthesized *en bloc* in the cytoplasm linked to an undecaprenyl lipid carrier, which is then polymerized and added to the LPS as the O-antigen during the transport of LPS across the bacterial cell wall. It is also known that in other bacteria, WbaP is responsible for addition of the first sugar to the isoprenoid lipid carrier to which the remainder of the sugars of the repeating unit are condensed. The analysis of *H. parainfluenzae* mutant strain 20wbaP indicated that WbaP performs a similar function in *H. parainfluenzae* strain 20. When *wbaP* was inactivated, the mutant did not express the O-antigen, consistent with its predicted function.

3.4 *Paper IV.*

Structural studies of the O-antigen from *Haemophilus parainfluenzae* strain 16.

LPS from *H. parainfluenzae* strain *Hp*16 was extracted by hot phenol/water method and purified by dialysis. Due to its hydrophilic property, LPS could be analyzed by NMR spectroscopy without prior derivatization.

The 1D C^{13} NMR spectum showed 12 signals, two of them at δ_C 107.5 p.p.m and at δ_C 104.8 p.p.m. were observed in the anomeric region (Fig. 26A). The 1D H^1 spectrum showed only one anomeric protons at δ_H 5.18 p.p.m. (Fig. 25B).
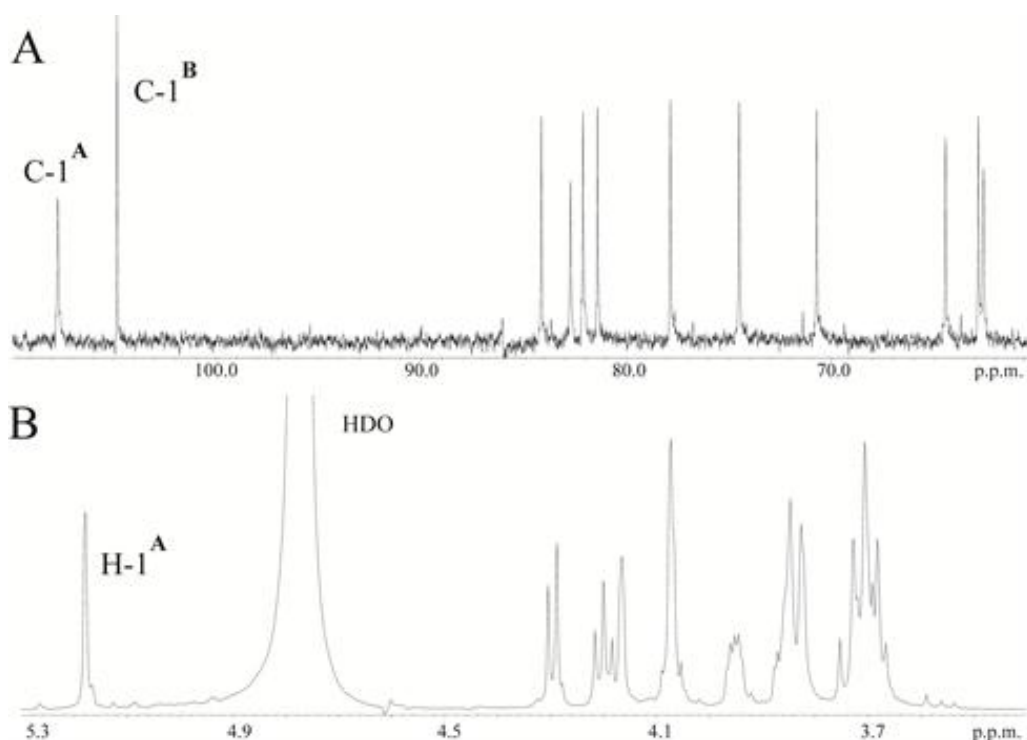
Fig. 25. 1D ^{13}C (A) and 1D ^{1}H (B) spectra of LPS from Hp 16.

According to this result it was suggested that the O-antigen of *H. parainfluenzae* 16 consists of a repeating disaccharide containing one ketose. Since sugar analysis showed the presence of glucose, galactose and mannose, it was concluded that this keto sugar is fructose, given that reduction of fructose gives mannitol and glucitol (Fig. 26).
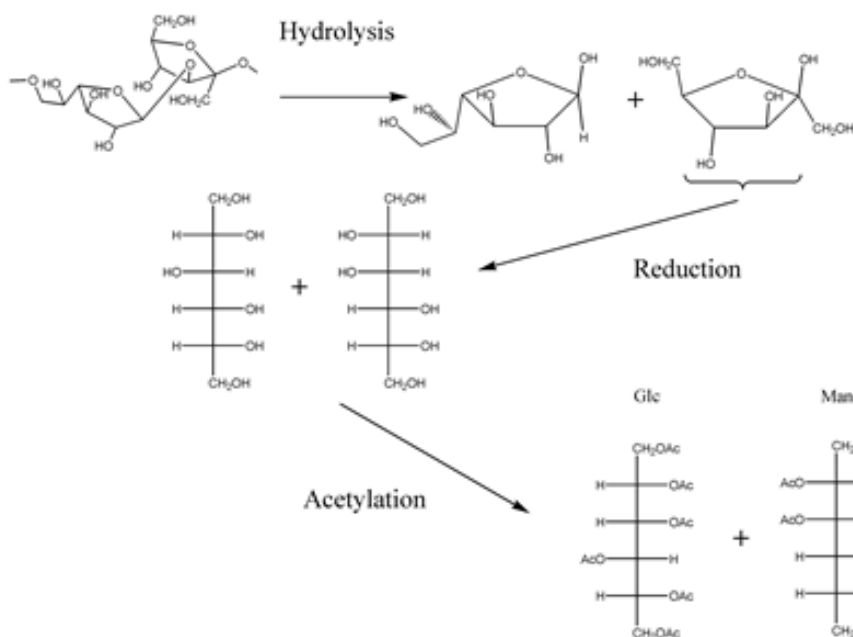
*Fig. 26. Sugar analysis of fructose residue from the O-antigen of Hp16.*

Analysis of 2D TOCSY and COSY spectra, which clearly showed the presence of two spin systems, confirmed that the O-antigen from *H. parainfluenzae* strain 16 is built on two sugar residues (Fig. 27).
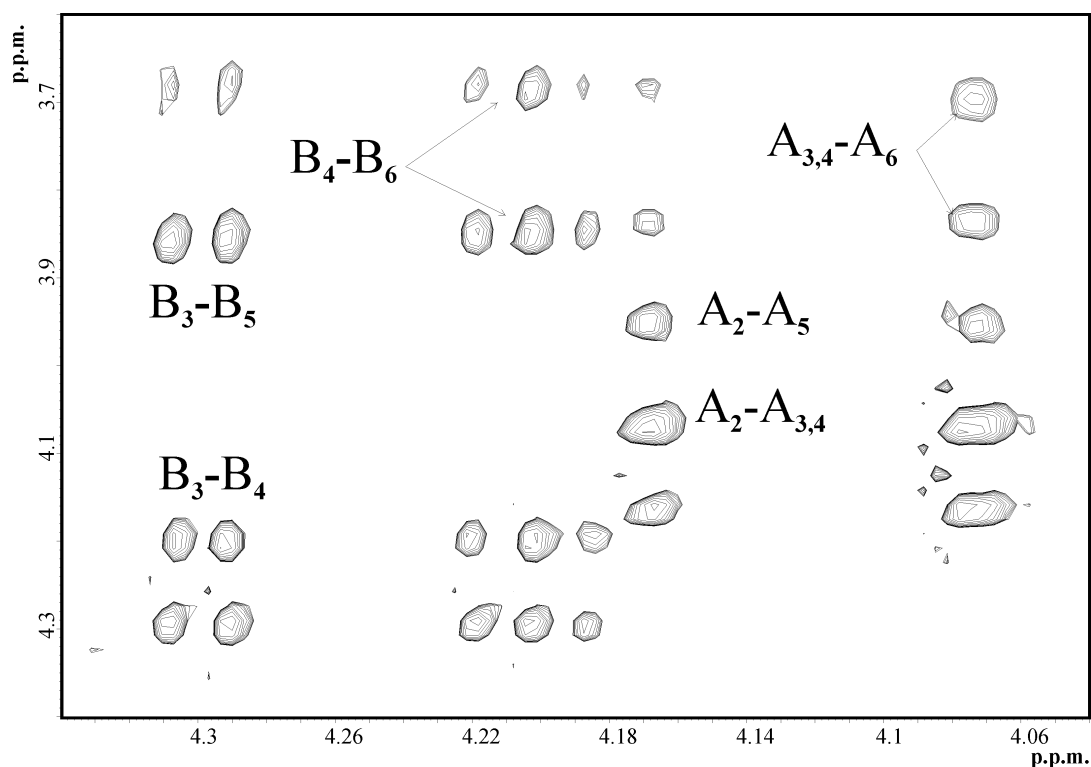
Fig. 27. Selected region of the 2D ^1H-^1H TOCSY spectrum of LPS from Hp16.
A→ β-D-Galf
B→ β-D-Fruf.

Carbon chemical shifts, obtained by an HMQC experiment, indicated that both sugar residues are furanosidic and have β- anomeric configuration. 2D COSY, TOCSY, NOESY, HMQC and HMBC experiments indicated that the anomeric signal at δ_C 107.5 p.p.m. belongs to β-D-Gal$f$ and the signal at δ_C 105 p.p.m. to β-D-Fru$f$.

The sequence analysis was achieved by HMBC experiment (Fig. 28) and was confirmed by a NOESY spectrum, which showed a cross-peak between H-1 of β-D-Gal$f$ and H-3 of β-D-Fru$f$.
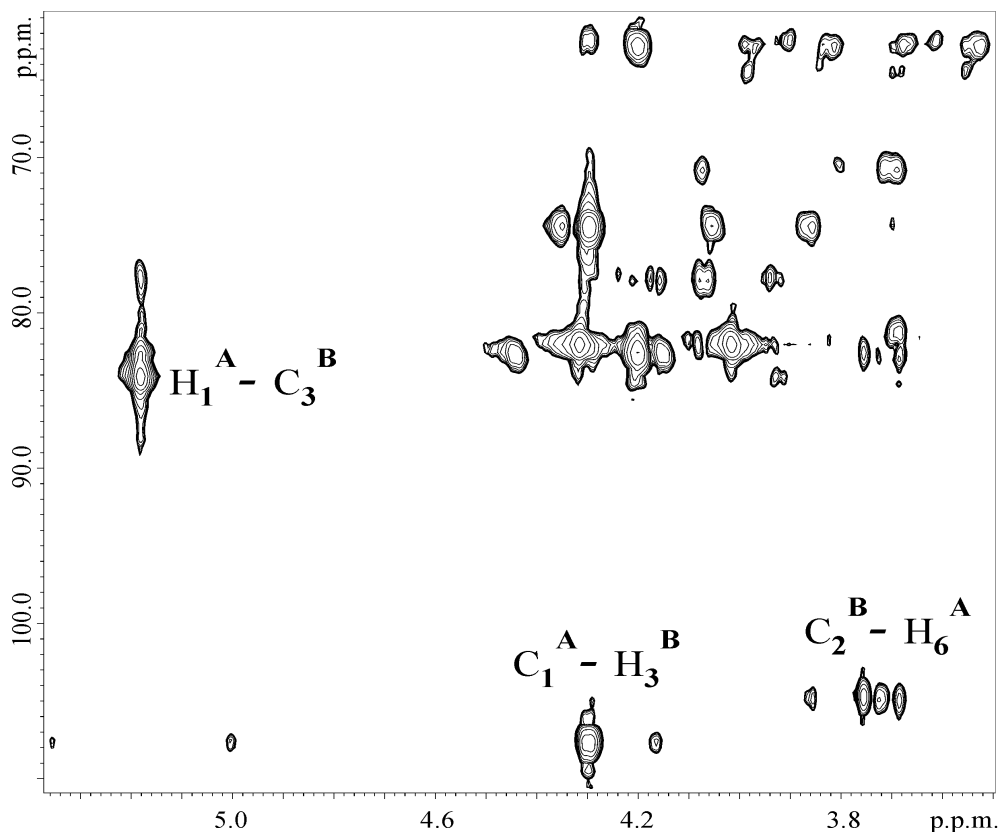
Fig. 28. The ^{1}H-^{13}C HMBC spectrum of LPS from H. parainfluenzae strain 16.

The cross-peaks between H-1 of A and C-3 of B, as well as C-1 of A and H-3 of B indicated that B was substituted by A at O-3 position. Cross-peaks between C-2 of B and H-6 of A indicated that A was substituted by B at O-6 position. Thus the structure was established as :

→6)-β-D-Gal$f$ –(1→3)-β-D-Fru$f$ –(2→
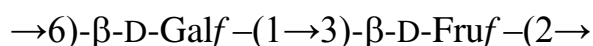
A B

4 SUMMARY AND CONCLUSIONS

The structural studies of NT*Hi* strains 1158, 1159 and 1232 extend our knowledge of structural motives extending from HepI. NT*Hi* strains 1158, 1159 and 1232 express the D,D-Hep residue in the outer core region.
Elucidation of LPS structures of NT*Hi* strain 1232 revealed di-substituted D,D-Hep.

α-D-Gal*p*-(1→4)-β-D-Gal*p*

1

↓

7

β-D-Glc*p*-(1→4)-D-α-D-HepIV*p*-(1→6)-β-D-GlcI*p*-(1→4)-L-α-D-HepI*p*-(1→

↓
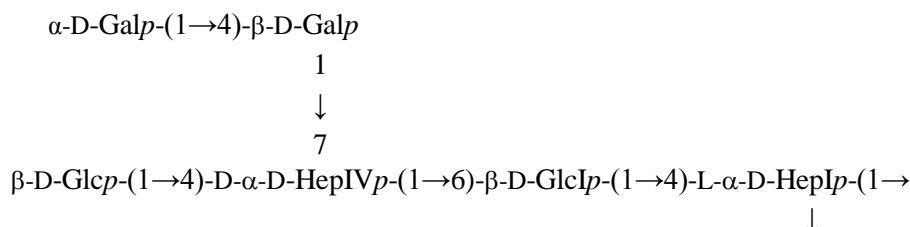
Results from this study together with previously published data have indicated that the D,D-Hep residue can be substituted not only at the O-4 position by β-D-Glc*p* (NT*Hi* strains: 1232, 2846 (*44*)) or β-D-Gal*p* (NT*Hi* strain 981(*35*)), but can also be substituted at O-7 position by β-D-Gal*p*, which can be further elongated by α-D-Gal*p* at O-4 position.
Hence a new site has been found for the expression of the digalactoside [α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→], an epitope known to increase resistance to killing by naturally acquired antibody and complement present in human serum (*57*).

Two gene pairs: *losB1/losA1* and *losB2 /losA2* are responsible for addition and substitution of D,D-Hep and L,D-Hep, respectively (*44*).
NT*Hi* strain 1232 like strain 2846, contains both the *losB1/losA1* and *losB2 /losB2* gene pairs (*116*). It has been shown that *losA1* is needed for elongation from D,D-Hep by β-D-Glc*p* at O-4 position.

Interestingly, NT*Hi* strains 1158/1159 do not have the *losB1/losA1* gene pair, only *losB2/losA2*, but still express D,D-Hep. We could confirm that instead of *losB1*, *losB2* is responsible for addition of the D,D-Hep residue in strain 1158. Sequence analysis of *losB2* would suggest that the function of 1158LosB2 more closely related to function of LosB1 in the other NT*Hi* strains (*116*). The O-7 position of the D,D-Hep residue can be also occupied by *P*Cho in this strain. This novel site for *P*Cho was found by structural analysis of 1158*lpsA* mutant strain.

The structural studies of NT*Hi* strains 1158/1159 and mutant strain 1158*losB2* have indicated that the chain elongation from HepII is only possible in glycoforms lacking the external heptose. This suggests that the D,D-Hep residue sterically hinders chain extension from HepII and vice versa.

The structural studies on LPS from *H. parainfluenzae* complement our investigations about the role of LPS of *H. influenzae* in diseases. It is known that *H. parainfluenzae* is the part of normal flora and causes diseases only very rarely. On the other hand the lipopolysaccharide from *H. parainfluenzae* comprises the same inner-core structural element as *H. influenzae*, consisting of the *P*Etn-substituted triheptosyl unit linked via the Kdo unit to the lipid A moiety (*33, 82*).

As it was shown earlier, some strains of *H. parainfluenzae* express O-antigen containing LPS on its surfaces (*81*). *H. parainfluenzae* strains 20 and 16 are among these strains. This study was the first to characterize the O-antigen structure of smooth type LPS from *H. parainfluenzae*.

The O-antigen from *Hp*20 has the following structure:

→*P*-(O→6)-β-D-Glc*p*-(1→3)-α-D-Fuc*p*NAc4N-(1→4)-α-D-Gal*p*NAc-(1→
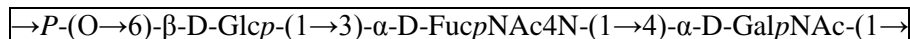
Interestingly, this structure is closely related to that of teichoic and lipoteichoic acids of *Streptococcus pneumoniae*:

→*P*→6)-β-D-Glc*p*-(1→3)-α-D-Fuc*p*NAc4N-(1→ 4)-α-D-Gal*p*NAc-(1→ 4)-β-D-Gal*p*NAc -(1→

1)-D-ribitol-5→
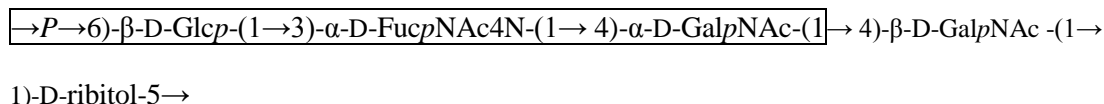
In S. *pneumonia* this structure is further substituted by one or two *P*Cho residues per repeating unit. However, *P*Cho was not found in the O-antigen structure of *Hp*20.

The O-specific polysaccharide of *H. parainfluenzae* strain 16 has the following structure:

→6)-β-D-Gal*f* –(1→3)-β-D-Fru*f* –(2→.
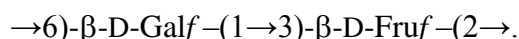
It is noteworthy, that fructofuranosyl unit is very rarely present in bacterial polysaccharides.

From a comparison of LPS structures from the NT*Hi* and *H. parainfluenzae* strains investigated in this thesis, it can be said that known outer-core virulence factors present in *H. influenzae* are absent in *H. parainfluenzae*. Instead in the investigated strains reported here, *H. parainfluenzae* express unique O-polysaccharide antigens which may play a role in colonization or persistence of the organism in the nasopharynx.

Previous studies of LPS from rough type *H. parainfluenzae* strains also show no evidence of virulence factors in their outer-core structures.
Future studies would include inserting known LPS virulence genes from *H. influenzae* into *H. parainflluenzae*. Since the triheptosyl inner-core region of the LPS are identical, it would be possible to add genes responsible for addition of globoside, phosphocholine or sialyllactose, epitopes known to promote virulence by NT*Hi*. The expected LPS structures in these genetically modified *H.parainfluenzae* strain would be confirmed by methods outlined in this thesis followed by studies to determine whether the modified organisms are pathogenic.

5 ACKNOWLEDGEMENTS

6 REFERENCES

1. Hamed, K. A., Dormitzer, P. R., Su, C. K., and Relman, D. A. (1994) *Haemophilus parainfluenzae* endocarditis: application of a molecular approach for identification of pathogenic bacterial species, *Clin Infect Dis 19*, 677-683.
2. Pillai, A., Mitchell, J. L., Hill, S. L., and Stockley, R. A. (2000) A case of *Haemophilus parainfluenzae* pneumonia, *Thorax 55*, 623-624.
3. Betriu, C., Coronel, F., Martin, P., and Picazo, J. J. (1999) *J Clin Microbiol 37*, 3074-3075.
4. Frankard, J., Rodriguez-Villalobos, H., Struelens, M. J., and Jacobs, F. (2004) *Eur J Clin Microbiol Infect Dis 23* . 46-48.
5. Murphy, T. F., Faden, H., Bakaletz, L. O., Kyd, J. M., Forsgren, A., Campos, J., Virji, M., and Pelton, S. I. (2009) Nontypeable *Haemophilus influenza*e as a pathogen in children, *Pediatr Infect Dis J 28*, 43-48.
6. Murphy, T. F. (2003) Respiratory infections caused by non-typeable *Haemophilus influenzae*, *Curr Opin Infect Dis 16*, 129-134.
7. Foxwell, A. R., Kyd, J. M., and Cripps, A. W. (1998) Nontypeable *Haemophilus influenzae*: pathogenesis and prevention, *Microbiol Mol Biol Rev 62*, 294-308.
8. Silhavy, T. J., Kahne, D., and Walker, S. (2010) The bacterial cell envelope, *Cold Spring Harbor perspectives in biology 2*, a000414.
9. Tomasz, A. (1971) The bacterial cell surface, *Nature 234*, 389-392.
10. Hancock, I. C. (1997) Bacterial cell surface carbohydrates: structure and assembly, *Biochem Soc Trans 25*, 183-187.
11. Bayer, M. E. (1974) Ultrastructure and organization of the bacterial envelope, *Annals of the New York Academy of Sciences 235*, 6-28.
12. Brandley, B. K., and Schnaar, R. L. (1986) Cell-surface carbohydrates in cell recognition and response, *Journal of leukocyte biology 40*, 97-111.
13. Jann, K., and Jann, B. (1978) Bacterial polysaccharides, *Methods Enzymol 50*, 251-272.
14. Costerton, J. W., Ingram, J. M., and Cheng, K. J. (1974) Structure and function of the cell envelope of gram-negative bacteria, *Bacteriological reviews 38*, 87-110.
15. Rietschel, E. T., Kirikae, T., Schade, F. U., Mamat, U., Schmidt, G., Loppnow, H., Ulmer, A. J., Zahringer, U., Seydel, U., Di Padova, F., and et al. (1994) Bacterial endotoxin: molecular relationships of structure to activity and function, *Faseb J 8*, 217-225.
16. Schletter, J., Heine, H., Ulmer, A. J., and Rietschel, E. T. (1995) Molecular mechanisms of endotoxin activity, *Archives of microbiology 164*, 383-389.
17. Alexander, C., and Rietschel, E. T. (2001) Bacterial lipopolysaccharides and innate immunity, *J Endotoxin Res 7*, 167-202.
18. Caroff, M., and Karibian, D. (2003) Structure of bacterial lipopolysaccharides, *Carbohydr Res 338*, 2431-2447.
19. Hitchcock, P. J., Leive, L., Makela, P. H., Rietschel, E. T., Strittmatter, W., and Morrison, D. C. (1986) Lipopolysaccharide nomenclature--past, present, and future, *J Bacteriol 166*, 699-705.
20. Galanos, C., Luderitz, O., Rietschel, E. T., Westphal, O., Brade, H., Brade, L., Freudenberg, M., Schade, U., Imoto, M., Yoshimura, H., and et al. (1985) Synthetic and natural *Escherichia coli* free lipid A express identical endotoxic activities, *Eur J Biochem 148*, 1-5.
21. Freudenberg, M. A., and Galanos, C. (1990) Bacterial lipopolysaccharides: structure, metabolism and mechanisms of action, *International reviews of immunology 6*, 207-221.
22. Hoffman, W. D., and Natanson, C. (1993) Endotoxin in septic shock, *Anesthesia and analgesia 77*, 613-624.
23. Fischer, W., Behr, T., Hartmann, R., Peter-Katalinic, J., and Egge, H. (1993) Teichoic acid and lipoteichoic acid of *Streptococcus pneumoniae* possess

identical chain structures. A reinvestigation of teichoid acid (C polysaccharide), *Eur J Biochem 215*, 851-857.

24.   Karlsson, C., Jansson, P. E., and Skov Sorensen, U. B. (1999) The pneumococcal common antigen C-polysaccharide occurs in different forms. Mono-substituted or di-substituted with phosphocholine, *Eur J Biochem 265*, 1091-1097.

25.   Fischer, W. (2000) Phosphocholine of pneumococcal teichoic acids: role in bacterial physiology and pneumococcal infection, *Research in microbiology 151*, 421-427.

26.   Kaplan, M. H., and Volanakis, J. E. (1974) Interaction of C-reactive protein complexes with the complement system. I. Consumption of human complement associated with the reaction of C-reactive protein with pneumococcal C-polysaccharide and with the choline phosphatides, lecithin and sphingomyelin, *J Immunol 112*, 2135-2147.

27.   Moxon, E. R., and Rappuoli, R. (1990) *Haemophilus influenzae* infections and whooping cough, *Lancet 335*, 1324-1329.

28.   Cody, A. J., Field, D., Feil, E. J., Stringer, S., Deadman, M. E., Tsolaki, A. G., Gratz, B., Bouchet, V., Goldstein, R., Hood, D. W., and Moxon, E. R. (2003) High rates of recombination in otitis media isolates of non-typeable *Haemophilus influenzae*, *Infect Genet Evol 3*, 57-66.

29.   Musser, J. M., Barenkamp, S. J., Granoff, D. M., and Selander, R. K. (1986) Genetic relationships of serologically nontypable and serotype b strains of *Haemophilus influenzae*, *Infect Immun 52*, 183-191.

30.   Porras, O., Caugant, D. A., Gray, B., Lagergard, T., Levin, B. R., and Svanborg-Eden, C. (1986) Difference in structure between type b and nontypable *Haemophilus influenzae* populations, *Infect Immun 53*, 79-89.

31.   Smith-Vaughan, H. C., Sriprakash, K. S., Leach, A. J., Mathews, J. D., and Kemp, D. J. (1998) Low genetic diversity of *Haemophilus influenzae* type b compared to nonencapsulated *H. influenzae* in a population in which H. influenzae is highly endemic, *Infect Immun 66*, 3403-3409.

32.   Hood, D. W., Richards, J. C., and Moxon, E. R. (1999) *Haemophilus influenzae* lipopolysaccharide, *Biochem Soc Trans 27*, 493-498.

33.   Schweda, E. K., Richards, J. C., Hood, D. W., and Moxon, E. R. (2007) Expression and structural diversity of the lipopolysaccharide of *Haemophilus influenzae*: implication in virulence, *Int J Med Microbiol 297*, 297-306.

34.   Li, J., Bauer, S. H., Mansson, M., Moxon, E. R., Richards, J. C., and Schweda, E. K. (2001) Glycine is a common substituent of the inner core in *Haemophilus influenzae* lipopolysaccharide, *Glycobiology 11*, 1009-1015.

35.   Mansson, M., Hood, D. W., Moxon, E. R., and Schweda, E. K. (2003) Structural characterization of a novel branching pattern in the lipopolysaccharide from nontypeable *Haemophilus influenzae*, *Eur J Biochem 270*, 2979-2991.

36.   Yildirim, H. H., Li, J., Richards, J. C., Hood, D. W., Moxon, E. R., and Schweda, E. K. (2005) Complex O-acetylation in non-typeable *Haemophilus influenzae* lipopolysaccharide: evidence for a novel site of O-acetylation, *Carbohydr Res 340*, 2598-2611.

37.   Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J. F., Dougherty, B. A., Merrick, J. M., and et al. (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd, *Science 269*, 496-512.

38.   Erwin, A. L., Nelson, K. L., Mhlanga-Mutangadura, T., Bonthuis, P. J., Geelhood, J. L., Morlin, G., Unrath, W. C., Campos, J., Crook, D. W., Farley, M. M., Henderson, F. W., Jacobs, R. F., Muhlemann, K., Satola, S. W., van Alphen, L., Golomb, M., and Smith, A. L. (2005) Characterization of genetic and phenotypic diversity of invasive nontypeable *Haemophilus influenzae*, *Infect Immun 73*, 5853-5863.

39.   Harrison, A., Dyer, D. W., Gillaspy, A., Ray, W. C., Mungur, R., Carson, M. B., Zhong, H., Gipson, J., Gipson, M., Johnson, L. S., Lewis, L., Bakaletz, L. O., and Munson, R. S., Jr. (2005) Genomic sequence of an otitis media isolate

of nontypeable *Haemophilus influenzae*: comparative study with *H. influenzae* serotype d, strain KW20, *J Bacteriol 187*, 4627-4636.

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

41. 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., and Richards, J. C. (2001) Genetic basis for expression of the major globotetraose-containing lipopolysaccharide from *H. influenzae* strain Rd (RM118), *Glycobiology 11*, 957-967.

42. Hood, D. W., Deadman, M. E., Cox, A. D., Makepeace, K., Martin, A., Richards, J. C., and Moxon, E. R. (2004) 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 150*, 2089-2097.

43. Wright, J. C., Hood, D. W., Randle, G. A., Makepeace, K., Cox, A. D., Li, J., Chalmers, R., Richards, J. C., and Moxon, E. R. (2004) *lpt6*, a gene required for addition of phosphoethanolamine to inner-core lipopolysaccharide of *Neisseria meningitidis* and *Haemophilus influenzae, J Bacteriol 186*, 6970-6982.

44. Lundstrom, S. L., Li, J., Deadman, M. E., Hood, D. W., Moxon, E. R., and Schweda, E. K. (2008) Structural analysis of the lipopolysaccharide from nontypeable *Haemophilus influenzae* strain R2846, *Biochemistry 47*, 6025-6038.

45. Twelkmeyer, B., Deadman, M. E., Haque, E., Li, J., Hood, D. W., and Schweda, E. K. (2011) The role of *lic2B* in lipopolysaccharide biosynthesis in *Haemophilus influenzae* strain Eagan, *Carbohydr Res 346*, 1262-1266.

46. Kimura, A., and Hansen, E. J. (1986) Antigenic and phenotypic variations of *Haemophilus influenzae* type b lipopolysaccharide and their relationship to virulence, *Infect Immun 51*, 69-79.

47. Peak, I. R., Jennings, M. P., Hood, D. W., and Moxon, E. R. (1999) Tetranucleotide repeats identify novel virulence determinant homologues in *Neisseria meningitidis*, *Microb Pathog 26*, 13-23.

48. Bayliss, C. D., Field, D., and Moxon, E. R. (2001) The simple sequence contingency loci of *Haemophilus influenzae* and *Neisseria meningitidis*, *J Clin Invest 107*, 657-662.

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

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

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

52. Moxon, R., Bayliss, C., and Hood, D. (2006) Bacterial contingency loci: the role of simple sequence DNA repeats in bacterial adaptation, *Annual review of genetics 40*, 307-333.

53. Gilsdorf, J. R. (1998) Antigenic diversity and gene polymorphisms in *Haemophilus influenzae*, *Infect Immun 66*, 5053-5059.

54. Griffin, R., Bayliss, C. D., Herbert, M. A., Cox, A. D., Makepeace, K., Richards, J. C., Hood, D. W., and Moxon, E. R. (2005) Digalactoside expression in the lipopolysaccharide of *Haemophilus influenzae* and its role in intravascular survival, *Infect Immun 73*, 7022-7026.

55. Deadman, M. E., Hermant, P., Engskog, M., Makepeace, K., Moxon, E. R., Schweda, E. K., and Hood, D. W. (2009) Lex2B, a phase-variable glycosyltransferase, adds either a glucose or a galactose to *Haemophilus influenzae* lipopolysaccharide, *Infect Immun 77*, 2376-2384.

56. Engskog, M. K., Yildirim, H. H., Li, J., Richards, J. C., Deadman, M., Hood, D. W., and Schweda, E. K. (2009) A dual role for the *lex2* locus: identification of galactosyltransferase activity in non-typeable *Haemophilus influenzae* strains 1124 and 2019, *Carbohydr Res 344*, 632-641.

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

58. Moran, A. P., Prendergast, M. M., and Appelmelk, B. J. (1996) Molecular mimicry of host structures by bacterial lipopolysaccharides and its contribution to disease, *FEMS Immunol Med Microbiol 16*, 105-115.

59. Hood, D. W., Cox, A. D., Gilbert, M., Makepeace, K., Walsh, S., Deadman, M. E., Cody, A., Martin, A., Mansson, M., Schweda, E. K., Brisson, J. R., Richards, J. C., Moxon, E. R., and Wakarchuk, W. W. (2001) Identification of a lipopolysaccharide alpha-2,3-sialyltransferase from *Haemophilus influenzae, Mol Microbiol 39*, 341-350.

60. Fox, K. L., Cox, A. D., Gilbert, M., Wakarchuk, W. W., Li, J., Makepeace, K., Richards, J. C., Moxon, E. R., and Hood, D. W. (2006) Identification of a bifunctional lipopolysaccharide sialyltransferase in *Haemophilus influenzae*: incorporation of disialic acid, *J Biol Chem 281*, 40024-40032.

61. Houliston, R. S., Koga, M., Li, J., Jarrell, H. C., Richards, J. C., Vitiazeva, V., Schweda, E. K., Yuki, N., and Gilbert, M. (2007) A *Haemophilus influenzae* strain associated with Fisher syndrome expresses a novel disialylated ganglioside mimic, *Biochemistry 46*, 8164-8171.

62. Yildirim, H. H., Li, J., Richards, J. C., Hood, D. W., Moxon, E. R., and Schweda, E. K. (2005) An alternate pattern for globoside oligosaccharide expression in *Haemophilus influenzae* lipopolysaccharide: structural diversity in nontypeable strain 1124, *Biochemistry 44*, 5207-5224.

63. Lundstrom, S. L., Twelkmeyer, B., Sagemark, M. K., Li, J., Richards, J. C., Hood, D. W., Moxon, E. R., and Schweda, E. K. (2007) Novel globoside-like oligosaccharide expression patterns in nontypeable *Haemophilus influenzae* lipopolysaccharide, *Febs J 274*, 4886-4903.

64. Cox, A. D., Hood, D. W., Martin, A., Makepeace, K. M., Deadman, M. E., Li, J., Brisson, J. R., Moxon, E. R., and Richards, J. C. (2002) Identification and structural characterization of a sialylated lacto-*N*-neotetraose structure in the lipopolysaccharide of *Haemophilus influenzae*, *Eur J Biochem 269*, 4009-4019.

65. Hood, D. W., Randle, G., Cox, A. D., Makepeace, K., Li, J., Schweda, E. K., Richards, J. C., and Moxon, E. R. (2004) Biosynthesis of cryptic lipopolysaccharide glycoforms in *Haemophilus influenzae* involves a mechanism similar to that required for O-antigen synthesis, *J Bacteriol 186*, 7429-7439.

66. Jones, P. A., Samuels, N. M., Phillips, N. J., Munson, R. S., Jr., Bozue, J. A., Arseneau, J. A., Nichols, W. A., Zaleski, A., Gibson, B. W., and Apicella, M. A. (2002) *Haemophilus influenzae* type b strain A2 has multiple sialyltransferases involved in lipooligosaccharide sialylation, *J Biol Chem 277*, 14598-14611.

67. Bauer, S. H., Mansson, M., Hood, D. W., Richards, J. C., Moxon, E. R., and Schweda, E. K. (2001) 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 335*, 251-260.

68. Bouchet, V., Hood, D. W., Li, J., Brisson, J. R., Randle, G. A., Martin, A., Li, Z., Goldstein, R., Schweda, E. K., Pelton, S. I., Richards, J. C., and Moxon, E. R. (2003) Host-derived sialic acid is incorporated into *Haemophilus influenzae* lipopolysaccharide and is a major virulence factor in experimental otitis media, *Proc Natl Acad Sci U S A 100*, 8898-8903.

69. Hood, D. W., Makepeace, K., Deadman, M. E., Rest, R. F., Thibault, P., Martin, A., Richards, J. C., and Moxon, E. R. (1999) Sialic acid in the lipopolysaccharide of *Haemophilus influenzae*: strain distribution, influence on serum resistance and structural characterization, *Mol Microbiol 33*, 679-692.

70. Lundstrom, S. L., Li, J., Mansson, M., Figueira, M., Leroy, M., Goldstein, R., Hood, D. W., Moxon, E. R., Richards, J. C., and Schweda, E. K. (2008)

Application of capillary electrophoresis mass spectrometry and liquid chromatography multiple-step tandem electrospray mass spectrometry to profile glycoform expression during *Haemophilus influenzae* pathogenesis in the chinchilla model of experimental otitis media, *Infect Immun 76*, 3255-3267.

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

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

73. Fox, K. L., Li, J., Schweda, E. K., Vitiazeva, V., Makepeace, K., Jennings, M. P., Moxon, E. R., and Hood, D. W. (2008) Duplicate copies of *lic1* direct the addition of multiple phosphocholine residues in the lipopolysaccharide of *Haemophilus influenzae*, *Infect Immun 76*, 588-600.

74. Weiser, J. N., Pan, N., McGowan, K. L., Musher, D., Martin, A., and Richards, J. (1998) 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 187*, 631-640.

75. Swords, W. E., Buscher, B. A., Ver Steeg Ii, K., Preston, A., Nichols, W. A., Weiser, J. N., Gibson, B. W., and Apicella, M. A. (2000) Non-typeable *Haemophilus influenzae* adhere to and invade human bronchial epithelial cells via an interaction of lipooligosaccharide with the PAF receptor, *Mol Microbiol 37*, 13-27.

76. Tong, H. H., Blue, L. E., James, M. A., Chen, Y. P., and DeMaria, T. F. (2000) Evaluation of phase variation of nontypeable *Haemophilus influenzae* lipooligosaccharide during nasopharyngeal colonization and development of otitis media in the chinchilla model, *Infect Immun 68*, 4593-4597.

77. Lysenko, E. S., Gould, J., Bals, R., Wilson, J. M., and Weiser, J. N. (2000) Bacterial phosphorylcholine decreases susceptibility to the antimicrobial peptide LL-37/hCAP18 expressed in the upper respiratory tract, *Infect Immun 68*, 1664-1671.

78. Fox, K. L., Yildirim, H. H., Deadman, M. E., Schweda, E. K., Moxon, E. R., and Hood, D. W. (2005) Novel lipopolysaccharide biosynthetic genes containing tetranucleotide repeats in *Haemophilus influenzae*, identification of a gene for adding O-acetyl groups, *Mol Microbiol 58*, 207-216.

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

80. Mikhail, I., Yildirim, H. H., Lindahl, E. C., and Schweda, E. K. (2005) Structural characterization of lipid A from nontypeable and type f *Haemophilus influenzae*: variability of fatty acid substitution, *Anal Biochem 340*, 303-316.

81. Roberts, M. C., Mintz, C. S., and Morse, S. A. (1986) Characterization of *Haemophilus parainfluenzae* strains with low-Mr or ladder-like lipopolysaccharides, *J Gen Microbiol 132*, 611-616.

82. Pollard, A., St Michael, F., Connor, L., Nichols, W., and Cox, A. (2008) Structural characterization of *Haemophilus parainfluenzae* lipooligosaccharide and elucidation of its role in adherence using an outer core mutant, *Can J Microbiol 54*, 906-917.

83. Westphal, O., and Jann, K. (1965) Bacterial lipopolysaccharide. Extraction with phenol water and further applications of the procedure,, *Methods Carbohyd. Chem. 5*, 88-91.

84. Galanos, C., Luderitz, O., and Westphal, O. (1969) A new method for the extraction of R lipopolysaccharides, *Eur J Biochem 9*, 245-249.

85. Holst, O., Brade, L., Kosma, P., and Brade, H. (1991) Structure, serological specificity, and synthesis of artificial glycoconjugates representing the genus-specific lipopolysaccharide epitope of Chlamydia spp, *J Bacteriol 173*, 1862-1866.

86. Schweda, E. K., Hegedus, O. E., Borrelli, S., Lindberg, A. A., Weiser, J. N., Maskell, D. J., and Moxon, E. R. (1993) Structural studies of the saccharide part of the cell envelope lipopolysaccharide from *Haemophilus influenzae* strain AH1-3 (lic3+), *Carbohydr Res 246*, 319-330.
87. Hellerqvist, C. G., and Sweetman, B. J. (1990) Mass spectrometry of carbohydrates, *Methods Biochem Anal 34*, 91-143.
88. Sawardeker, J. S., Sloneker, J. H., and Jeanes, A. (1965) Quntitative determination of monosaccharides as their alditol acetates by gas liquid chromatography., *Anal Chem 37*, 1602-1604.
89. Hellerqvist, C. G. (1990) Linkage analysis using Lindberg method, *Methods Enzymol 193*, 554-573.
90. Morelle, W., Slomianny, M. C., Diemer, H., Schaeffer, C., van Dorsselaer, A., and Michalski, J. C. (2004) Fragmentation characteristics of permethylated oligosaccharides using a matrix-assisted laser desorption/ionization two-stage time-of-flight (TOF/TOF) tandem mass spectrometer, *Rapid Commun Mass Spectrom 18*, 2637-2649.
91. Morelle, W., Faid, V., and Michalski, J. C. (2004) Structural analysis of permethylated oligosaccharides using electrospray ionization quadrupole time-of-flight tandem mass spectrometry and deutero-reduction, *Rapid Commun Mass Spectrom 18*, 2451-2464.
92. Sheeley, D. M., and Reinhold, V. N. (1998) Structural characterization of carbohydrate sequence, linkage, and branching in a quadrupole Ion trap mass spectrometer: neutral oligosaccharides and N-linked glycans, *Anal Chem 70*, 3053-3059.
93. Reinhold, V. N., and Sheeley, D. M. (1998) Detailed characterization of carbohydrate linkage and sequence in an ion trap mass spectrometer: glycosphingolipids, *Anal Biochem 259*, 28-33.
94. Reinhold, V. N., Reinhold, B. B., and Chan, S. (1996) Carbohydrate sequence analysis by electrospray ionization-mass spectrometry, *Methods Enzymol 271*, 377-402.
95. Li, J., Dzieciatkowska, M., Hood, D. W., Cox, A. D., Schweda, E. K., Moxon, E. R., and Richards, J. C. (2007) Structural characterization of sialylated glycoforms of H. influenzae by electrospray mass spectrometry: fragmentation of protonated and sodiated O-deacylated lipopolysaccharides, *Rapid Commun Mass Spectrom 21*, 952-960.
96. Li, J., Cox, A. D., Hood, D., Moxon, E. R., and Richards, J. C. (2004) Application of capillary electrophoresis- electrospray-mass spectrometry to the separation and characterization of isomeric lipopolysaccharides of *Neisseria meningitidis*, *Electrophoresis 25*, 2017-2025.
97. Kohler, M., and Leary, J. A. (1995) LC/MS/MS of carbohydrates with postcolumn addition of metal chlorides using a triaxial electrospray probe, *Anal Chem 67*, 3501-3508.
98. Li, J., and Richards, J. C. (2007) Application of capillary electrophoresis mass spectrometry to the characterization of bacterial lipopolysaccharides, *Mass spectrometry reviews 26*, 35-50.
99. Dzieciatkowska, M., Brochu, D., van Belkum, A., Heikema, A. P., Yuki, N., Houliston, R. S., Richards, J. C., Gilbert, M., and Li, J. (2007) Mass spectrometric analysis of intact lipooligosaccharide: direct evidence for O-acetylated sialic acids and discovery of O-linked glycine expressed by Campylobacter jejuni, *Biochemistry 46*, 14704-14714.
100. Kawasaki, N., Ohta, M., Hyuga, S., Hyuga, M., and Hayakawa, T. (2000) Application of liquid chromatography/mass spectrometry and liquid chromatography with tandem mass spectrometry to the analysis of the site-specific carbohydrate heterogeneity in erythropoietin, *Anal Biochem 285*, 82-91.
101. Ruhaak, L. R., Deelder, A. M., and Wuhrer, M. (2009) Oligosaccharide analysis by graphitized carbon liquid chromatography-mass spectrometry, *Anal Bioanal Chem 394*, 163-174.
102. Olling, A., Breimer, M. E., Peltomaa, E., Samuelsson, B. E., and Ghardashkhani, S. (1998) Electrospray ionization and collision-induced

dissociation time-of-flight mass spectrometry of neutral glycosphingolipids, *Rapid Commun Mass Spectrom 12*, 637-645.

103. Wilm, M., and Mann, M. (1996) Analytical properties of the nanoelectrospray ion source, *Anal Chem 68*, 1-8.

104. Bahr, U., Pfenninger, A., Karas, M., and Stahl, B. (1997) High-sensitivity analysis of neutral underivatized oligosaccharides by nanoelectrospray mass spectrometry, *Anal Chem 69*, 4530-4535.

105. Domon, B., and Costello, C. E. (1988) A Systematic Nomenclature for Carbohydrate Fragmentations in FAB-MS/MS Spectra of Glycoconjugates, *Glycoconjugate journal 5* 397 -409

106. Delaney, J., and Vouros, P. (2001) Liquid chromatography ion trap mass spectrometric analysis of oligosaccharides using permethylated derivatives, *Rapid Commun Mass Spectrom 15*, 325-334.

107. Viseux, N., de Hoffmann, E., and Domon, B. (1998) Structural assignment of permethylated oligosaccharide subunits using sequential tandem mass spectrometry, *Anal Chem 70*, 4951-4959.

108. Viseux, N., de Hoffmann, E., and Domon, B. (1997) Structural analysis of permethylated oligosaccharides by electrospray tandem mass spectrometry, *Anal Chem 69*, 3193-3198.

109. Gerwig, G. J., Kamerling, J. P., and Vliegenthart, J. F. (1979) Determination of the absolute configuration of mono-saccharides in complex carbohydrates by capillary G.L.C, *Carbohydr Res 77*, 10-17.

110. Mulloy, B. (1996) High-field NMR as a technique for the determination of polysaccharide structures, *Mol Biotechnol 6*, 241-265.

111. Agrawal, P. K. (1992) NMR spectroscopy in the structural elucidation of oligosaccharides and glycosides, *Phytochemistry 31*, 3307-3330.

112. Duus, J., Gotfredsen, C. H., and Bock, K. (2000) Carbohydrate structural determination by NMR spectroscopy: modern methods and limitations, *Chem Rev 100*, 4589-4614.

113. Vliegenthart, J. F. G. (2006) Introduction to NMR spectroscopy of carbohydrates, *Acs Sym Ser 930*, 1-19.

114. van Halbeek, H. (1994) 1H nuclear magnetic resonance spectroscopy of carbohydrate chains of glycoproteins, *Methods Enzymol 230*, 132-168.

115. Yu, R. K., Koerner, T. A., Scarsdale, J. N., and Prestegard, J. H. (1986) Elucidation of glycolipid structure by proton nuclear magnetic resonance spectroscopy, *Chem Phys Lipids 42*, 27-48.

116. Hood, D. W., Deadman, M. E., Engskog, M. K., Vitiazeva, V., Makepeace, K., Schweda, E. K., and Moxon, R. (2010) Genes required for the synthesis of heptose-containing oligosaccharide outer core extensions in *Haemophilus influenzae* lipopolysaccharide, *Microbiology 156*, 3421-3431.

117. Arbatsky, N. P., Kondakova, A. N., Senchenkova, S. N., Siwinska, M., Shashkov, A. S., Zych, K., Knirel, Y. A., and Sidorczyk, Z. (2007) Structure of a new ribitol teichoic acid-like O-polysaccharide of a serologically separate Proteus vulgaris strain, TG 276-1, classified into a new Proteus serogroup O53, *Carbohydr Res 342*, 2061-2066.

118. Lipkind, G. M., Shashkov, A. S., Knirel, Y. A., Vinogradov, E. V., and Kochetkov, N. K. (1988) A computer-assisted structural analysis of regular polysaccharides on the basis of 13C-n.m.r. data, *Carbohydr Res 175*, 59-75.

119. Bacterial Carbohydrate Structure DataBase, csdb.glycoscience.ru/bacterial/.