From Clinical Research Centre,
Department of Laboratory Medicine,
Karolinska Institutet, Stockholm, Sweden

Structural Studies on Lipopolysaccharides from *Haemophilus* Species

Brigitte Twelkmeyer

Karolinska Institutet

Stockholm 2011

All previously published papers were reproduced with permission from the publisher.

To my family

ABSTRACT

Background: Carbohydrates are indispensable mediators for a variety of cellular interactions. In biological systems carbohydrates are usually linked to a carrier as e.g. proteins or lipids. One example is the molecule lipopolysaccharide (LPS). LPS is one of the major outer membrane constituents found in Gram-negative bacteria and they play key roles in the biology of these organisms. Notably, they have been found to be important virulence factors in pathogenic species.

Aims: In this thesis the biosynthesis and the molecular structures of LPS expressed by *Haemophilus influenzae* and *Haemophilus parainfluenzae* were investigated. These two bacteria colonize the human nasopharynx. *H. influenzae* with capsule type b is involved in invasive diseases such as meningitidis and epiglottis while non-encapsulated *H. influenzae* (non-typeable, NTHi) often cause otitis media, and acute and chronic lower respiratory tract infections in infants. Though, closely related *H. parainfluenzae* is a commensal.

Materials and Methods: Structural elucidation of LPS involves initial de-lipidation to obtain water-soluble material that is suitable for subsequent analyses by chemical, nuclear magnetic resonance (NMR) and mass spectrometric (MS) methods.

Results: The function of the gene *lic2B* in *H. influenzae* type b strain Eagan was investigated. The LPS expressed by the mutant strain Eagan*lic2B*lic2C+ was analyzed and found to encode for a glucosyltransferase responsible for the addition of β-D-Glc*p* to O-4 of α-D-Glc*p*-(1→ elongating from the middle inner core heptose. Further, the LPS structures of NTHi strains 1247 and 1008 were determined. NTHi strain 1247 expressed globotetraose elongating from the phosphocholine bearing GlcI, [β-D-Gal*p*NAc-(1→3)-α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→4)-[*P*Cho→6]-β-D-GlcI*p*-(1→], or truncated versions thereof. Globotetraose was also found to elongate from the distal heptose in NTHi 1247. The *lpsA* mutant of strain 1247 allowed the identification of a novel disialyllactose epitope, [α-Neu5Ac-(2→8)-α-Neu5Ac-(2→3)-β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→], from the proximal inner core heptose. Alternatively, a globotetraose was found to elongate from GlcI. NTHi strain 1008 showed almost identical glycoforms to NTHi strain 1247 but lacked terminal *N*-acetyl galactoseamine.

All of the three investigated *H. parainfluenzae* strains were shown to express the same lipid A and inner core as NTHi. *H. parainfluenzae* genome strain T3T1 and strain 22 were shown to express rough-type LPS having novel outer core structures elongating from GlcI that were [α–Neu5,9Ac₂-(2→6)-β-D-Gal*p*NAc-(1→4)-β-D-Gal*p*-(1→3)-β-D-Fuc*p*NAc4N-(1→] in strain T3T1 and [Neu5Ac-(2→6)-α-D-Gal*p*-(1→6)-β-D-Glc*p*-(1→3)-β-D-Fuc*p*NAc4N-(1→] in strain 22. *H. parainfluenzae* strain 13 expressed an O-repeating chain with the structure [→6)-[Ac→3]-β-D-Gal*f*-(1→3)-[*P*Etn→6]-β-D-Glc*p*NAc-(1→].

Conclusion: It was showed that the *H. parainfluenzae* LPS investigated here lacks all virulence determining LPS attributes expressed by *H. influenzae* such as phosphocholine and phase variable expression of outer core structures. This may provide further insight into the factors relating to commensal vs. pathogenic behavior inside the host.

LIST OF PUBLICATIONS

I.   The role of *lic2B* in lipopolysaccharide biosynthesis of *Haemophilus influenzae* strain Eagan.
Brigitte Twelkmeyer, Mary E. Deadman, Ehsanul Haque, Jianjun Li, Derek W. Hood, Elke K.H. Schweda
Carbohydrate Research, Vol. 346, 10, (2011) 1262-1266

II.   Expression of a new disialyllacto structure in the lipopolysaccharide of non-typeable *Haemophilus influenzae*.
Brigitte Twelkmeyer, Peter K. Burström, Jianjun Li, E. Richard Moxon, Derek W. Hood, Elke K.H. Schweda
Carbohydrate Research, *in press*

III. Rough type lipopolysaccharide expression by *Haemophilus parainfluenzae*; The lipopolysaccharide structure of genome sequence strain T3T1, and strain 22.
Brigitte Twelkmeyer, Jianjun Li, Rosanna Young, Derek W. Hood, Elke K.H. Schweda
Manuscript

IV.   The lipopolysaccharide structure expressed by *Haemophilus parainfluenzae* strain 13.
Brigitte Twelkmeyer, Rosanna Young, Derek W. Hood, Elke K.H. Schweda
Manuscript

Related Publications-not included in the present thesis

I.   Novel globoside-like oligosaccharide expression patterns in nontypeable *Haemophilus influenzae* lipopolysaccharide.
Susanna L. Lundström, Brigitte Twelkmeyer, Malin K. Sagemark, Jianjun Li, James. C. Richards, Derek W. Hood, E. Richard Moxon, Elke K.H. Schweda
European Journal of Biochemistry, 2007, 274(18): 4886-4903

II.   Profiling structural elements of short-chain lipopolysaccharide of non-typeable *Haemophilus influenzae*.
Elke K.H. Schweda, Brigitte Twelkmeyer, Jianjun Li
Innate Immunity, 2008, 14(4):199-211

III. Structural studies of the lipopolysaccharide from *Haemophilus parainfluenzae* strain 20.
Varvara Vitiazeva, Brigitte Twelkmeyer, Rosanna Young, Derek W. Hood, Elke K.H. Schweda
Carbohydrate Research, *in press*

CONTENTS

LIST OF ABBREVIATIONS

1D	one-dimensional
2D	two-dimensional
Ac	*O*-acetyl group
*An*Kdo-ol	reduced anhydro Kdo
CE-ESI-MS/MS	capillary electrophoresis tandem ESI-MS
COSY	correlation spectroscopy
CPS	capsular polysaccharide
Gal	galactose
GalNAc	2-acetamido-2-deoxy-galactose
Glc	glucose
DMSO	dimethyl sulfoxide
ESI-MS	electrospray ionization mass spectrometry
Fuc*p*NAc4N	2-acetamino-4-amino-2,4,6-trideoxy-galactopyranose
GC-MS	gas chromatography mass spectrometry
Gly	glycine
Hep	heptose
Hex	hexose
HexNAc	*N*-acetyl-hexosamine
HMBC	heteronuclear multiple bond correlation
HMQC	heteronuclear multiple quantum correlation
HPAEC	high performance anion exchange chromatography
HSQC	heteronuclear single quantum correlation
Kdo	3-deoxy-D-*manno*-oct-2-ulosonic acid
LC	liquid chromatography
L-D-Hep	L-*glycero*-D-*manno*-heptose
LPS	lipopolysaccharide
LPS-OH	O-deacylated LPS
MS^n	multiple step tandem MS
Neu5Ac	5-*N*-acetylneuraminic acid, sialic acid
NTHi	non-typeable *Haemophilus influenzae*
NMR	nuclear magnetic resonance
NOESY	nuclear Overhauser enhancement spectroscopy
OS	oligosaccharide
P	phosphate
PAD	pulsed amperometric detection
*P*Cho	phosphocholine
*P*Etn	phosphoethanolamine
*PP*Etn	pyrophosphoethanolamine
PMAA	partially methylated alditol acetate
SDS PAGE	sodium dodecylsulphate polyacrylamide gel electrophoresis
TOCSY	total correlation spectroscopy

1 Introduction

Carbohydrates are ubiquitos in nature. In cells they can be found on the outside attached to proteins and lipids as part of the so called glycocalix. They are indispensable mediators for biological interaction during cell to cell communication, adhesion and recognition. They are also important for the correct functioning of proteins and often used for the differentiation of self and non-self in the human body [1].

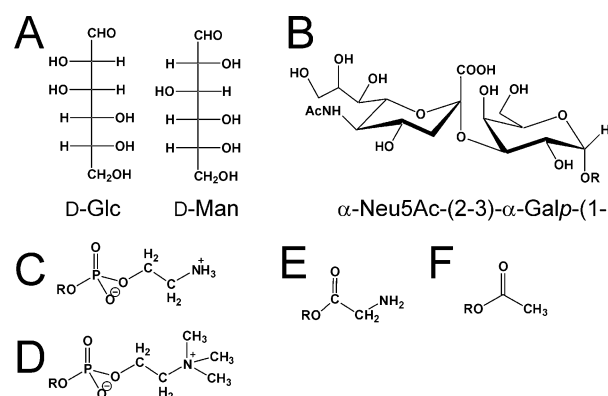Variation is the key attribute of carbohydrate molecules that can be combined in multiple linkages. Carbohydrates differ in many properties giving rise to diastereomers as illustrated in Figure 1A on the example of glucose (Glc) and mannose (Man). Further, the structures can differ in the absolute configuration (D or L), sugars that are open or in the half acetal ring formation and two different anomeric configurations (α or β), Fig. 1B. Natural decoration with carbohydrates normally includes oligo- or polysaccharides which in their turn differ in the attachment points, the possibility of branching, length and auxiliary elongation with non-carbohydrate substituents such as phosphate (P) and its derivatives phosphoethanolamine (PEtn, Fig. 1C), phosphocholine (PCho, Fig. 1D) as well as simple amino acids such as glycine (Gly, Fig. 1E) or acetate (Ac, Fig. 1F).

Figure 1: Carbohydrates and functional groups found in LPS. A) D-Glucose and D-Mannose, two diastereomers; B) α sialic acid bound to α-D-Gal$p$; C) PEtn; D) PCho; E) glycine and F) acetate.

Carbohydrates are often linked to proteins or lipids as for example in lipopolysaccharide (LPS). LPS is built up by lipid A that anchors the molecule into the outer membrane, the inner and outer core and an optional O-repeating unit (Fig. 2A). LPS is found in Gram-negative bacteria, that in contrast to Gram-positive bacteria, are surrounded by an outer membrane (Fig. 2B and C). Gram-positive bacteria have a thick cell wall that absorbes a dye as discovered by Gram in 1884, while Gram-negative bacteria do not. This is due to the outer membrane that inhibits the inclusion of the dye [2, 3].
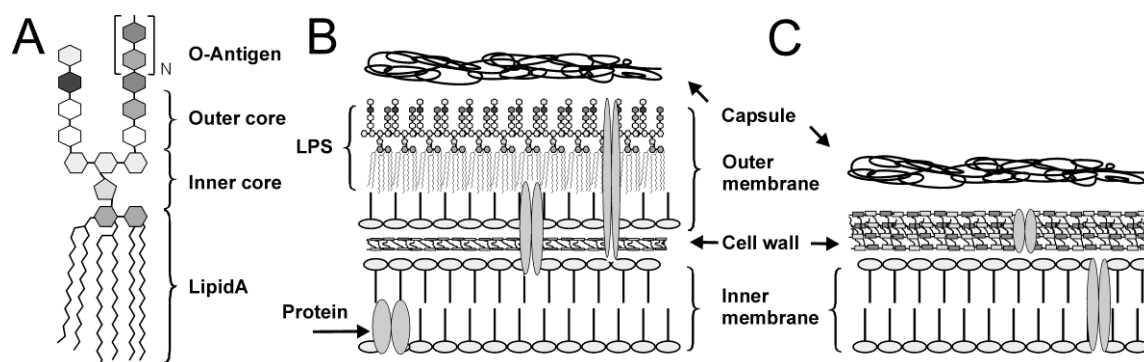
Figure 2: Schematic representation of A) lipopolysaccharide with optional O-antigen B) the cell wall of Gram-negative bacteria, and C) Gram-positive bacteria.

LPS is essential for growth of Gram-negative bacteria. Lack of LPS expression is lethal for most of them (with the only exception of a *lpxA* genetic mutant of *Neisseria meningitidis* that is reported to be vital even without LPS [4]). The correct assembly of the outer membrane requires LPS which imparts membrane stability. This is accomplished by lipid A that is substituted by phosphate groups. They are connected to each other using doubly charged positive ions such as Mg^{2+}, Ca^{2+} via ionic bonds. As a result the outer membrane is a strong and viscous layer [5].

LPS fulfills several functions such as a barrier allowing only small and selected molecules to enter via porines, a group of proteins that regulates entry. Environmental stress such as antibiotics are not affecting the Gram-negative bacteria in the same way as they affect Gram-positive bacteria, since their entrance is hindered. Due to LPS´s location on the outer membrane it comes into contact with the surrounding environment. As a result, the carbohydrate moiety of LPS is important in host-pathogen interactions, including adhesion, colonization and in provoking immune responses. Additionally, LPS can express human carbohydrate patterns known as a form of molecular mimicry [6, 7].

1.1   Two bacteria in the nasopharynx of humans

H. influenzae and *H. parainfluenzae* colonize the human nasopharynx from where they can migrate to the inner ear and respiratory tract. They are transmitted from one individual to the next by droplet infection. Once inside the new host, *H. influenzae* can cause several diseases while *H. parainfluenzae* normally does not.

Over the past decades research on *H. influenzae* increased the knowledge and understanding about its behavior considerably. However, only little is known about *H. parainfluenzae*.

1.1.1   *Haemophilus influenzae*

The strictly human pathogen *H. influenzae* has been the subject of intense scientific investigations due to its high pathogenicity and the high lethal rate of meningitis caused by *H. influenzae* type b (Hib). *H. influenzae* is a small (0.2-0.3×0.5-0.8 μm), non-motile Gram-negative coccobacillus, that was first discovered in 1890 by Pfeiffer who named it "blood loving", *Haemophilus*, as he established its need for haemin and nicotinamide adenine dinucleotide for growth. Since Pfeiffer thought it was the cause of pandemic influenzae he named it *Haemophilus influenzae*. It became later clear that *H. influenzae* colonizes the nasopharynx and to a lesser extent the trachea and bronchi of humans as this is its only habitat. *H. influenzae* can be a commensal, but frequent pathogenic behaviour is also known.

H. influenzae can be categorized into two groups, encapsulated (typeable) and non-encapsulated (non-typeable) forms. Capsular serotypes are categorized into six classes a - f, depending on the molecular structure of the capsular polysaccharide [8]. *H. influenzae* type b (Hib) is the most virulent and able to invade the bloodstream to cause meningitis (infection of the membranes enveloping the brain and spinal cord), pneumonia (lung infection), epiglottis (life-threatening inflammation of the flexible cartilage that covers the gap in the vocal cords during swallowing), septic arthritis (inflammation of the joints),

septicemia and bacteremia (blood poisoning), and cellulitis (inflammation of connective tissue). The other capsule bearing *H. influenzae* types rarely cause invasive disease.

The non encapsulated *H. influenzae* (non-typeable, NTHi) are also frequently pathogenic, causing otitis media (middle ear infection), Fisher syndrome (a post infectious autoimmune disorder) and acute and chronic respiratory tract infections [9, 10]. Other bacteria causing otitis media are *Moraxella catarrhalis* and *Streptococcus pneumoniae*. As this is one of the major reasons for childhood pediatric visits and antibiotic treatment, understanding of the molecular basis of the disease mechanism is needed in order to develop better diagnostic tools and to design alternative, preventative methods and treatments.

Surface expressed carbohydrate antigens such as capsular polysaccharide (CPS) and LPS are important factors for the pathogenicity and carrier state of *H. influenzae* [11, 12]. A glycoconjugate vaccine has been developed against the CPS structure of *H. influenzae* type b. The introduction of the vaccine in the 1990`s has resulted in a dramatic decline of Hib induced meningitidis in infants [13-17]. Never the less, these vaccines do not provide protection against diseases caused by NTHi since they lack the specific type b capsule. LPS carbohydrate structures could provide the basis for a vaccine against NTHi caused diseases.

1.1.2 *Haemophilus parainfluenzae*

H. parainfluenzae is another bacterium populating the nasopharynx and closely related to *H. influenzae*. It is a highly adapted human commensal. Most 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. Despite being closely related to *H. influenzae*, *H. parainfluenzae* very rarely cause disease, and few cases have been reported recently [18-25].

Little work has been done on the structure, genetics and biology of the *H. parainfluenzae* LPS molecules. The complete genome sequence of one *H. parainfluenzae* strain, T3T1, has recently been completed at the Sanger Centre, UK, in collaboration with researchers at University of Oxford, UK [DW Hood, unpublished data]. This data now provides the opportunity to investigate the difference in LPS structures and biosynthesis between the two species.

1.1.3 *H. influenzae* lipopolysaccharide: structure and genetics

LPS biosynthesis in *H. influenzae* is carried out with the outer core building blocks attached step by step (see page 11).

The outer membrane anchor, the lipid A, of *H. influenzae* was first investigated by Helander *et. al.* [26] establishing β-2-amino-2-deoxy-D-glucopyranose-(1→6)-α-2-amino-2-deoxy-D-glucopyranose backbone with phosphate (*P*) substitution on C-1 and C-4` on the reducing and non-reducing glucosamine residue, respectively. Amine linked and ester linked 3-hydroxy tetradecanoic and tetradecanoic fatty acids substitute the C-2/C-2` and C-3/C-3` positions , respectively. Further ester linked tetradecanoic acids are attached on C-2` and C-3`. Structural studies on lipid A done in our lab established a pronounced diversity in the lipid part of NTHi LPS in which *O*-acetyl groups, and fatty acid chain length variation including dodecanoic acid and decanoic acid were identified [27].

Extensive structural studies of LPS from *H. influenzae* have led to the identification of a conserved inner core region shown in Figure 3 consisting of a triheptosyl moiety L-α-D-Hep*p*-(1→2)-[*P*Etn→6]-L-α-D-Hep*p*-(1→3)-L-α-D-Hep*p* which is linked to lipid A via one 3-deoxy-D-*manno*-oct-2-ulosonic acid (Kdo) substituted at C-4 with *P* or pyrophosphoethanolamine (*PP*Etn). The determination of Kdo was hindered in early investigations due to the phosphate group [28]. However, dephosphorylation of Kdo lead to the unambiguous detection of this acid labile residue [29]. The stereochemical assignment of the inner core heptose as L-*glycero*-D-*manno*-heptose is based upon biosynthetic studies established by Coleman *et. al.* [30]. In every investigation to date, the first heptose has always been found to be extended by β-D-Glc*p*-(1→ (GlcI) at the 4 position.
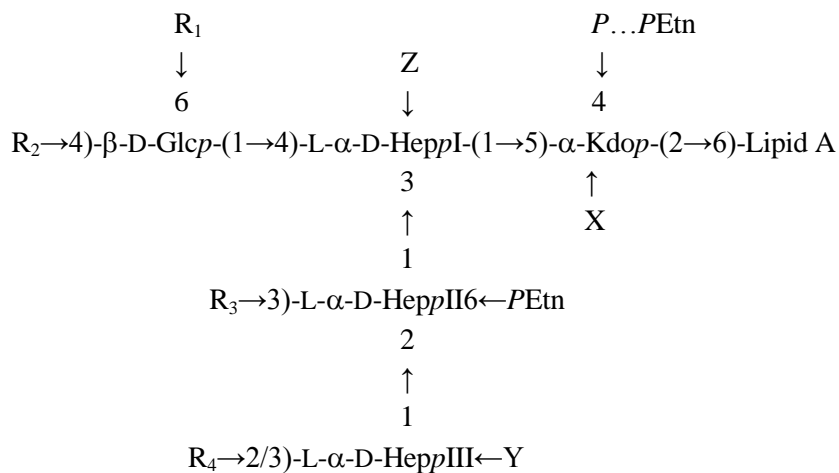
```
        R₁                              P…PEtn
        ↓                Z               ↓
        6                ↓               4
R₂→4)-β-D-Glcp-(1→4)-L-α-D-HeppI-(1→5)-α-Kdop-(2→6)-Lipid A
                         3               ↑
                         ↑               X
                         1
           R₃→3)-L-α-D-HeppII6←PEtn
                         2
                         ↑
                         1
           R₄→2/3)-L-α-D-HeppIII←Y
```

Figure 3: Structure of LPS from *H. influenzae* where R₁ is H, *P*Cho or D-α-D-Hep*p*, L-α-D-Hep*p*; R₂ is H, β-D-Glc*p*, β-D-Gal*p* or Ac; R₃ is H or α-D-Glc*p*; R₄ is H, β-D-Glc*p* or β-D-Gal*p*; X is H or Gly; Y is H, *P*Etn or Gly and Z is H or Ac.

The inner core of NTHi LPS serves as a template for further carbohydrate extensions as well as non-carbohydrate substitution where *P*, *P*Cho, *P*Etn, Gly and acetate groups have been identified [31]. The described outer core structures for NTHi display a great deal of diversity between and within strains [31].

Branching in NTHi LPS is commonly found in three positions: on HepI, HepII and GlcI that can be further elongated by outer core heptoses and hexoses. Furthermore, the outer heptose (HepIV) in NTHi strain 1232 has been found to be branched [32].

The outer core carbohydrate extensions have been found to mimic human glycolipid structures as shown in Figure 4 including globotetraose (1.), globoside (2.), galabiose (3.) lactose (4.), sialyllactose (5.), sialyllacto-*N*-neotetraose (6.) and its related structure *P*Etn→6)-α-D-Gal*p*NAc-(1→6)-β-D-Gal*p*-(1→4)-β-D-Glc*p*NAc-(1→3)-β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→ (7.) [6, 10, 33]. Structures resembling globotetraose and globoside have been observed extending from HepII. Since the first glucose substituting HepII is in α configuration these structures have been named globoteraose-like.

Interestingly, the outer core oligosaccharide structures of NTHi are all related having the same basis and getting more complex in building blocks of one further sugar

unit. Truncated versions are typically observed due to incomplete biosynthesis (micro heterogeneity) and phase variation which increases the observed number of glycoforms while the genetic background is kept simple.

It has been found that the expression of outer core structures are associated with increased virulence. These oligosaccharide structures are all controlled by specific loci or genes, namely *lic1*, *lic2*, *lic3*, *lgtC*, *lex2*, *oafA*, *siaA2* that allow the bacteria to adapt to changing environmental situations within the host [12, 34, 35].

All of these loci and genes are subjected to phase variation. Phase variable genes include tetranucleotide repeats that consist of either CAAT (*lic1A*, *lic2A*, *lic3A, lic3B*), GCAA (*lex2*, *oafA*), or GACA (*lgtC*) which are located at the 5´end of the open reading frame of the gene. These areas of the genome are unstable and can easily be misaligned and rearranged such that one tetranucleotide unit is easily lost or gained which results in an altered reading pattern of the genome, i.e. a frameshift. Hence, gene function is reversibly lost or gained due to slipped-strand mispairing, a process described as the "on-off switching" of genes. The outcome is the complete or incomplete biosynthesis of the outer core chain extensions. This random process occurs at high frequency and contributes to the virulence of *H. influenzae* by constantly changing the oligosaccharide extensions of the bacteria. [36].

Phase variation has been observed in other bacteria such as *Helicobacter pylori* [37], *Neisseria meningitidis* [38] and *Haemophilus somnus* [39]. It believed to facilitate colonization of these bacteria and the repression of the host immune system.

1. β-Gal*p*NAc-(1→3)-α-Gal*p*-(1→4)-β-Gal*p*-(1→4)-β-Glc*p*-(1→
2. α-Gal*p*-(1→4)-β-Gal*p*-(1→4)-β-Glc*p*-(1→
3. α-Gal*p*-(1→4)-β-Gal*p*-(1→
4. β-Gal*p*-(1→4)-β-Glc*p*-(1→
5. α-Neu5Ac-(2 →3)–β-Gal*p*-(1→4)-β-Glc*p*-(1→
6. α-Neu5Ac-(2→3)-β-Gal*p*-(1→4)-β-Glc*p*NAc-(1→3)-β-Gal*p*-(1→4)-β-Glc*p*-(1→
7. *P*Etn→6)-α-Gal*p*NAc-(1→6)-β-Gal*p*-(1→4)-β-Glc*p*NAc-(1→3)-β-Gal*p*-(1→4)-β-Glc*p*-(1→

Figure 4: The outer core glycoforms found in *H. influenzae* LPS. All sugars are in D configuration.

The genome of *H. influenzae* strain Rd was the first free living organism to be completely genome sequenced [40]. In addition *H. influenzae* strain Eagan and RM 7004 were extensively used to identify additional LPS biosynthesis genes. Recently, NTHi strains 86-028NP, R2846 and R2866 were also sequenced [41, 42]. Furthermore, the study of *H. influenzae* involved a systematic analysis of clinical isolates that were chosen to be representative for the biological diversity of the species. The majority of these NTHi strains originated from a Finish otitis media study [43].

The combined genetic information allowed a more detailed view of the genomic background. Taken together it formed a stable platform for the investigation of genes involved in the LPS biosynthesis in *H. influenzae*. Candidate genes could be proposed that are responsible for the biosynthesis of LPS and, by analytical chemistry, the expressed

LPS forms could be analyzed in order to prove the biosynthetic hypotheses [12, 44]. Moreover, genes of validated function could be compared among strains thereby facilitating the identification of genes common to all *H. influenzae*.

Some LPS biosynthetic genes were found to be invariably present in *H. influenzae*. These are the genes responsible for the synthesis of lipid A and inner core, see Figure 5. Furthermore, all strains posses the *lgtF* and *lpsA* genes.

*P*Cho   *P*…*P*Etn

lic1D→ ↓   **lgtF**   **kdkA→**↓

6   ↓   4

β-Gal*p*-(1→4)-β-Glc*p*-(1→4)-L-α-D-Hep*p*I-(1→5)-α-Kdo-(2→6)-LipidA

↑   3   ↑   ↑

lex2B   **rfaF→**↑   **opsX**   **kgtA**

1

β-Gal*p*-(1→4)-β-Glc*p*-(1→4)-α-Glc*p*-(1→3)-L-α-D-Hep*p*II6←*P*Etn

2   ↑

orfH→↑   **lpt6**

1

β-Gal*p*NAc-(1→3)-α-Gal*p*-(1→4)-β-Gal*p*-(1→4)-β-Glc*p*-(1→2)-L-α-D-Hep*p*III3←*P*Etn

↑   ↑   ↑   ↑   ↑

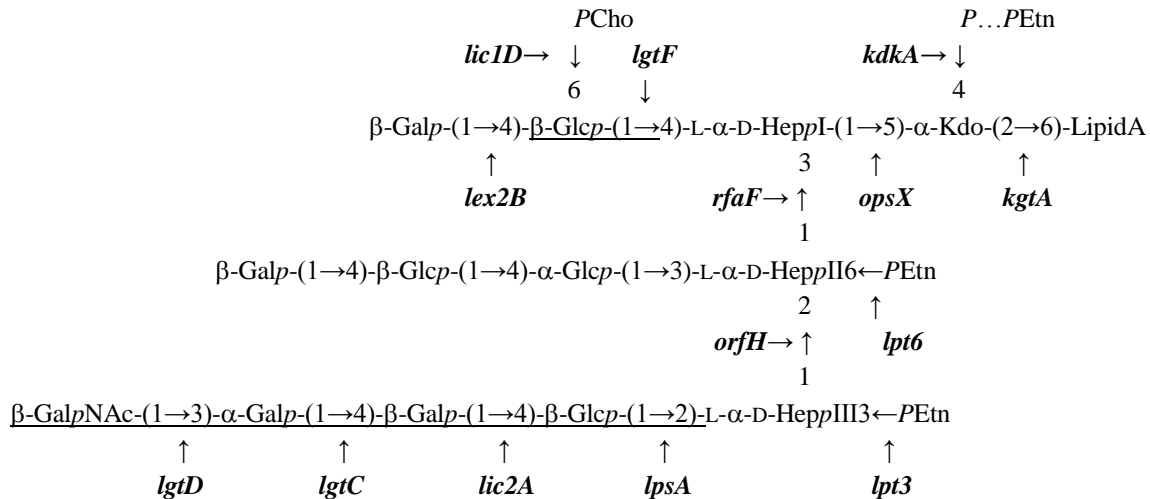**lgtD**   **lgtC**   **lic2A**   **lpsA**   **lpt3**

Figure 5: The combined schematic representation of LPS structures of *H. influenzae* strain Rd, type d strain (RM 118) and strain 2019 [45, 46]. The biosynthetic genes are in bold and italics. Strain Rd specific outer core oligosaccharide structures are underlined. All sugars are in D configuration, unless stated otherwise.

1.1.3.1 Outer core extentions from GlcI
The most variety of outer core extentions were observed from GlcI. Elongation at O-4 of GlcI is dependent on the phase variable *lex2* locus as it was established in NTHi strain 1124 and 2019 (Figure 5) [47, 48]. It encodes either for a glucosyltransferase or a galactosyltransferase depending on the amino acid at position 157 [47]. Further extension of gobotetraose and its truncated versions, as well as the sialylated and disialylated derivates of lactose were observed extending from Glc or Gal [49] (**Paper II**).

Furthermore, GlcI can be substituted by *P*Cho at O-6, or elongated with outer core heptose.

1.1.3.1.1 Outer core heptoses
About 20% of NTHi strains contain the genes *losB1* and *losB2* that code for heptosyltransferases [41, 50]. They result in carbohydrate extension at O-6 of GlcI with the incorporation of an outer core heptose (HepIV) [31]. HepIV has been found to possess either the L-α-D-Hep*p* or D-α-D-Hep*p* configuration. Responsible for implementing D-α-D-Hep*p* is the gene *losB1* and the expression of L-α-D-Hep*p* is governed by *losB2* [51]. L-α-D-Hep*p* has been observed in NTHi strains 1209 [52] and R2866 [Engskog, MKR *et. al.*, unpublished results]. D-α-D-Hep*p* was described for the first time in NTHi strain 9274 [53] and was later found in NTHi strains 981 [54] and 1232 [32]. In NTHi strain R2846 both configurations for HepIV have been observed [55].

Incorporation of D,D-Hep into the outer core LPS appears to be common for *Haemophilus* species and was also found in *H. ducreyi* and *H. parainfluenzae* (see page 10) in which it also substitutes at O-6 of GlcI [56, 57].

1.1.3.1.2 Outer core structures from GlcI added blockwise

Approx. 40% of the investigated *H. influenzae* strains contained the *hmg* locus. It enables the bacterium to synthesize the epitopes sialyllacto-*N*-neotetraose and its related structure that contains *P*Etn→6)-α-D-Gal*p*NAc-(1→6 at O6 of β-D-Gal*p* instead of Neu5Ac, see Fig. 4. These structures are synthesized by the addition of a complete tetrasaccharide unit to O-4 of GlcI. They are the only known structures added *en bloc* in contrast to the otherwise stepwise addition of the outer core oligosaccharide structures [58].

1.1.3.2   Outer core extentions from HepII

The *lic2* locus is present in about 50% of the investigated strains and consists of four genes, *lic2A, ksgA, lic2C* and *lic2B*. *lic2C* is responsible for extension of HepII by α-D-Glc*p*. α-D-Glc*p* can be substituted by *P*Cho and further substituted as part of globotetraose-like structure [59, 60]. The enzyme adding the next β-D-Glc*p*- was characterized to be Lic2B (**Paper I**). Lic2A and LgtC add sequentially the hexoses to give the globoside epitope [α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→] extending from HepII in *H. influenzae* strain Eagan and *H. influenzae* type b strain RM 7004, Figure 6 [61-63]. LgtC can also be active in the globoside extending from HepIII [49].

Alternatively, elongation with β-D-Gal*p*NAc-(1→3)-α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→4)-α-D-Glc*p*-(1→ elongating from O-3 of HepII has been observed in NTHi strains 1268 and 1200 [60].

1.1.3.3   Outer core extentions from HepIII

Oligosaccharide chains extending from HepIII are depending on *lpsA*. It is present in all strains investigated. The gene encodes for a glycosyltransferase directing either β-D-Gal*p*-(1→ or β-D-Glc*p*-(1→ to HepIII [64]. If the amino acid at position 151 is threonine, *lpsA* was found to encode a galactosyltransferase, while cysteine, alanine or methionine at the same position resulted in a glucosyltransferase. Depending on the 3´end the hexose was added to either O-2 or O-3 of HepIII [64]. Interestingly β-D-Gal can be substituted by *P*Cho and Ac but is otherwise terminal. Extensions from β-D-Glc include globotetraose and its truncated versions as well as the sialylated and disialylated forms of lactose.

1.1.3.4   Phosphocholine

H. influenzae LPS are frequently decorated with phosphocholine (*P*Cho). In NTHi it was found at O-6 of GlcI [49, 65-67], O-6 of terminal Gal on HepIII [68, 69], O-6 of α-Glc on HepII [59, 60], but also attached to an outer core heptose, i.e. at O-7 of L-α-D-Hep*p* in NTHi strain R2866 [Engskog, MKR *et.al*. unpublished results].

The genes that are responsible for the incorporation and addition of *P*Cho are located in the *lic1* operon that is subjected to phase variation [70]. *Lic1* is divided into four genes, *lic1A* to *D* [70]. *Lic1B* is responsible for the uptake of choline from the bacterium's

surrounding. The enzyme Lic1A phosphorylates the choline giving *P*Cho, Lic1C is a pyrophosphorylase that activates *P*Cho. The last step involves Lic1D which directs the transfer to the LPS molecule [71-73].

Once incorporated, *P*Cho has been shown to be targeted by the C-reactive protein (CRP) that binds to *P*Cho and helps to kill the bacterium via activation of complement [74, 75]. In a distinct subset of NTHi, duplicate copies of the *lic1* are found, resulting in the expression of two *P*Cho residues in the same LPS molecule [73].

*P*Cho has also been found in *Streptococcus pneumoniae* [76, 77] and *Neisseria meningitidis* [78]. Hence, it might provide a general advantage for the bacteria when occupying the human respiratory tract.

α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→4)-β-D-Glc*p*-(1→4)-L-α-D-Hep*p*I-(1→5)-*An*Kdo-ol

$$3$$
$$\uparrow$$
$$1$$

α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→4)-α-D-Glc*p*-(1→3)-L-α-D-Hep*p*II6←*P*Etn

$$2$$
$$\uparrow$$
$$1$$

*P*Cho→6)-β-D-Gal*p*-(1→2)-L-α-D-Hep*p*III

Figure 6: The LPS structure of *H. influenzae* type b strain RM 7004 [62].

1.1.3.5 Sialic acid

N-acetyl-neuraminic acid (sialic acid, Neu5Ac) is found in almost all NTHi strains [79]. Sialic acid is incorporated from the host environment since NTHi cannot synthesize it *de novo* [80]. NTHi utilizes Neu5Ac either to gain energy or adds it to its LPS. The *lic3* locus encodes for the sialyltransferases, *lic3A* and *lic3B*. The gene *lic3A* was shown to encode for a α-2,3-sialyltransferase adding α-Neu5Ac to O-3 of β-D-Gal*p* in lactose [81]. Lic3A competes with the galactosyltransferase LgtC for the lactose receptor, giving two possible outcomes either sialylated lactose or digalactosyl. Lic3B can act the same way as Lic3A but additionally it can add α-Neu5Ac to O-8 of another sialic acid giving α-Neu5Ac-(2→8)-α-Neu5Ac-(2→ [82].

Sialic acid was found to cap lactose (α-Neu5Ac-(2→3)–β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→) for the first time in NTHi strain 375 [83]. NTHi strain DH1 was shown to express the structure [α-Neu5Ac-(2→8)-α-Neu5Ac-(2→3)-β-D-Gal*p*-(1→4)-L-α-D-Hep*p*-(1→] extending from GlcI. A knock-out mutant DH1*lic3B* did not express any sialylation while DH1*lic3A* did not show any alteration compared to the wildtype. Interestingly, the sialylated structure was only found bound to HepIV with no evidence for sialylation from the lactose structure extending from HepIII [7]. The structural difference of the sialylation site might explain why this strain is related to Fisher syndrome while other NTHi are not [7, 9].

Sialylated glycoforms are important for the resistance of the bacteria to normal human serum [79, 81]. They have a neglecting effect on the attachment of cultured human

epithelial cells and neutrophils [79]. Sialylated LPS glycoforms were shown to play a crucial role during the course of infection in the chinchilla model of otitis media [84, 85].

1.1.3.6   Non-carbohydrate substituents

In every NTHi strains published to date, *P*Etn was found substituting HepII at O-6 and in some strains also at O-3 of HepIII. The gene *lpt6* was proven to be responsible for the substitution at O-6 of HepII with *P*Etn [86]. *P*Etn substitution at O-3 of HepIII was shown to be directed by *lpt3* as shown in a mutant strain of NTHi strain 2019 [46]. Its biological significance is not established yet.

Phosphate addition as part of the inner core is mediated by the gene *kdkA* and is always found on Kdo at the 4 position where it is substituted by *P*Etn [12]. Phosphate groups have also been observed on other positions mainly substituting the distal heptose as *e.g.* in *H. influenzae* strain Eagan [61].

O-acetyl groups have been identified in NTHi LPS, both substituting the inner and outer core. Ac can substitute GlcI, HepI, HepIII, the outer core Glc-HepIII, Gal-HepIII, and substituting the outer core chain extending from GlcI in NTHi strain R2846. The location is typically determined by MS/MS. The exact positions can be determined by NMR. However, the substitution site is difficult to define due to migration. Yet, studies on NTHi strain 723 have been successfully accomplished showing Ac to substitute O-3 and O-4 of GlcI, at O-2 of HepI and on HepIII (position not determined) [67] and in strain 486 HepIII were substitution occurred at O-2 [59] and in strain 1209 at O-3 [52].

It was shown that *O*-acetylation increases the resistance of the bacteria towards host immune clearance when compared to the *O*-acetyl deficient mutants [87]. The phase variable gene *oafA* has been identified to be involved in the addition of outer core acetate groups. In a knock out mutant, NTHi strain 285*oafA*, GlcI was still found to be *O*-acetylated while acetylation on HepIII was not observed. This indicated that there are more genes yet unidentified mediating *O*-acetylation [87].

Glycine was found to substitute the inner core of NTHi and typeable *H. influenzae* strains by Li *et. al.* [88]. In that study, Gly was detected using high performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD), and the substitution was predominantly found on HepIII and to a lesser extent on Kdo by CE-ESI-MS/MS. However, NTHi strain 162 and strain 1233 showed Gly on HepII and HepI, respectively. The exact linkage positions of Gly have not been established in any *H. influenzae* strain and the genes involved in Gly addition have not yet been elucidated.

1.1.3.7   Non-disease associated NTHi

Non-disease associated NTHi strains have also been investigated in regard to their LPS structure. NTHi strain 11 and 16 [68] have been established to express the same lipid A as the pathogenic strains. The triheptosyl inner core was substituted with GlcI on HepI, α-D-Glc*p* on HepII and β-D-Gal*p* on HepIII. Both strains showed Gly on HepIII and strain 11 showed additionally Neu5Ac. Strain 11 showed up to two *P*Cho extending from α-D-Glc*p* and β-D-Gal*p*. Disialylated glycoforms were identified by CE-

ESI-MS precursor ion scanning which indicated Neu5Ac2Hex3 glycoforms with and without *P*Cho. Strain 16 showed Hex3 glycoforms with up to two *P*Cho. It needs to be mentioned that the established LPS structures do not show any significant difference to the disease causing strains.

1.1.4 The lipopolysaccharide of *H. parainfluenzae*

Early investigation on *H. parainfluenzae* LPS by sodium dodecylsulphate polyacrylamide gel electrophoresis (SDS PAGE) showed that there are rough type or smooth type LPS [89]. Analysis and comparison of *H. parainfluenzae* strains by 16S rRNA accomplished by our colleagues at University of Oxford [DW Hood, unpublished results] demonstrated the O-repeating units to be synthesized by either the Wzy or the ABC dependent mechanisms.

Structural details of *H. parainfluenzae* rough type LPS were described for the first time by Pollard *et.al*. [57]. These authors investigated two clinical isolates of *H. parainfluenzae* (strains 4201 and 4282) as well as a *pgm* mutant of strain 4201. They established for both strains a triheptosyl inner core, L-α-D-Hep*p*-(1→2)-[*P*Etn→6]-L-α-D-Hep*p*-(1→3)-L-α-D-Hep*p* which is linked to lipid A *via* one Kdo. The first heptose is substituted by the outer core starting with β-D-Glc*p*-(1→. This glucose is substituted at O-6 with β-D-Glc*p*-(1→4)-D,D-Hep in strain 4201 and with β-D-Gal*p*-(1→6)-β-D-Glc*p*-(1→4)-D,D-Hep in strain 4282.

Both wildtype strains were found to elaborate exopolysaccharide. *H. parainfluenzae* strain 4201 was found to express a galactan, [→6)-β-Gal*f*-(1→]$_n$, and strain 4282 showed a β-1→6-linked glucopyranosyl polymer [57].

The phosphoglucomutase, Pgm, has been shown to convert glucose-6-phosphate to glucose-1-phosphate in *Neisseria meningitidis* (strain NMB-R6) and *N. gonorrhoeae* (strain 1291-R6) [90]. This step is necessary for incorporation of the activated sugar into the LPS. The same function of Pgm was proposed for NTHi 2019 and *H. parainfluenzae* strain 4201 [57, 91].

The lack of outer core in *H. parainfluenzae* strain 4201*pgm* had severe consequences for the adherence to immortalized human bronchial ephitelial cells. The mutant showed decreased adherence to the ephitelial cells when compared to the wildtype LPS of strain 4201 and NTHi 2019 [57].

Recently, *H. parainfluenzae* strain 20 isolated from a healthy individual, was shown to expresses LPS that contains O-chain polysaccharide. Structural analysis of the inner core showed it to be identical to that in *H. parainfluenzae* strains 4201 and 4282. This strain was shown to have additional outer core β-D-Gal*p* elongating from O-2 of HepIII. The O-repeating unit was shown to be a trisaccharide of [→4)-α-D-Gal*p*NAc-(1→*P*→6)-β-D-Glc*p*-(1→3)-α-D-Fuc*p*NAc4N-(1→] in which FucNAc4N is 2-acetamino-4-amino-2,4,6-trideoxy-galactopyranose and β-linked when substituting O-4 of GlcI [92].

1.1.4.1 LPS –O-Antigen synthesis

LPS biosynthesis in *H. influenzae* is carried out in several steps where the LPS first is synthesized at the inner membrane with the outer core building blocks attached step by step. The complete molecule is then transported through the cell wall to the outer membrane by several proteins including LptA to LptG, see Fig. 7A [93]. O-Antigen synthesis takes place on the cytoplasm site of the inner membrane, see Figure 7. Two mechanisms for the O-antigen assembly have been described, namely the ABC and the Wzy dependent systems.

1.1.4.1.1 ABC-transporter dependent pathway

ATP-binding cassette (ABC)-transporters comprise a protein super family that plays many roles in biological processes and have been investigated in *Escherichia coli* [94]. They consist of cylinder like membrane proteins that, in the case of Gram-negative bacteria, extend to the cell wall. They are not only involved in the export of molecules such as oligo- and polysaccharides, proteins and lipids but also in the uptake of nutrients.

The ABC transporter mechanism is characterized by continuous O-chain synthesis, Fig. 7B, involving the transfer of phosphorylated sugars to the lipid carrier undecaprenyl-phosphate (Und-P). The assembled structures are simpler than those that involve the Wzy pathway. The detailed process is still a matter of investigation.

Briefly, the phosphorylated sugar residues are sequentially assembled onto a primer linked to Und-P. Following assembly, the molecule is exported via a barrel like protein probably consisting of KpsM and TMD, the ABC transporter, through the cell wall to the outer membrane. It is unclear what factors determine the biosynthesis and whether completion of the O-chain synthesis is mandatory [95].

Paper IV elucidates the LPS structure of *H. parainfluenzae* strain 13 that expresses LPS containing O-chain polysaccharide assembled with the ABC dependent pathway.
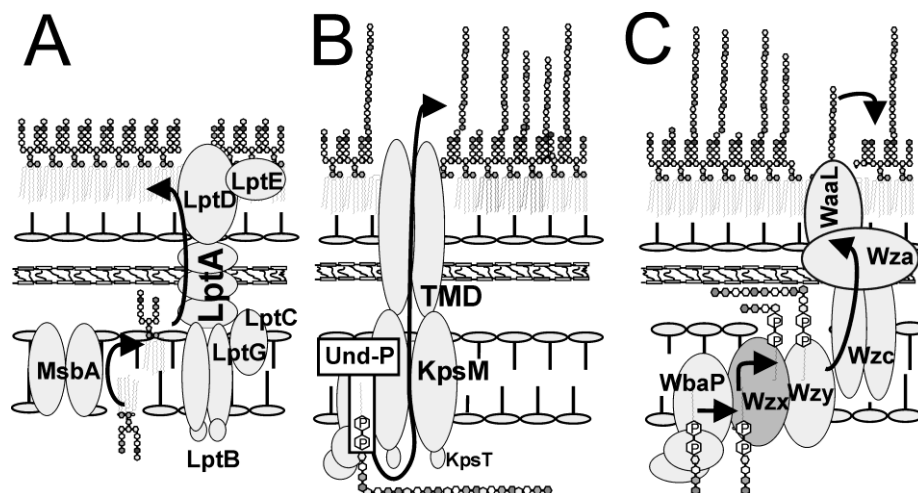

Figure 7: The LPS transport from the inner membrane to the outer membrane. Shown are the current working model according to [94], ovals and rectangles represent the same molecules as in Fig. 2. A) The transport mechanism for rough type LPS, B) the ABC dependent LPS transport mechanism and C) the Wzy dependent mechanism.

1.1.4.1.2 The Wzy dependent mechanism

The first step in the Wzy dependent mechanism has been identified in a number of bacteria. In the beginning a phosphorylated sugar is transfered to the lipid carrier Und-P. In some bacteria, e.g. *Salmonella enterica*, the first sugar transferred to Und-P is galactose-1-phosphate (Gal-1P) catalyzed by WbaP [96]. In other bacteria, e.g. *E. coli*, WecA is responsible for the addition of *N*-acetylglucoseamine-1-phosphate to Und-P [97]. This provides the primer on which the O-repeating unit is synthesized. That unit is then transported across the membrane by a Wzx-translocase, presumably polymerized by Wzy and transferred to the outer membrane by a Wzc, Wza. Thereafter O-chain is transferred to the carrier (Lipid A-coreOS) by a conjugating enzyme, WaaL, Fig. 7C [94].

The LPS of *H. parainfluenzae* strain 20 is accomplished according to the Wzy dependent mechanism as suggested by genetic analysis. This was confirmed by structural analysis of a *H. parainfluenzae* strain 20*wbaP* mutant [92]. Structural analysis evidenced the β-D-Glc*p*-(1→ and β-D-Gal*p*-(1→ substituted triheptosyl inner core moiety that has been also detected in the wildtype core OS. The material lacked the O-antigen and confirmed that WbaP is involved in the Wzy dependent mechanism of the biosynthesis of O-repeating unit [92].

2 Aims of this Thesis

The scope of this thesis was to determine the detailed LPS molecular structures expressed by several strains of *H. influenzae* and *H*. *parainfluenzae*. Full understanding of the LPS biosynthesis in *H. influenzae* demanded establishing of the function of the gene *lic2B*. In order to further advance the understanding of possible LPS structures expressed by NTHi, two clinical isolates were also investigated.

H. *parainfluenzae* can express rough type LPS and LPS including the O-repeating unit. With the recent determination of the genome of *H*. *parainfluenzae* strain T3T1 it became possible to elucidate LPS profiles from commensal *H*. *parainfluenzae* and to compare them to NTHi. Two *H*. *parainfluenzae* strains with rough type LPS were investigated. In addition, one *H*. *parainfluenzae* strain expressing an O-antigen was analyzed.

Structural elucidation of LPS involved initial mild acid hydrolysis to obtain water-soluble material that was suitable for subsequent analyses by chemical, nuclear magnetic resonance (NMR) and mass spectrometric (MS) methods.

The following bacterial strains were investigated:

   I.   *Haemophilus influenzae* strain Eagan*lic2B*lic2C+
  II.   Non-typeable *Haemophilus influenzae* strains 1247 and 1008
 III.   *Haemophilus parainfluenzae* genome sequence strain T3T1 and strain 22
 IV.   *Haemophilus parainfluenzae* strain 13

3 Materials and Methods

This chapter gives an introduction to the structural analysis of LPS. There are several aspects to be covered:

A. The identity of the single residues that build up the oligosaccharide or the repeating unit of O-polysaccharide,
B. The presence of pyranoses, furanoses or open ring form,
C. The anomeric configuration,
D. The absolute configuration,
E. The positions of glycosidic linkages,
F. Branching points and sequence,
G. The position of non-carbohydrate modification.

3.1 Bacterial cultivation and LPS isolation

Bacterial cultivation and construction of mutant strains were accomplished by DW Hood and his group at University of Oxford, UK. NTHi strains were grown in brain-heart infusion (BHI) broth supplemented with haemin (10 µg·ml^{-1}) and nicotinamide adenine dinucleotide (2 µg·ml^{-1}). *H. parainfluenzae* was grown in the same medium as NTHi without added haemin, since it is not required for growth. The bacteria were grown overnight at 37°C until late logarithmic phase and were then harvested. Alternatively, *H. parainfluenzae* strain T3T1 and *H. parainfluenzae* strain 22 were grown on solid BHI medium (agar 1% w/v) supplemented with 10% Levinthals reagent. After growth over night at 37°C, the bacteria were harvested from the plates in a combined volume of about 50 ml of PBS with 0.5% phenol then were pelleted by centrifugation. Following growth on either liquid or solid medium, the cell pellet was resuspended and sequentially washed with ethanol, acetone and ether prior to being lyophilized.

Mutant strains were created by transformation with plasmid constructs containing the respective cloned and inactivated gene [98]. *H. influenzae* strain Eagan*lic2B*lic2C+ was additionally transformed with a self-replicating plasmid containing a functioning *lic2C* gene. Successful transformation was confirmed by Southern analyses and PCR amplification [98, 99]. For mutant strains, kanamycin, 10 µg·ml^{-1}, was added to the growth medium.

LPS extraction for rough type LPS is accomplished by the phenol-chloroform- light petroleum method as described by Galanos *et. al.* [100]. To a freshly made mixture of phenol/chloroform/ light petroleum (PCP, 2:5:8, v/v/v) crystalline phenol was added until the solution was clear. Then the PCP solution was added to the lyophilized bacteria and the mixture stirred for 24h in the dark. The bacteria and the solution were separated by centrifugation (6 100 g, 4°C, 10 min). The bacteria pellet was again suspended in PCP and stirred under the same conditions. The solution phases were combined and chloroform and light petroleum were evaporated. The method was modified by precipitating the LPS using

diethyl ether/acetone (1:5, v/v; 6 vol). LPS was purified by ultracentrifugation (82 000 g, 4°C, 16h) and finally lyophilized.

For smooth type LPS the hot water phenol method was employed [101]. Briefly, the bacteria were dissolved in hot water (70°C) and the same volume of 90% aq. phenol was added. After stirring for 1 h, the mixture was cooled in an ice bath. Both phases were separated by centrifugation (6 100 g, 4°C, 20 min). The water phase was retained and new water was added to the phenol solution. Again the mixture was stirred for 1 h and centrifuged. The interphase was rejected and the combined water phases and the phenol phase were separately dialyzed against running tap water (3 days) and running deionized water (1-2 days) in order to remove the phenol. The remaining materials were purified by enzymatic digestion with DNase, RNase and proteinase K. Further purification was achieved by ultracentrifugation (82 000 g, 4°C, 16h).

3.2 Derivatization

The amphiphilic nature of LPS is due to the hydrophobic lipid and the hydrophilic carbohydrate parts of the molecule. Therefore, LPS is only sparsely dissolvable in both aqueous and organic solvents and easily forms micelles in the former. Analysis of native LPS is a challenge, but has been achieved when applying capillary electrophoresis (CE) electrospray ionization mass spectrometry (ESI-MS) [102] and when the LPS was hydrophilic due to a large O-antigen allowing NMR investigations [92].

Thus, structural work on LPS includes routinely partial cleavage of the fatty acids by O-deacylation or total removal of the lipid A to afford water soluble oligo- and polysaccharides components.

3.2.1 O-deacylation

For removal of all ester linked fatty acids, dry LPS was dissolved in anhydrous hydrazine and stirred for 1 h at 37°C [103]. Then the sample was placed in an ice bath and the excess hydrazine was destroyed by adding chilled acetone. The O-deacylated LPS (LPS-OH) was allowed to precipitate for 1h, washed twice with acetone and finally lyophilized.

3.2.2 Mild acid hydrolyses

Release of core oligosaccharide, OS, was achieved by hydrolysis of LPS with dilute acetic acid (pH 3.1, 100°C, 2h). The ketosidic Kdo linkage between lipid A and OS was cleaved while other aldosidic linkages were not hydrolysed due to higher acid stability. All other acid labile substituents such as *PP*Etn and keto sugars including adjacent Kdo`s and sialic acids were also released.

Several forms of anhydro Kdo (*An*Kdo-ol) were formed due to β-elimination of the phosphate group at O-4 of Kdo [104, 105]. The amount of resulting *An*Kdo-ol forms was decreased by reduction with borane-*N*-methylmorpholine which terminates the keto-enol-tautomerism between C-2 and C-3 (see Figure 8) [106].

For purification the insoluble lipid A was removed from the mixture by centrifugation (6 100 g, 4°C, 10 min). The crude OS was separated from salts by

chromatography as gel permeation chromatography (GPC) or via liquid chromatography-MS (LC-MS).

Figure 8: The formation of *An*Kdo-ol upon mild acid hydrolysis of LPS with simultaneous reduction. For simplification the ring protons are not shown.

For the determination of sialic acid, the LPS or LPS-OH was treated with 0.1 M HCl (aq.) at 80°C for 1h and analyzed by high performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD) [79]. Alternatively, sialic acid was released with 1 M propionic acid [107], lyophilized, pertrimethylsilylated and analyzed by GC-MS. Challenges of sialic acid analyses included fine balancing of the reaction conditions for complete cleavage of the sialic acid while minimizing *O*-acetyl migration and loss. Hydrolysis with propionic acid has the advantage of low *O*-acetyl group migration and loss in comparison to treatment with HOAc [107, 108].

3.2.3 Dephosphorylation

Removal of all phosphate bound substituents was accomplished by treatment of LPS-OH or OS with 48% hydrofluoric acid at 4°C for 48 h to give LPS-OH-P or OS-P. The acid was removed under a stream of air in an ice bath and the remaining material was lyophilized. The use of low temperature is obligatory, so that the glycosidic linkages are not hydrolysed.

3.2.4 Methylation

Methylation of the available hydroxyl groups on a sugar residue is a common derivatization procedure. The critical step is the complete methylation of all available hydroxyl groups, as the results might be misleading otherwise. Therefore, the whole sample has to be dissolved or at least dispersed in a suitable solvent such as dimethyl sulfoxide (DMSO), to avoid partial methylation. Hence, the use of dry DMSO (with molecular sieves to capture H_2O) and purging of the solution with N_2 (g) to remove O_2 and CO_2 is necessary. In order to ensure solvation of LPS-OH, prior to methylation the sample is peracetylated using acetic anhydride and 4-dimethylaminopyridine (21°C, 4h).

In a first step of methylation, the hydroxyl groups of the sugars were deprotonated to form carbanions. This was achieved by adding a strong base such as butyl lithium to

16

DMSO. This gave rise to lithium methylsulphinyl carbanion at 40 °C. Methylation was achieved by addition of methyl iodide (21°C, 16h). [109, 110].

The methylated compounds were recovered on a C-18 SepPak column. As picturized in Figure 9, the sample was then either analyzed directly with LC-ESI-MS^n or subjected to the procedure of the methylation analysis (see below).

During methylation the pH is typically > 7 which leads to the loss of all esterbound *O*-acetyl and glycine groups. Methylation under neutral conditions has been described by Prehm [111]. With trimethyl phosphate as solvent 2,6-di-tert-butylpyridine acted as a proton scavenger and allowed methylation by trifluoromethanesulphonate. In an attempt to rescue the *O*-acetyl groups and determine their location via LC-ESI-MS^n, this method was employed. However, in our hands the yields were not sufficient to give interpretable amounts.

3.3   Chemical Analysis
3.3.1   Compositional Analysis

In order to establish the monosaccharide composition of LPS, the material was hydrolysed (2M triflouroacetic acid (TFA), 120 °C, 2h), reduced with sodium borohydride (NaBH$_4$) in 500 µl 1M ammonia solution (16h, room temperature) and acetylated with pyridine/acetic acid anhydride (90 °C, 30 min) as described before [112]. After extraction of the acetylated alditol derivatives, the sample was analyzed by GC-FID and GC-MS (Figure 9).

The acid labile Kdo is phosphorylated in *H. influenzae* and *H. parainfluenzae*, which hinders its detection by GC-MS. However, once dephosphorylated, Kdo can be determined by methylation analysis or the thiobarbiturate assay [113].

3.3.2   Methylation Analysis

Identification of the substitution site of the oligosaccharide residues was accomplished by methylation analysis. After tagging of all free hydroxyl groups with methyl groups, the oligosaccharide was hydrolyzed and reduced with NaBD$_4$, as illustrated in Figure 9. The use of a deuterated reducing agent allowed for the identification of the aldolidic carbon. The remaining free hydroxyl groups (from the substitution site and the ring formation) were acetylated and the partially methylated acetylated alditol derivates (PMAA) were analyzed by GC-MS. Retention times and characteristic fragmentation patterns gave information about substitution sites.

A comparison of the methylation analyses of native and dephosphorylated material allows for the identification of phosphorylated sugars as e.g. HepII of the inner core in NTHi.

3.3.3   Absolute Configuration

Determination of the D or L configuration of the sugars was accomplished by glycosidation of the monosaccharides with *S*-2-(+)-butanol according to Gerwig *et. al.* [114]. In contrast to the use of 2-(+)-octanol this procedure has the advantage of giving better results for hexoseamine sugars.
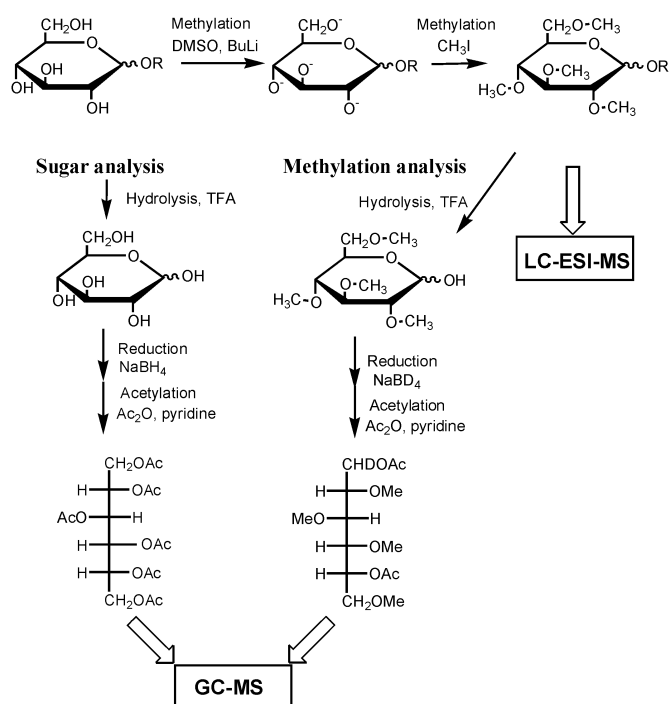
Figure 9: Overview on the example of glucopyranose being subjected to sugar analysis, or methylation after which the material either is analyzed by LC-ESI-MS or subjected to methylation analysis.

3.4   Mass Spectrometry

Mass spectrometry (MS) allows the determination of the molecular mass. In this thesis two ionization techniques were used: the "soft" electrospray ionization (ESI) and "hard" electron impact ionization (EI, used in GC-MS). EI induces fragmentation and gives rise to mass spectra that are used to identify compounds in sugar and methylation analysis.

ESI-MS allows for the detection of whole molecular ions of the respective glycoforms. In ESI the sample was typically dissolved in a mixture of water and an organic solvent such as acetonitrile or methanol in order to ensure proper spray formation. Further, some additives such as ammonia or sodium acetate were added in order to facilitate ionization. Multiple charged species were generally observed. Native and derivatized LPS materials were analyzed by ESI-MS. From the protonated molecular ions $[M+nH^+]^{n+}$ or the sodium adduct ions $[M+nNa^+]^{n+}$ the molecular mass can be calculated from which the molecular compositions can be proposed.

In tandem MS specific molecular ions are chosen and fragmented to give detailed structural information of the glycoforms. The resulting fragments reveal information about substitution sites of sialic acid, non carbohydrate substituents and sequence. For increased sensitivity precursor ion scanning and/or derivatization is applied. In Figure 10 the several techniques and instrument combinations are listed.
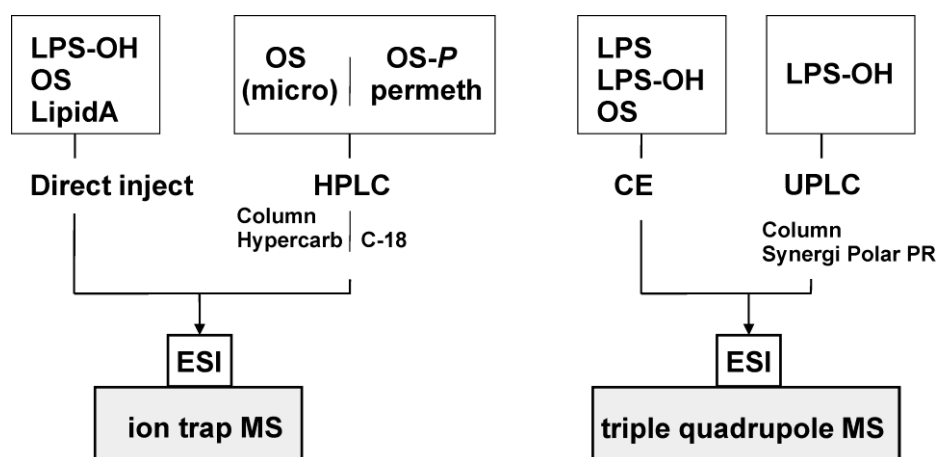
Figure 10: Overview of the mass spectrometric methods employed.

3.4.1 LC-ESI-MS^n

H. influenzae expresses a great variety of different glycoforms that have to be unraveled for the structural analysis. Important aspects are the sequence, *i.e.* the order in which the different building blocks are attached to each other, and the position of branching points. Dephosphorylated and permethylated OS was analyzed by LC-ESI multi step tandem MS (MS^n) after separation on a C-18 column. The solvent consists of water/methanol with 0.1mM sodium acetate [115].

The analysis of dephosphorylated and permethylated OS in LC-ESI-MS^n has been proven to be sensitive method, due to the hydrophobic properties of permethylated OS which increases the rate of ionization in ESI. The sodiated adduct ions were easily detected and allowed for the precise structural analysis of isomeric glycoforms and trace amounts as for example in NTHi strain 1247 (**Paper II**). The samples were analyzed by LC-ESI-MS^n after separation on a C-18 column. The solvent consisted of water/methanol with 0.1mM sodium acetate [115].

For example the MS^2 mass spectrum of the HexNAcHex4Hep3Kdo glycoform at *m/z* 2121 from NTHi strain 1247 is shown in Figure 11A. The OS structures were determined using the mass differences induced by methyl tagging of all free hydroxyl groups. As a result terminal residues differed from a substituted or branched residue by increments of 15 Da corresponding to one or more methyl group. The resulting fragments in Fig. 11A originated from two isomeric glycoforms shown in Fig. 11B. One glycoform was found to consist of a terminal (t) HexNAc-Hex-Hex-Hex- elongation from HepIII. This glycoform showed also the loss of tHex, *m/z* 218, and resulted in the fragment ion at *m/z* 1903. The second isomer, the HexNAc containing extension elongated from GlcI. This glycoform resulted only in fragments from the loss of terminal (t) HexNAc, *m/z* 259, and substituted Hex, *m/z* 204.

For unambiguous structural determination, more information was gained using a MS^3 on the ion *m/z* 1611 corresponding to the loss of the tHepIII-HepII, see Fig. 11C. The resulting fragment ions confirmed the tHexNAc-Hex-Hex-Hex-Hex elongation from HepI. No loss of tHex could be observed.
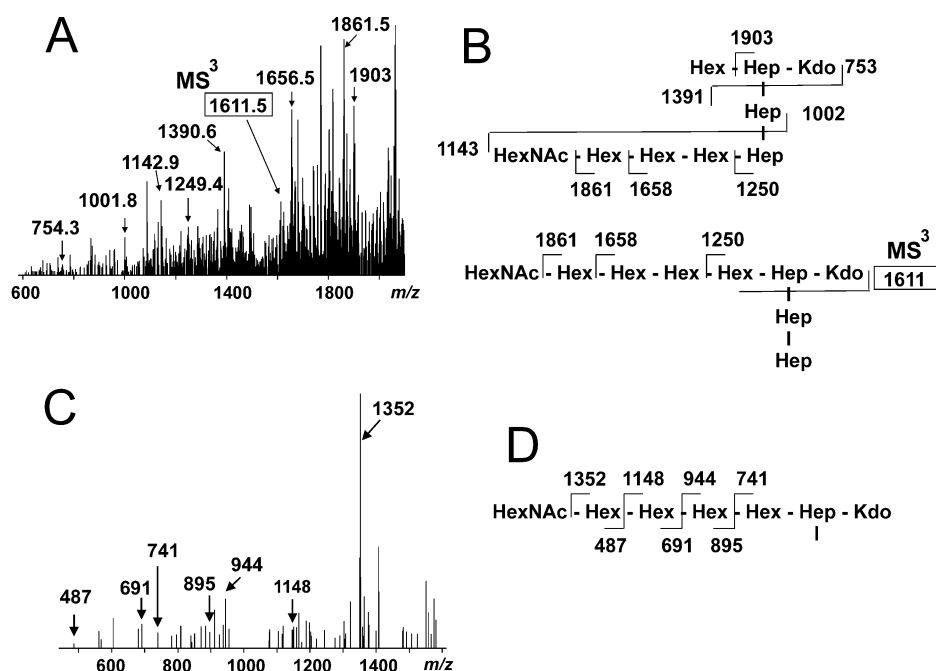
Figure 11: LC-ESI-MS^n analysis of dephosphorylated and permethylated core OS from NTHi strain 1247. A) The MS^2 spectrum of the HexNAc•Hex$_4$•Hep$_3$•Kdo glycoform, m/z 2121, B) the two isomeric glycoforms contributing to the MS^2 spectrum, C) MS^3 on m/z 1611 confirming the HexNAcHex4 extension from GlcI and D) the proposed structure.

The use of graphitized carbon columns in LC-ESI-MS has been described before as a useful method for oligosaccharide analysis [116]. Some applications even allowed for the separation of isomeric glycoforms as reported for *N*-linked oligosaccharides from murine kidneys [117]. We adopted the method in an attempt to meet the limited amounts of sample. In a micro method, LPS (1-3 mg) was treated with diluted acetic acid with simultaneous reduction by borane-*N*-methylmorpholine complex. The precipitated lipid A was removed by centrifugation and the OS containing supernatant was lyophilized. The sample was dissolved in water and directly introduced to the LC-MS. Salts were separated from OS using a Hypercarb carbon column with a gradient of water/acetonitrile with 5mM ammonia. As the salts eluted first and were let into the waste, the later eluting sample was led into the ESI-MS. Longer retention times for larger glycoforms were observed. However, separation of isomeric glycoforms was not successful. Using this approach the required sample amount was reduced by approx. 10 times, since purification of OS by GPC was not necessary.

Identification of substitution locations of non carbohydrate substituents can be achieved by LC-ESI-MS^n on OS. This method was applied on core OS from NTHi strain 1247*lpsA*. The mass spectrum resulted in doubly and singly charged ions as shown in Figure 12A. The doubly protonated ion m/z 686.2 corresponding to the glycoform composition of Ac$_2$•PCho•Hex$_1$•Hep$_3$•PEtn•AnKdo-ol was analyzed in MS^2. The resulting mass spectrum is shown in Fig. 12B. Marker ions at m/z 234.7, and 550.0 corresponding to Ac-Hep and Ac-Hep-Hep-PEtn, allowed for the assignment of an *O*-acetyl group at HepIII. The ion m/z 369.9 was indicative of acetylation at the PCho bearing GlcI.

Care has to be taken to rule out potential rearrangements that can occur on protonated parent ions [118]. Here, the use of the sodiated adduct ions has been shown to delivered stable fragments [119].

Comparative analysis of OS and LPS-OH in ESI-MS allowed the identification of acid or base labile moieties such as sialic acid and *O*-acetyl groups.
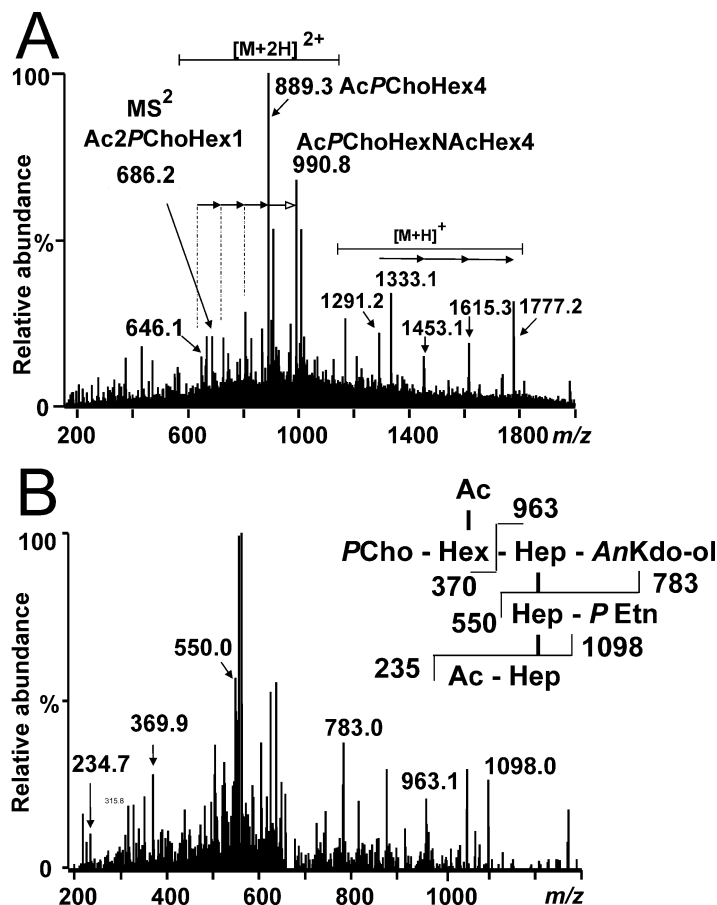
Figure 12: LC-ESI-MS^n analysis of core OS from NTHi 1247*lpsA*. A) The mass spectrum observed in positive mode and B) MS^2 analysis of the glycoform Ac$_2$•*P*Cho•Hex$_1$•Hep$_3$•*P*Etn•*An*Kdo-ol, *m/z* 686.2.

3.4.2 Lipid A and fatty acid distribution

The mass spectrometric analysis of lipid A provided information about the fatty acid substitution pattern. In a first step lipid A obtained following the mild acid hydrolysis of the LPS was purified by extraction with water/methanol/chloroform (1:1:2, v/v/v). The lower organic phase was saved and the solvent evaporated. The purified lipid A was then dissolved in methanol/chloroform (1:1, v/v) and directly injected into the ESI-ion trap MS. Tandem MS experiments ranging from MS^2 to MS^4 were conducted in order to elucidate the complete structure as described before [27].

In order to facilitate structural determination the results from fatty acid analysis were used to confirm the identity of the present fatty acids.

3.4.3 GC-MS

In the GC-MS analysis, the retention time and the MS fragmentation pattern were used for identification of the analytes. The sample was derivatized by acetylation or

pertrimethylsilylation as in sugar analysis, methylation analysis, the determination of the absolute configuration, and fatty acid methyl esters were analyzed by GC-MS.

Next to the retention time, specific mass spectra reveal the identity of the sample compound. As illustrated in Fig. 13A on the example of pertrimetylsilylated Neu5Ac of *H. parainfluenzae* strain 22. The fragment ion *m/z* 624 was indicative for 5-*N*-acetyl-neurmainic acid. In contrast, *H. parainfluenzae* strain T3T1 (Fig. 13B) resulted in fragment ions at *m/z* 694, 536 and 175 indicating *O*-acetylation at O-9 of sialic acid (**Paper III**).
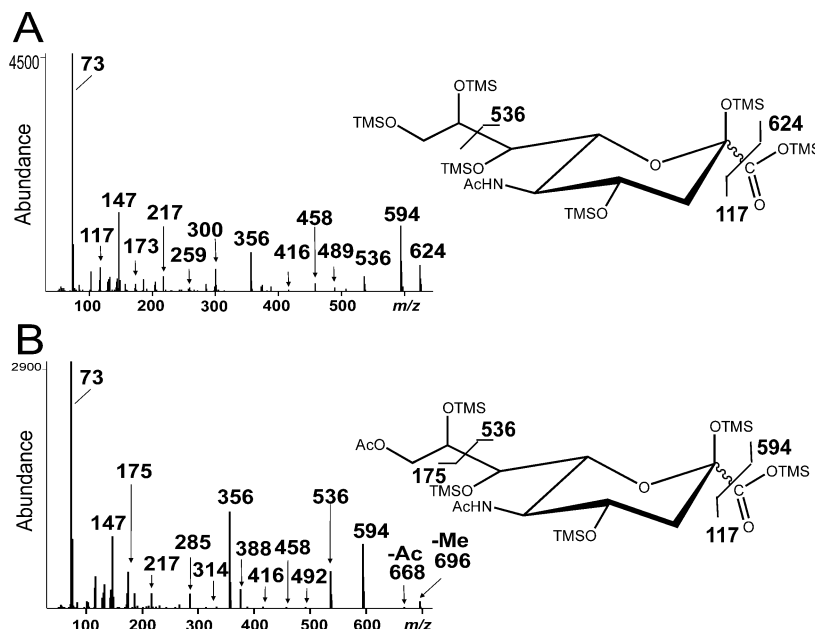
Figure 13: Extracted mass spectra of the pertrimetylsilylated sialic acid derivatives of LPS from two *H. parainfluenzae*. Shown is A) 5-*N*-acetyl-neuraminic acid of strain 22 with a retention time of 27.7 min and B) O-9 acetylated 5-*N*-acetyl-neuraminic acid of strain T3T1 with a retention time of 28.7 min.

3.4.4 CE-ESI-MS

Capillary electrophoresis (CE) coupled to triple quadrupol ESI-MS has been shown to provide a useful tool for multiple analyses of LPS and its derivatives [85, 88, 120]. CE separates the sample from impurities and gives therefore increased sensitivity. For CE a polymicro capillary column (1m x 375 μm x 50 μm) was used. Typically the sample is dissolved in water. This sensitive method needed only minimal sample amounts using sample volume of 1μl. Separation of the isomeric glycoforms before MS/MS investigation had been achieved as e.g. for NTHi strain 375*lic1*, where the position of *P*Etn could be determined to be on HepIII and *P*Kdo, respectively [121].

CE-ESI-MS/MS has been shown to be essential in particular for the analysis of sialylated glycoforms which are important virulence factors for NTHi and only expressed in trace amounts. The CE-ESI-MS spectrum of LPS-OH from NTHi strain 1247*lpsA* in negative ion mode is shown in Fig. 14A (**Paper II**). The major glycoforms are easily detected as doubly, triply and quadruply deprotonated ions. Precursor ion scanning in negative ion mode using *m/z* 291 (Neu5Ac) or *m/z* 274 (Neu5Ac-H$_2$O) gave sensitive detection of sialic acid bearing glycoforms. This is demonstrated on LPS-OH from NTHi strain 1247*lpsA*. The precursor ions scanning on the ion *m/z* 290, i.e. sialic acid, gave rise

to the resulting mass spectrum that is shown in Fig. 14B. Mono- and disialylated glycoforms were easily detected.
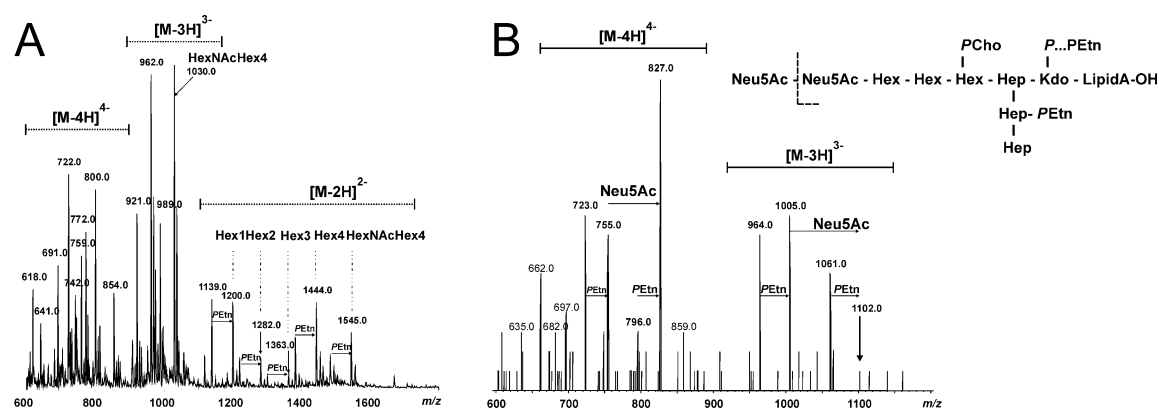
Figure 14: CE-ESI-MS analysis of LPS-OH from NTHi strain 1247*lpsA*. Shown is in A) the mass spectrum in negative ion mode and B) the precursor ion scanning on *m/z* 290, i.e. sialic acid. The proposed structure is indicated in the inset.

The analysis by tandem MS experiments (CE-ESI-MS/MS) gives structural information on the substitution location of non carbohydrate substituents and sialic acid. Glycoforms of interest were further fragmentated by allowing for the detection of specific marker ions e.g. *m/z* 494 and 453 originating from Neu5Ac-HexNAc and Neu5Ac-Hex, respectively as observed for *H. parainfluenzae* strain T3T1 and 22 [122] (**Paper III**).

As one of only a few methods, CE-ESI-MS can be applied on intact LPS as described before for *Campylobacter jejuni* [123]. LPS from *H. parainfluenzae* strain T3T1 has been investigated and precursor ion scanning on *m/z* 332, i.e. Ac-Neu5Ac, gave rise to one molecular ion at *m/z* 1331.5, see Figure 15A. The MS/MS experiment on *m/z* 1331.5, i.e. Ac•Neu5Ac•dHexNAcN•HexNAc•Hex$_2$•Hep$_3$•Kdo•*P*•*P*Etn$_2$•LipidA, allowed for assignment of the *O*-acetyl group that substitutes sialic acid in its outer core, Fig. 15B (**Paper III**).
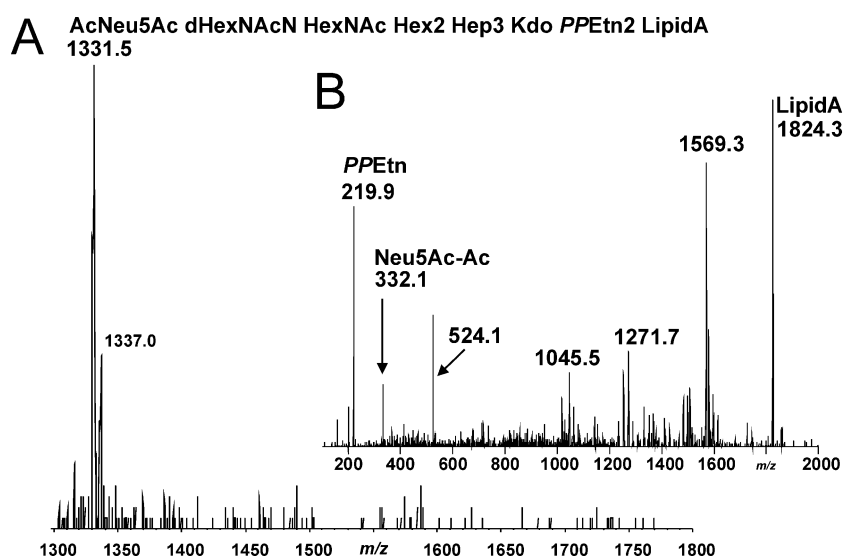
Figure 15: CE-ESI-MS/MS analysis on native LPS from *H. parainfluenzae* strain T3T1. A) The mass spectrum resulting from precursor ion scanning on the ion *m/z* 332, i.e. Ac-Neu5Ac, in negative ion mode and B) the MS/MS spectrum of the acetylated LPS molecule, *m/z* 1331.5, indicating *O*-acetylation on Neu5Ac.

The substitution location of non carbohydrate substituents was obtained for core OS in positive mode giving insite in Ac, Gly, *P*, *P*Cho, *P*Etn substitution [88, 124](**Paper III**).

3.5 Nuclear Magnetic Spectroscopy

NMR experiments were accomplished on core OS or LPS-OH dissolved in D_2O or D_2O with 2mM perdeutero-EDTA and perdeutero-SDS, 10 mg·ml^{-1}, respectively. In order to exchange the hydroxyl group protons with deuterium the samples were lyophilized twice with D_2O. Chemical shifts (δ) are reported in ppm, by referencing to standard compounds using as internal standard sodium 3-trimethylsilylpropanoate-*d*4 (δ_{1H} 0.00), or as external standard acetone in D_2O (δ_{13C} 31.0) or 85% phosphoric acid in D_2O (δ_{31P} 0.00). Depending on the position of signal of interest, ^1H NMR spectra were recorded at a temperature so that the water signal did not obscure it. Typically, spectra were recorded at 25°C or 22°C.

3.5.1 One dimensional NMR experiments

The simplest NMR experiment is a one dimensional (1D) spectrum as shown in Figure 16 on the example of a proton spectrum of core OS from NTHi strain 1247*lpsA*. Since NMR is rather insensitive, the same experiment is run in a set of repeats and all the gained signals are added up in order to increase the signal to noise ratio.

The 1D proton NMR spectrum of core OS can be divided in several regions (see Fig 16). Methyl group protons of deoxy sugars are found around 1.5 ppm while the methyl protons from acetate groups resonate around 2 ppm. If present, *P*Cho methyl protons give a strong signal at approx. 3.28 ppm. The ring proton region is recognizable as the bulky region ranging from approx. 3.0 to 4.2 ppm. Most information can be gained from the anomeric protons that give rise to signals in the region from 4.3 to 6.0 ppm. This region can roughly be further divided in to the region of α pyranosidic anomeric of *gluco/galacto* configurated protons that resonate from approx 4.8 ppm to 6.0 ppm and the β pyranosidic anomeric protons that resonate from 4.3 to 4.8 ppm. Furanosidic anomeric protons are not as easily generalizable.
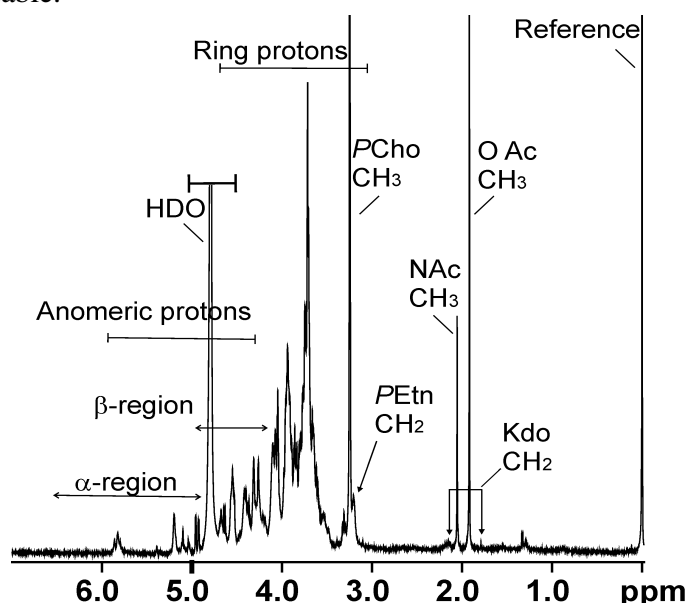
Figure 16: The 1D ^1H spectrum of core oligosaccharide from NTHi strain 1247*lpsA* run at 25 °C.

Most of the information in a 1D proton spectrum is gained in the anomeric region, e.g. the signal intensity of the anomeric proton that can be used to estimate the ratio of the sugars in the sample.

Due to scalar coupling, *J*, the chemical shift of a specific proton can appear as doublet, triplets etc. In proton NMR, *J* reveals information about the stereochemistry of two protons, as e.g. in the case for anomeric protons that couple to the respective H2. Depending on the torsion angle the coupling constant can range from about 8 Hz for β pyranosidic anomeric protons, to ~3 Hz for α pyranosidic anomeric protons of D-glucose (Table 1). In *manno* configurated sugars, e.g. L,D-Hep from the inner core of *H. influenzae* and *H. parainfluenzae*, the dihedral angle between H1 and H2 is so small that the coupling is typically not resolved. The dependence between the torsion angel and the ^3J is illustrated by the Karplus curve [125].

Table 1: Scalar coupling constants of some common hexoses, adopted from [126].

	$^3J_{\text{H-1-H-2}}$	$^1J_{\text{H-1-C-1}}$
	Pyranose	Pyranose
α-Glc	~3	~173
β-Glc	~7	~165
α-Man	~1-2	~172

As an example the anomeric protons of the core oligosaccharide from NTHi strain 1247*lpsA* is shown in Figure 17. H1 and the vicinal proton H2 are connected via three covalent linkages. They can exchange magnetization with each other resulting in doublets. The unresolved coupling from the L-D-Hep*p* residues (Hep) in contrast to the smaller coupling from α-Gal*p* and the larger coupling constant from β-Gal*p*NAc is easily visible.

The strong overlap of the ring protons does not allow for the complete chemical shift assignment which is why two dimensional (2D) experiments are regularly applied (see page 26).
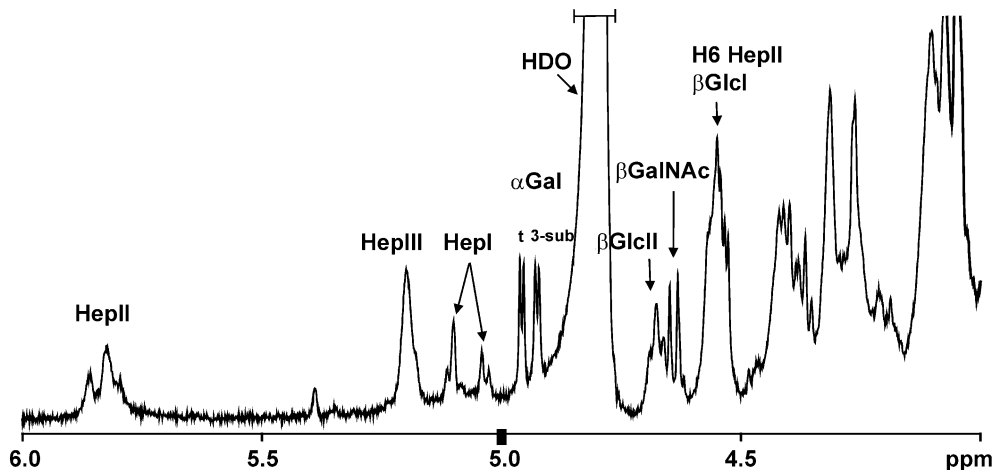
Figure 17: Zoom on the anomeric region of the 1D ^{1}H spectrum of OS from NTHi 1247*lpsA*. The α-Gal*p* residue revealed two spin systems, one for terminal (t) α-Gal*p* and one from 3-substitutetd (3-sub) α-Gal*p*.

1D NMR spectra on carbon, ^{13}C, can also be acquired. Due to the low natural abundance of carbon these experiments need more scans and a higher sample amount when compared to 1D ^{1}H spectrum. Since ^{13}C resonates over a higher spectral width, than ^{1}H, the information gained from these experiments can be more informative. Figure 18 illustrates the 1D ^{13}C NMR spectrum of O-antigen containing OS from *H. parainfluenzae* strain 13 (**Paper IV**). The chemical shift data reveals information about the number of anomeric carbons, that resonated in an area from ~96 -110 ppm. Carbonyl carbons were detected at ~180 ppm. Carbons that were part of methyl groups resonated at around 15- 23 ppm. Ring carbons were detected in an area ranging from approx. 60 to 72 ppm. The chemical shift of a carbon was downfield shifted when it was part of an elongation site. Nitrogen bearing carbons from *N*-acetyl-hexoseamines were detected around 50-55 ppm.

In order to gain information about which proton was connected to a specific carbon, two dimensional NMR experiments were carried out.
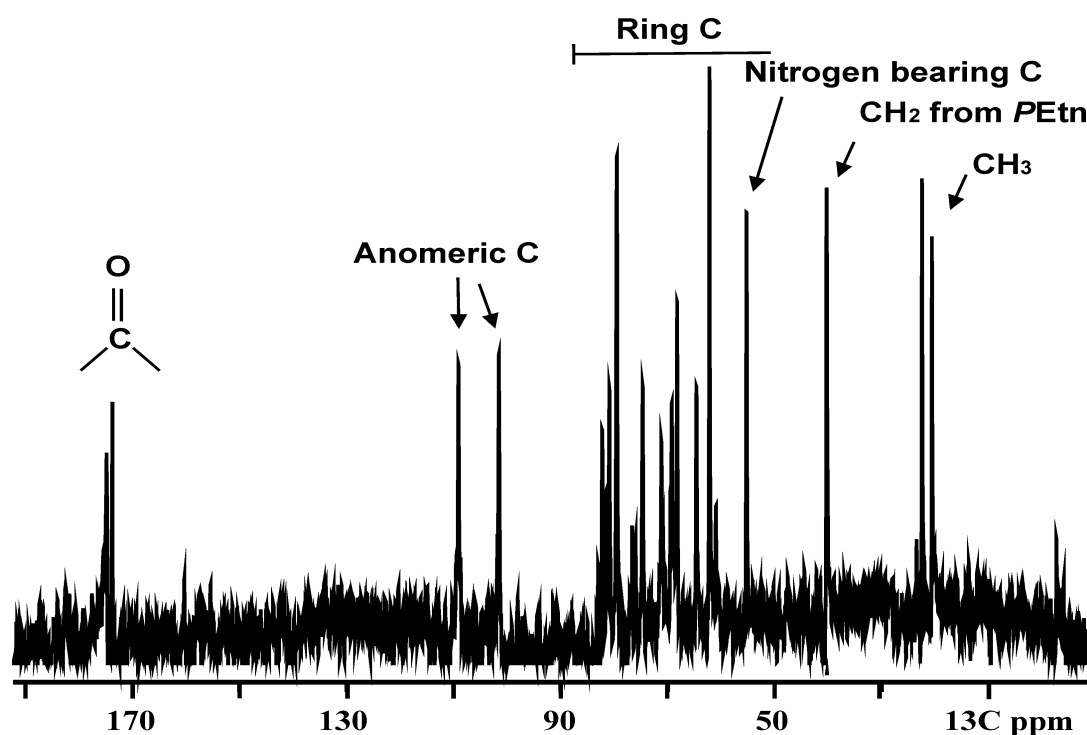
Figure 18: 1D ^{13}C NMR spectrum of O-antigen material from *H. parainfluenzae* strain 13 obtained after mild acid hydrolysis.

3.5.2 Two dimensional NMR experiments

As shown in Figure 16 the chemical shifts from ring protons in oligo- and polysaccharides overlap excessively. Therefore, 2D NMR experiments are needed in order to unravel specific spin systems. Delineation of a single spin system starts with determination of the anomeric proton in the 1D experiment. In the correlation spectroscopy (COSY) experiments vicinal protons exchange magnetization. A selected part of the resulting spectrum is shown in Fig. 19A on the example of *H. parainfluenzae* strain 22 (**Paper III**). Clearly the cross peak between H1 and H2 of β-D-Fuc*p*NAc4N, β-D-Glc*p*II and β-D-Glc*p*I are recognizable. Other cross peaks from vicinal proton as H6

from HepII and H5 and $H7_{A/B}$, respectively, are also observed. However, again the overlap might hinder full assignment which is solved by employing the total correlation (TOCSY) experiment.

The TOCSY experiment includes a mixing period where the magnetization is allowed to be transferred through the spin system. Depending on the length of the mixing time (50-180 ms) and the scalar coupling the whole spin system can be revealed. As shown in Fig 19B on the GlcII example from core OS of *H. parainfluenzae* strain 22 (**Paper III**), glucose reveals typically its whole spin system at a mixing time of 180 ms ranging from H1 to H6 $_{A/B}$. Obstacles in the assignment of the proton spin systems can be given by overlap of the ring protons as in the case of the GlcI residue in Fig. 20B. In *galacto* configured residues as FucNAc4N an unfavorable small coupling constant of ~1 Hz between H4 and H5 will prevent transfer of magnetization (Fig. 19B). Hence, even under long mixing times the spin system will not be indicated further than from H1 to H4, which makes the complete assignment of the spin system difficult.
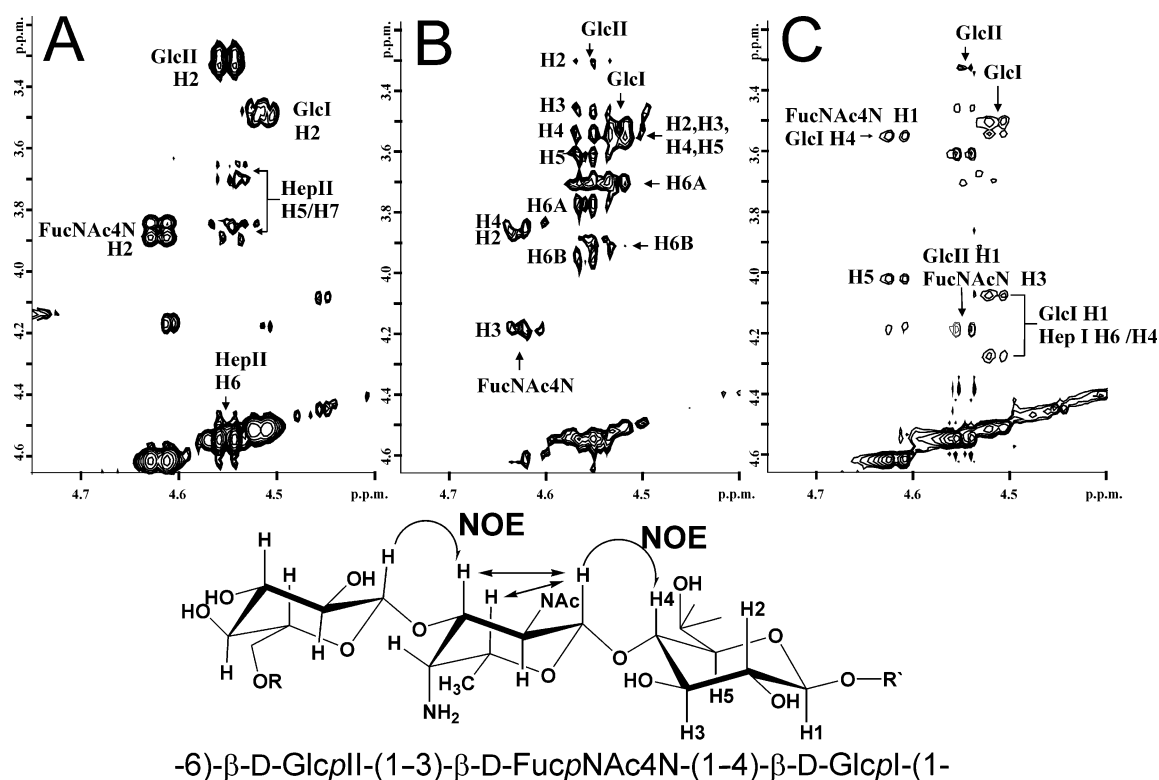
Figure 19: Selected regions of 2D NMR experiments on core OS from *H. parainfluenzae* strain 22. A) the DQF COSY, B) TOCSY, mixing time 180 ms and C) NOESY spectrum, mixing time 250 ms. The structure in the inset illustrates the selected part of the OS molecule with the intermolecular NOE connectivities displayed as round arrows and the intramolecular NOE connectivities depicted as straight arrows.

Once the whole proton and carbon spin system is assigned the order of the residues can be delineated using nuclear Overhauser spectroscopy (NOESY). This was used in **Paper III** and is illustrated in Fig. 19C. Nuclei that are close in space, 3-5 Å, can under these experimental conditions exchange magnetization and give rise to a cross peak. This information is used in order to delineate the sequence as shown in Fig. 19C. The anomeric protons are only close to the proton of the respective attachment site of the reducing

residue, e.g H1 of β-D-Glc*p*II reveals a cross peak with H-3 of β-D-FucNAc4N and H1 of β-D-Fuc*p*NAc4N with H4 of β-D-Glc*p*II. Hence, the sequence is →6)-β-D-Glc*p*II-(1→3)-β-D-Fuc*p*NAc4N-(1→4)-β-D-Glc*p*I-(1→.

NOE connectivities exist also between protons inside the sugar residue. Intramolecular signals between H1-H2 and H4 are observed for α *galacto*/g*luco* pyranosidic residues while β *galacto/gluco* pyranosidic residues will show connectivities between H1-H3 and H5. This is also observed for β-D-Fuc*p*NAc4N and the glucoses of Fig. 19C. It allowed for the assignment of H5, which coupled to H6 in COSY and TOCSY (not shown). Hence, the whole spin system of β-D-Fuc*p*NAc4N could be assigned.

If the mixing time is too long, magnetization between vicinal protons will start to be transferred, hence giving rise to a phenomenon called spin diffusion (Fig. 19C, GlcI and GlcII).

2D hetero nuclear correlation spectroscopy as HMQC and HSQC are used to correlate the proton shifts to their respective carbons. These experiments are run with proton decoupling in order to reduce the number of signals for each carbon. For sugars as the inner core heptoses of NTHi, L-α-D-Hep, the anomeric configuration can be delineated by running a non decoupled experiment. The distance between the two signals is determined and measured in Hz, Table 1, as illustrated in Fig. 20 on the example of core OS from NTHi 1247. A hetero-coupling constant $^1J_{H1,C1}$ of ~170-175 Hz is characteristic for an α pyranosidic configuration as observed for the inner core heptoses L-α-D-Hep*p*I - III. A $^1J_{H1,C1}$ value of ~160-167 Hz is indicating a β pyranosidic configured sugar, as observed for e.g. β-D-Glc*p*.

Substitution sites are indicated by downfield shifted values for carbon by ~3-10 ppm while the neighboring carbons are upfield shifted by ~1-3 ppm.

Heteronuclear multiple bond correlation spectroscopy (HMBC) enables the observation of three bond connectivities between carbon and proton. The information is used for sequence assignment.
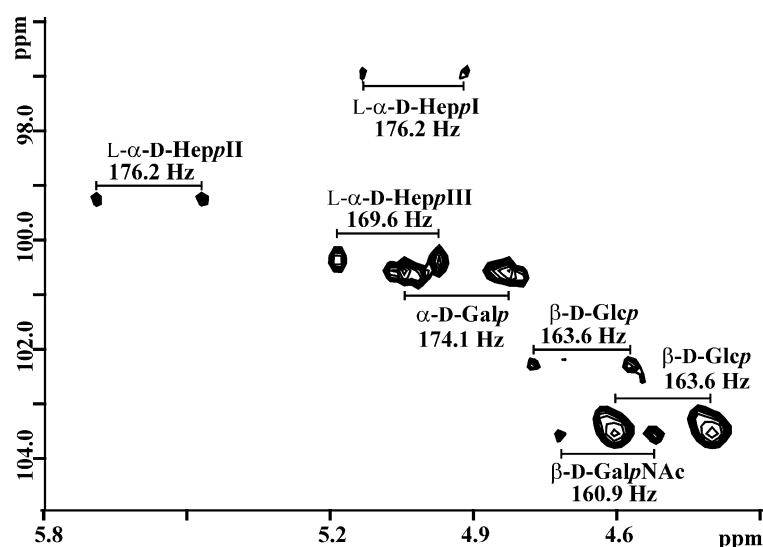
Figure 20: Selected region of the non-decoupled HSQC ^{1}H-^{13}C of core OS from NTHi strain 1247 run at 25°C and 800 MHz.

3.6   Anion Exchange Chromatography

High performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD) is a sensitive technique for the separation and characterization of carbohydrates.

The hydroxyl groups of carbohydrates are deprotonated, *i.e.* in anionic form, under strongly alkaline conditions and can be separated by anion exchange chromatography. For HPAEC-PAD the basic gradient solvent is 0.1 M sodium hydroxide solution. The column is a strong anion exchange stationary phase. The sample is driven out by a rising concentration of sodium acetate which competes with the sample in interaction to the column. The PAD detects carbohydrates by oxidation on a gold electrode surface.

HPAEC-PAD was used for the direct detection of trace amounts of sialic acids (down to pictogram/mg) [79]. The identity of Neu5Ac was confirmed by comparison with an authentic standard which also confirms the retention time of the Neu5Ac.

The sample needs no advanced, time consuming purification steps. However, the retention times can vary considerably. Therefore, a periodical standard should be run in between the sample sequence for confirmation of the retention time.

Using HPAEC-PAD, sialic acid was determined to be present in almost all NTHi strains from a collection of clinical isolates [79]. In the present thesis this technique was used to determine Neu5Ac in the LPS-OH from *H. parainfluenzae* strain T3T1 and 22 (**Paper III**).

An alternative application for HPAEC-PAD is the preparative separation and purifications of oligosaccharide mixtures as accomplished by Schweda *et.al.* [106]. However, the exposure of oligosaccharide to the very basic environment can initiate degradations, as complete loss of *O*-acetyl groups but this can be reduced by instant neutralization of the fraction prior to desalting by GPC.

4 Results

4.1 The gene function of *lic2B* in *H. influenzae* (Paper I)

Background:

Lic2B is encoded by a phase variable gene *lic2B* starting with multiple tandem repeats of 5´-CAAT-3´ (see page 7). This gene has been the focus for several investigations. Pettigrew *et. al.* screened a total of 210 type b and non-typeable *H. influenzae* isolates and found *lic2B* to be present in 52% of the middle ear isolates [127].

In Paper I we investigated the function of the gene *lic2B* by comparing the LPS of a mutant strain with that of wildtype Eagan [61]. Previous studies of *lic2B* were hampered due to a polar effect when mutating *lic2B*. As a consequence of mutation in *lic2B*, *lic2C* was also defective. This circumvented all hexose expression from HepII in the mutant strains. The bacteria in Paper I were restored with a functioning reintroduced *lic2C* copy as a plasmid resulting in the mutant *H. influenzae* strain Eagan*lic2B*lic2C+.

Results:

LPS-OH and core OS from *H. influenzae* strain Eagan*lic2B*lic2C+ were investigated by sugar and methylation analysis, MS and NMR experiments.

The compositional sugar analysis of OS material showed Glc, Gal and L,D-Hep. Sugar analysis on LPS-OH additionally showed GlcN that was confirmed to originate from the lipid A. Methylation analysis showed the substitution for the heptoses to be as expected for the branched triheptosyl moiety. Furthermore, it resulted into terminal Glc and terminal Gal.

The ESI-MS of LPS-OH material indicated the major glycoforms to be $Hex_3 \cdot Hep_3 \cdot Kdo \cdot P \cdot PEtn_{1-2} \cdot LipidA-OH$. The core oligosaccharide resulted in major $Ac_{0-1} \cdot Hex_3 \cdot Hep_3 \cdot PEtn \cdot AnKdo-ol$.
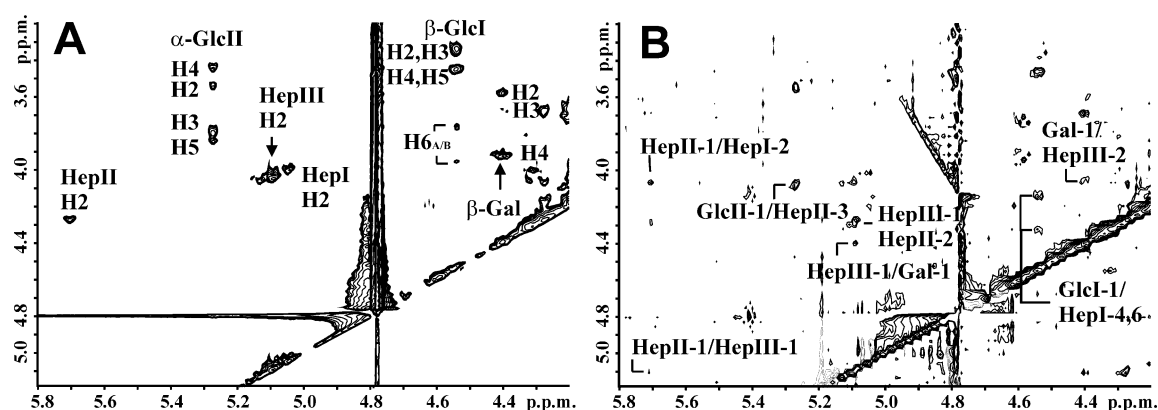
Figure 21: Selected regions of the 500 MHz 2D NMR experiments of *H. influenzae* strain Eagan*lic2B*lic2C+. Shown is A) the TOCSY, 180 ms and B) the NOESY, 200 ms.

The core OS was further investigated to determine its sequence and branching pattern. Details of the various glycoforms were established after dephosphorylation and by LC-ESI-MS[n]. In accordance with the suggested mutant it was shown that the three outer core hexoses were terminal. No further elongation of the middle hexose could be observed.

NMR experiments confirmed glucose substitution on Hep I and HepII in β and α configurations, respectively. HepIII was found to be elongated by β-galactopyranose (Fig. 21). Hence, comparison of the investigated mutant LPS with the wildtype LPS from strain Eagan established that Lic2B is responsible for the linkage of β-D-Glc$p$-(1→4)-α-D-Glc$p$-(1→ elongating from the middle heptose.

4.2   The LPS of NTHi strains 1247 and 1008 (Paper II)
Background:

Two distinct NTHi strains, namely 1247 and 1008, were isolated from patients with otitis media. The two strains are part of the selection of *H. influenzae* strains that represent the genetic diversity of this species. In Paper II we investigated their detailed LPS structure. A *lpsA* mutant of NTHi strain 1247 lacking all elongation from the distal heptose was constructed in order to facilitate structural analysis from HepI.

Results:

The LPS structure of NTHi 1247 was determined to consist of the conserved inner core structure L-α-D-Hep$p$-(1→2)-[PEtn→6]-L-α-D-Hep$p$-(1→3)-[β-D-Glc$p$-(1→4)]-L-α-D-Hep$p$ linked to lipid A via a single Kdo-4-phosphate; as it is common to all investigated NTHi. Sugar analysis of core OS resulted in Glc, Gal, L,D-Hep, GalN and in the LPS-OH additionally in GlcN. Methylation analysis and analysis of permethylated OS-P using LC-ESI-MS^n (Fig. 22A) confirmed the structure of the triheptosyl inner core. As expected HepI was substituted by GlcI. HexNAc-Hex-Hex-Hex was indicated to elongate from GlcI and HepIII. Truncated glycoforms were observed, which rationalized the terminal Glc and terminal Gal of the methylation analysis. Interestingly, LC-ESI-MS^n indicated HexNAc-Hex-Hex-Hex elongation from HepIII only in the HexNAcHex4 glycoform. The larger HexNAc containing glycoforms only showed substitution from GlcI and no HexNAc2 glycoforms were detected. Hence, the highest glycoform expressed only one globotetraose epitope from GlcI with elongation from HepIII consisting of a digalactoside epitope (see Figure 23).
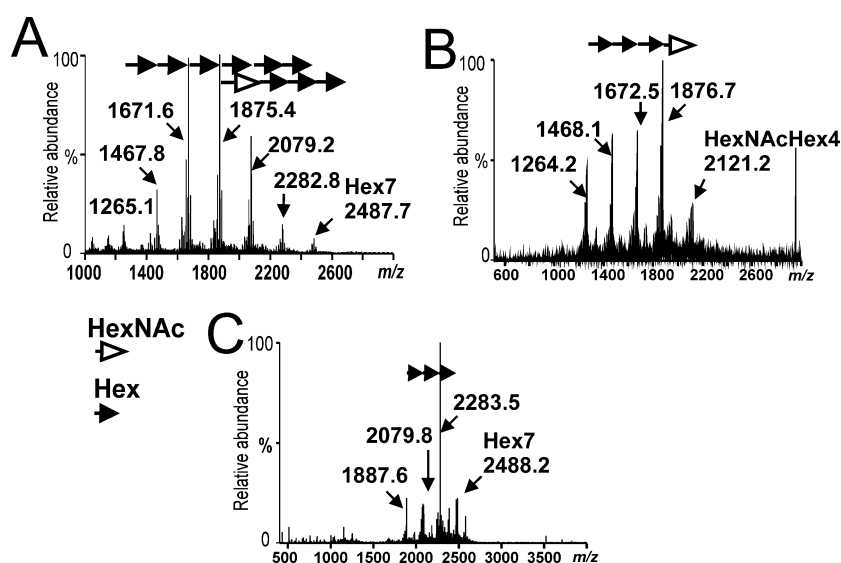
Figure 22: The mass spectra of dephosphorylated and permethylated core OS of A) NTHi 1247, B) NTHi 1247*lpsA* and C) NTHi 1008. Indicated are the sodiated adduct ions and the molecular composition all including Hep3*An*Kdo-ol.

31

2D NMR experiments demonstrated globotetraose to elongate from the phosphocholine bearing GlcI, [β-D-GalpNAc-(1→3)-α-D-Galp-(1→4)-β-D-Galp-(1→4)-β-D-Glcp-(1→4)-[PCho→6]-β-D-GlcIp-(1→]. The established LPS structure for NTHi strain 1247 is shown in Figure 23. Truncated glycoforms were indicated in the NMR spectra resulting in sub-glycoforms with globotriose and lactose chain extensions.

The highest glycoform detected in NTHi 1247*lpsA* was HexNAcHex4 in the dephosphorylated and permethylated core OS, Fig. 22B. NTHi strain 1247*lpsA* was found to express globotetraose chain extensions from GlcI corresponding to the Hep3 glycoform observed in NTHi strain 981 [54]. Methylation analysis showed increased abundance of 3-substitued Gal for LPS-OH when compared to the OS. This was consistent with the presence of glycoforms where GlcI is extended by the sialyllactose epitope, [α-Neu5Ac-(2→3)-β-D-Galp-(1→4)-β-D-Glcp-(1→] as shown by CE-ESI-MS/MS. Furthermore, disialyllactose, α-Neu5Ac-(2→8)-α-Neu5Ac-(2→3)-β-D-Galp-(1→4)-β-D-Glcp-(1→, was also found to substitute GlcI, which is a novel molecular environment for this structure, Fig. 23. The addition of β-D-Glcp rather than β-D-Galp to O-4 of GlcI is accounted for by polymorphism of Lex2 [47, 48].

Further heterogeneity was observed due to substitution with non-carbohydrate substituents with up to two acetates and/ or glycine. Analysis by LC-ESI-MS/MS showed the acetylation sites to be GlcI and HepIII (see Fig. 12B).
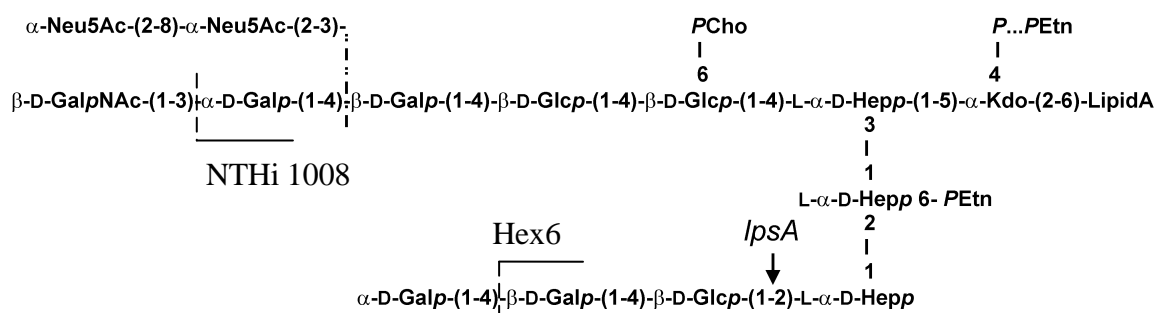
Figure 23: The structure of LPS expressed by NTHi 1247. In which the disialyllactose epitope determined in NTHi 1247*lpsA* is indicated. The LPS structure expressed by NTHi 1008 is also indicated, for the predominant Hex6 glycoform.

NTHi strain 1008 showed two major glycoforms with six or seven hexoses in the outer core, see Fig. 22C. Globotriose was found to extend from GlcI and HepIII. The structures of LPS from NTHi 1247 and 1008 were found to be essentially the same with the only difference being the lack of β-D-GalpNAc in the LPS of NTHi 1008, Fig. 23. The reason for this is unknown since *lgtD* is present. The reason for the gene´s malfunctioning was not further investigated.

CE-ESI-MS indicated sialylated glycoforms, presumably sialyllactose α-Neu5Ac-(2→3)-β-D-Galp-(1→4)-β-D-Glcp-(1→, by precursor ion scanning at *m/z* 290 (sialic acid).

4.3 *H. parainfluenzae* genome strain T3T1 and strain 22 (Paper III)

Background:

In this study the LPS of *H. parainfluenzae* genome sequence strain T3T1 and strain 22 were investigated. *H. parainfluenzae* strains T3T1 and 22 were collected from the throat of healthy individuals in the Gambia and the United Kingdom, respectively. The genome of strain T3T1 was completely sequenced at the Sanger Centre, UK, and genetic comparison to the closely related *H. influenzae* has been accomplished. These studies suggested that the genetic basis for both the lipid A and the inner core are very similar to NTHi [DW Hood, unpublished results]. Yet, the genes responsible for the outer core differed considerably, suggesting an O-repeating unit-like assembly. In order to investigate this further analysis of the LPS expressed by both strains was accomplished.

Results:
The LPS structure of *H. parainfluenzae* strain T3T1

The elucidation of the LPS from *H. parainfluenzae* strain T3T1 was accomplished on LPS, LPS-OH, OS and lipid A.

Fatty acid analysis of lipid A indicated 3-hydroxy-tetradecanoic and tetradecanoic acid. ESI-MS^n on lipid A in the negative ion mode established a diphosphorylated hexosamine backbone with phosphate groups at C-1 and C-4` and acylation by 3-hydroxy-tetradecanoic acids [14:0 (3-OH)] on C-2, C-3, C-2` and C-3`. The fatty acid chains elongating from C-2`and C-3` were further substituted by tetradecanoic acids (14:0).

NMR on LPS-OH indicated a double doublet in the α configurated and a β coupled anomeric signal at δ 5.38 and δ 4.68, respectively. These signals were assigned to originate from GlcNI and GlcNII of the glycosidically linked phosphorylated diglucosamine backbone β-D-Glc*p*NII-(1→6)-α-D-Glc*p*NI. The data confirmed this to be the same structure as it has been reported before for NTHi [26, 27].
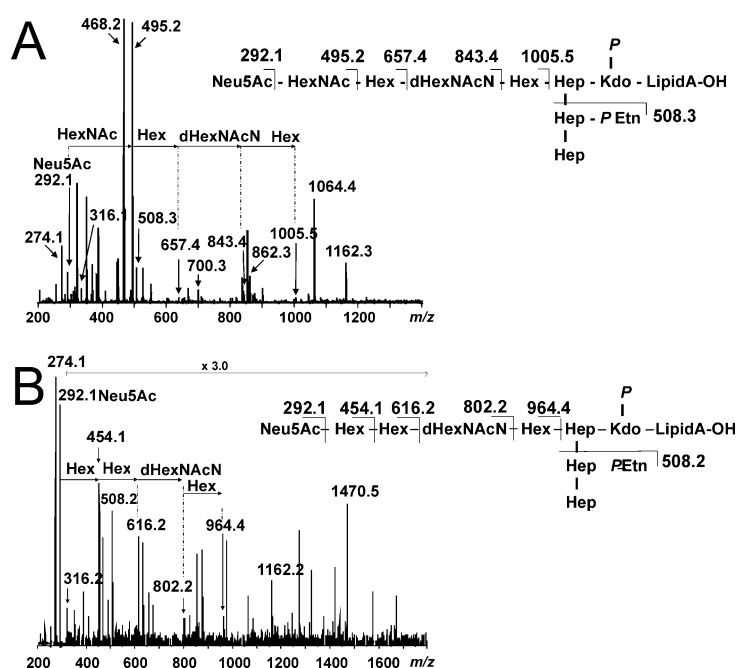
Figure 24: CE-ESI-MS/MS analyses of LPS-OH from A) *H. parainfluenzae* stain T3T1 and B) *H. parainfluenzae* stain 22. The proposed fragmentation is depicted in the respective inset.

The inner core oligosaccharide was found to consist of the triheptosyl moiety with glucose substitution on the proximal heptose, HepI, L-α-D-Hep$p$-(1→2)-[PEtn→6]-L-α-D-Hep$p$-(1→3)-[β-D-Glc$p$-(1→4)]-L-α-D-Hep$p$-(1→. It is attached to the lipid A via one phosphorylated Kdo as observed by inter alia CE-ESI-MS/MS on LPS-OH and 2D NMR. This is the same structure as reported for NTHi [57].

The outer core was found to contain the sugar 2-acetamido-4-amino-2,4,6-trideoxy-D-galactopyranose (Fuc$p$NAc4N) as determined by 2D NMR. Since this sugar is acid labile all chemical analyses failed. Never the less, CE-ESI-MS and CE-ESI-MS/MS showed the occurrence of the Fuc$p$NAc4N sugar, Fig. 24A. CE-ESI-MS indicated glycoform compositions of Hex$_1$·Hep$_3$·P·PEtn$_{1-2}$·Kdo·LipidA-OH and minor abundant glycoforms of the proposed compositions Neu5Ac·HexNAc·dHexNAcN·Hex$_2$·Hep$_3$·P·PEtn$_{1-2}$·Kdo·LipidA-OH. The latter were also detected when applying the precursor ion scanning for sialic acid, m/z 290. Since no ions for truncated glycoforms were detectable, this data suggested further that the outer core was attached *en bloc*.

Interestingly, CE-ESI-MS/MS on LPS in precursor ion mode on the ion m/z 332 gave evidence for O-acetylation on the sialic acid, see Fig. 15. The site of attachment of Ac was determined on mildly hydrolysed LPS after pertrimethylsilylation by GC-MS, and gave evidence for acetylation at O-9 of Neu5Ac, see Fig. 13B. 2D NMR on OS and LPS-OH material gave evidence for the outer core structure of [α–Neu5,9Ac$_2$-(2→6)-β-D-Gal$p$NAc-(1→4)-β-D-Gal$p$-(1→3)-β-D-Fuc$p$NAc4N-(1→4)-β-D-Glc$p$-(1→] elongating from HepI in accordance with methylation analysis (see Figure 25A).

A *lgtF* mutant of strain T3T1 was found to express no elongation from the triheptosyl inner core, demonstrating the *lgtF* gene product to encode for β-D-Glc$p$-(1→4 glucosyltransferase.
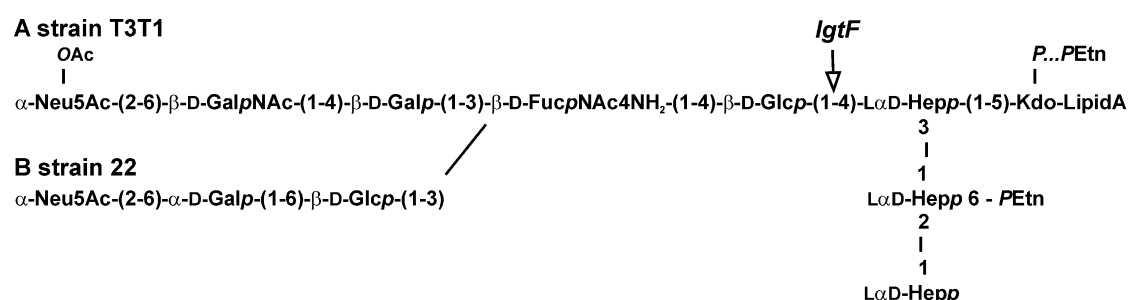
Figure 25: The proposed LPS structure of (A) *H. parainfluenzae* strain T3T1 and (B) *H. parainfluenzae* strain 22. The function in *H. parainfluenzae* strain T3T1 of *lgtF* is indicated.

The LPS structure of *H. parainfluenzae* strain 22

The LPS from *H. parainfluenzae* 22 was analyzed by mass spectrometry and NMR on lipid A, LPS-OH and OS material. The lipid A of strain 22 was investigated by fatty acid analysis and ESI-MS[n] and resulted in virtually the same structure as found in *H. parainfluenzae* strain T3T1 and NTHi.

CE-ESI-MS of the LPS-OH resulted in ions indicating glycoform compositions of Neu5Ac·dHexNAcN·Hex$_3$·Hep$_3$·Kdo·P·PEtn$_{1-2}$·LipidA-OH (major) and Hex$_1$·Hep$_3$·P·PEtn$_{1-2}$·Kdo·LipidA-OH (minor). Tandem MS on these glycoforms indicated the triheptosyl moiety inner core with hexose and PEtn substitution, respectively, on HepI and HepII as observed for strain T3T1. The outer core was shown to be built up by Neu5Ac·Hex·Hex·FucNAc4N elongating from the hexose of HepI, Fig. 24B. 2D NMR experiments on OS allowed the assignment of α-D-Gal$p$-(1→6)-β-D-Glc$p$-(1→3)-β-D-Fuc$p$NAc4N-(1→ attached to O-4 of GlcI (see Figure 25B). Methylation analysis of LPS-OH gave evidence for 6-substituted Gal. In contrast, methylation analysis of OS resulted only in terminal Gal. Hence, Gal was substituted by sialic acid which was cleaved during mild acid hydrolysis.

4.4 *H. parainfluenzae* strain 13 (Paper IV)

Background:

The commensal *H. parainfluenzae* strain 13 was isolated from a healthy individual and chosen for LPS investigation due to its genetic background that indicated an ABC transport mechanism for the elaboration of the O-repeating unit.

Results:

The LPS was extracted using the hot phenol-water method. Both phases contained LPS, with the rough LPS in the water phase and O-chain containing LPS in the phenol phase. Investigation of the rough type LPS by sugar analysis indicated Glc, Gal and L,D-Hep. Methylation analysis suggested a doubly branched core structure and terminal hexoses. LC-ESI-MS^n on permethylated OS-P showed two glycoforms with compositions of Hex1Hep3Kdo and Hex2Hep3Kdo, respectively. MS$^{2-4}$ showed hexose substitution in both glycoforms on HepI. In the Hex2 glycoform HepII was additionally substituted by one hexose residue.

NMR on core OS gave evidence for the triheptosyl moiety, [L-α-D-Hep$p$-(1→2)-[PEtn→6]-L-α-D-Hep$p$-(1→3)-L-α-D-Hep$p$-(1→]. The first heptose was found to be substituted by β-Glc$p$ at O-4. HepII was substituted at O-3 with α-configurated glucopyranose. The core oligosaccharide was connected to the lipid A via Kdo.

The lipid A was investigated by fatty acid analysis and ESI-MS^n were it was found to have a diphosphorylated dihexoseamine backbone with acylation by 3-hydroxy-tetradecanoic acids [14:0 (3-OH)] on C-2, C-3, C-2`and C-3´. The fatty acid chains elongating from C-2`and C-3` were further substituted by tetradecanoic acids (14:0). This is the same structure as reported before for *H. parainfluenzae* strain 20, T3T1 and 22 (**Paper III**) and NTHi [27].

O-antigen containing LPS was recovered from the phenol phase and showed Gal and GlcN in the sugar analysis. Methylation analysis indicated 6-substituted galactofuranose and 3-substituted GlcNAc. The O-antigen structure was investigated by NMR on material gained after mild acid hydrolysis. The 1D proton NMR spectrum is shown in Figure 26 and revealed heterogeneity in the structure that could be accounted for by the non-stochiometric presence of non-carbohydrate substituents. 2D NMR enabled us to propose the structure of the O-repeating unit as shown in Fig 26C.

Investigation of the point of attachment of the O-antigen to the core was hampered, due to low abundance and overlap in the NMR spectra. This is still under investigation. In comparison with the investigated *H. parainfluenzae* strains 20, T3T1 and 22 it is reasonable to assume that the O-antigen extends from GlcI.
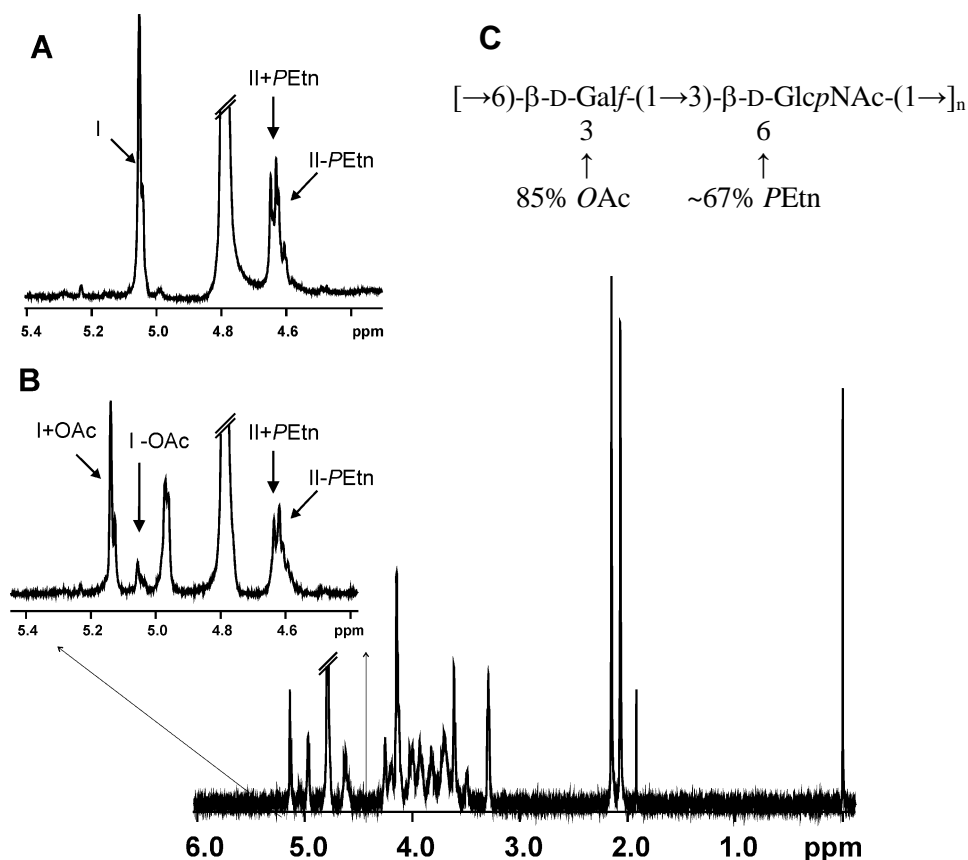
Figure 26: One dimensional ^{1}H spectrum of OSpp from *H. parainfluenzae* strain 13. A) the anomeric region of OSpp after O-deacetylation, B) the anomeric region of OSpp and C) the proposed structure of the O-antigen.

4.5   Additional structural studies

4.5.1   *H. parainfluenzae* strain 20*lic1+*

H. influenzae are frequently decorated with *P*Cho. *P*Cho plays an important role in the course of colonization of NTHi and other bacteria of the nasopharynx such as *S. pneumoniae* [76, 77] and *N. meningitidis* [78]. However, the LPS of all *H. parainfluenzae* strains investigated to date do not contain *P*Cho. In order to investigate the influence of *P*Cho incorporation the genome of *H. parainfluenzae* strain 20 [92] was extended with the *lic1* locus from *H. influenzae* Rd (RM 118). Strain Rd is a type d strain that is known to express *P*Cho at O-6 of GlcI [45].

The LPS from the mutant strain was extracted by the hot phenol-water method [101] with subsequent enzymatic purification and dialysis. Both phases contained LPS. The water phase contained the O-antigen containing LPS, while the phenol phase was shown to containe LPS with *P*Cho incorporation.

Structural information was obtained on core OS by LC-ESI-MS and MS/MS experiments. The mass spectrum indicated glycoforms with the compositions

PCho•Hex$_1$•Hep$_3$•PEtn•AnKdo-ol and PCho•Hex$_2$•Hep$_3$•PEtn•AnKdo-ol. The MS^2 on the doubly protonated ion m/z 625.2, $i.e.$ PCho•Hex$_1$•Hep$_3$•PEtn•AnKdo-ol, is shown in Figure 27. Fragment ions were observed in singly and doubly charged state. The marker ion m/z 327.9 and 742.1 corresponds to PCho-hexose (PCho-Hex) and PCho-Hex-Hep-Kdo. MS^2 on the singly protonated ion m/z 1411.2, $i.e.$ PCho•Hex$_2$•Hep$_3$•PEtn•AnKdo-ol, resulted in ions confirming phosphocholine on the hexose elongation from the first heptose. PCho substitution of the hexose on the third heptose could not be observed.
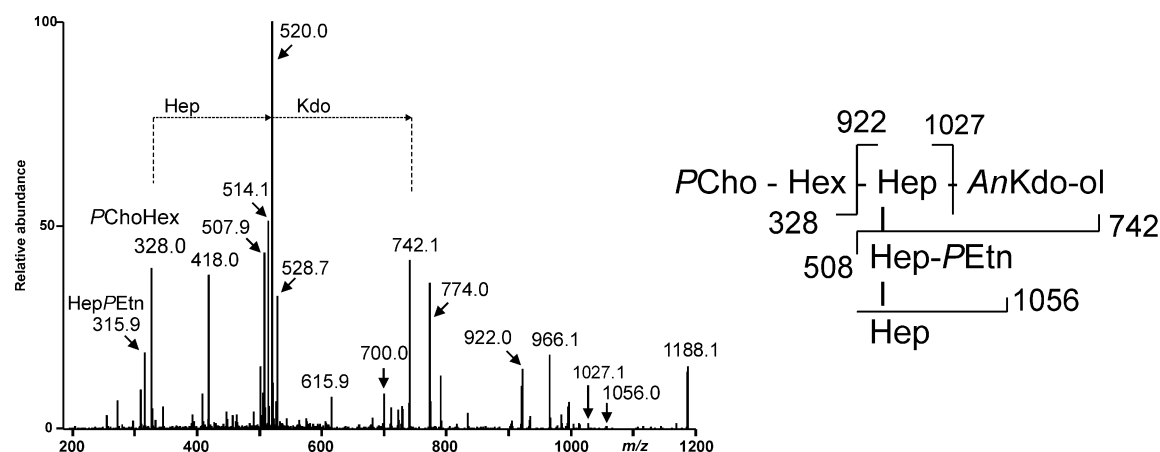
Figure 27: LC-ESI-MS^2 on the reduced core oligosaccharide of *H. parainfluenzae* strain 20*lic1+*.

NMR experiments (1D ^{1}H, COSY, ^{1}H-^{31}P HMQC) showed PCho specific resonances from the methyl groups at 3.24 ppm correlating with a carbon signal at 54.3 ppm. PCho substitution on GlcI was supported by the chemical shifts of GlcI, showing signals inter alia H-6 4.23 ppm and a high field shifted C-6 at 64.6 ppm indicating a substitution site.

Hence, PCho substitutes GlcI at C-6 as observed in the *H. influenzae* wildtype strain Rd. The impact of phosphocholine substitution on colonization and survival inside the host is a matter of ongoing investigations.

4.5.2 LC-ESI-MS method development for LPS-OH

A method for the structural analysis of LPS-OH has been developed. Commonly, CE-ESI-MS was employed to investigate LPS-OH from NTHi. However, CE-ESI-MS can be rather delicate to handle and difficult to automate. In order to overcome these drawbacks a LC-MS method was developed using a Phenomenex Synergi Polar PR column that was combined with ESI-triple quadrupole MS.

Results

The solvent acetonitrile and water with 10mM NH$_4$OAc had a flow rate of 0.05 ml·min^{-1}. The gradient conditions were the following: 10% acetonitrile at 0 min, rising to 40 % acetonitrile from 3-5min. The column was reconditioned between minute 6 and 7 at 10 % acetonitrile.

Analyses were performed on LPS-OH samples from NTHi strain 1003 for which LPS structures had been established [128]. The sample eluted as a sharp peak typically with a retention time of 4.8 min, Fig. 28A. The mass spectrum of LPS-OH from NTHi

1003 is shown in Figure 28B and resulted in triply and doubly deprotonated ions. The corresponding glycoforms are in accordance with the established LPS structure for this strain [128].

The method was applied in both the positive and negative ion mode and with advanced tandem MS methods including precursor ion scanning. All experiments resulted in satisfying mass spectra.

The linearity of the detector response was determined in multiple reaction monitoring (MRM) scan mode with triple replicates. The standard curve (18-40 ng·µl^{-1}, triple replicates) for the highest abundant ion, m/z 1281.3, resulted in a correlation coefficient of 0.9814.

In order to validate the approach of LPS-OH analysis by LC-MS, the applicability of the method was demonstrated on material from two further NTHi strains, 1200 and 375*lic1* [60, 121], and from three other bacteria, *H. parainfluenzae* strain 22 (**Paper III**), *Pasteurella multocida* strain AL 85 and *N. meningitidis* L4*galE*.

LC-MS was readily applicable on LPS-OH of NTHi strains 1200 and 375*lic1A*. The mass spectra resulted in triply and doubly deprotonated ions. For both strains, observed ion ratios and the corresponding glycoform compositions were comparable to those described before [60, 121]. The LPS-OH material from the three other bacteria resulted also in triply and doubly deprotonated ions with comparable ion ratios as observd before (**Paper III**, AD Cox, personal communication). All LPS-OH samples resulted in retention time of approx. 4.8 min.
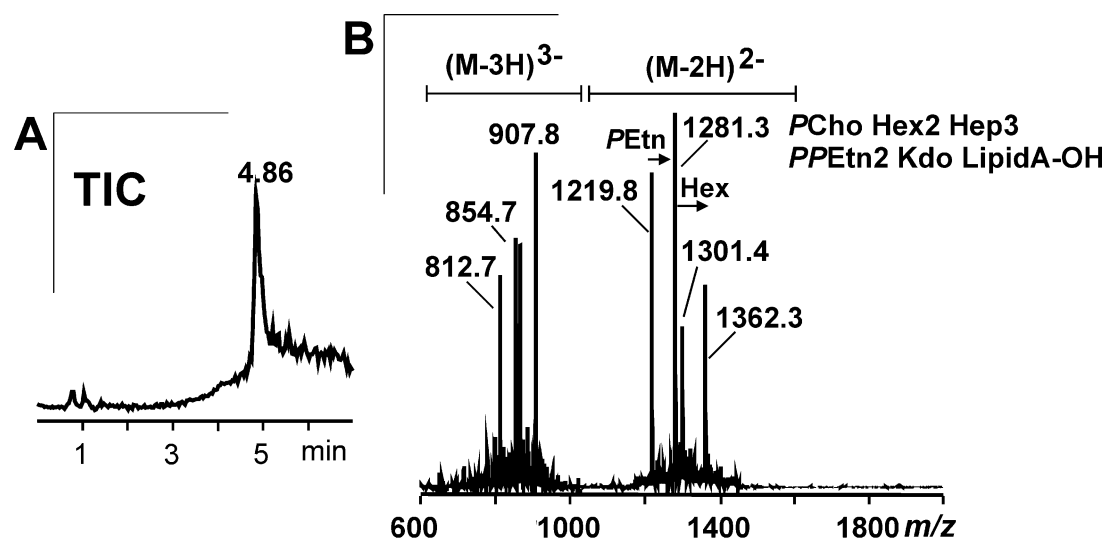
Figure 28: LC-ESI-MS analysis of LPS-OH derived from NTHi strain 1003. A) The total ion chromatogram, B) the extracted mass spectrum of 4.5-5 min.

LC-MS for the analysis of LPS-OH was found to be a fast and sensitive tool for structural analysis. It was found to be suitable for all advanced MS experiments and quantification. Furthermore, this method has proven to be robust and can easily be applied in many laboratories, since LC-MS is a common and widespread instrument combination.

5 Discussion

The aim of this thesis was focused on the determination of LPS structures expressed by the pathogen *H. influenzae* and the commensal *H. parainfluenzae*, in order to further understand the role of this surfaced exposed molecule in host interactions.

The *H. influenzae* strain investigated in Paper I was a mutant from the well established type b strain Eagan. The investigated NTHi strains were collected from the middle ear of patients with otitis media while the *H. parainfluenzae* strains were obtained from the throats of healthy individuals.

In Paper I it was established for the first time that *lic2B* encodes a glucosyltransferase. This was achieved by comparison of LPS from *H. influenzae* strain Eagan *lic2B*lic2C+ to that of the wildtype. While all other LPS coding genes in *H. influenzae* strain Eagan had been established, the function of *lic2B* was unknown. It is noteworthy that a recent study on NTHi strain NT172 conducted structural investigation on strain NT172 with *lic2B* deletion. These authors found *lic2B* to encode a galactosyltransferase [129]. Hence, the gene *lic2B* appears to express a similar promiscuity regarding the addition of β-D-Glc*p* or β-D-Gal*p* to O-3 of HepII as observed before for *lpsA* and *lex2* [47, 48, 64].

Structural studies of NTHi strains 1247 and 1008 revealed the presence of epitopes that are important in the pathogenic life of the bacteria. Globobiose was found to substitute GlcI and HepIII in strain 1008. Globotetraose was found to extend from GlcI and HepIII in strain 1247. Additionally, disialyllactose was unambiguously determined elongating from GlcI in strain 1247*lpsA*. Disialyllactose was also observed as an extension from HepI in NTHi strain 1124 [49], but in this strain β-D-Gal*p* substituted GlcI at O-4, while in strain 1247 and 1008 GlcI were substituted by β-D-Glc*p*. Disialylated glycoforms have been observed extending from HepIII in NTHi strain 162 (i.e. 1292) [66] and 375 [82]. The importance of sialic acid during the course of colonization has been established before in the chinchilla animal model of otitis media [84, 85]. The digalactoside structure was determined to increase the virulence of the bacteria in infant rats [130].

H. parainfluenzae strain T3T1 and 22 were shown to express rough type LPS with a single outer core oligosaccharide extending GlcI. In accordance with genetic analysis the lipid A and inner core of these *H. parainfluenzae* strains were found to be identical to NTHi. The complete genome sequence of strain T3T1 has been established. Genetic analysis suggested a Wzy-like elongating mechanism for strain T3T1 indicating the outer core to be attached *en bloc*. NTHi shows similar outer core extensions in the sialyllacto-*N*-neotetraose and its related *P*Etn-GalNAc bearing structure that are both encoded by the *hmg* locus. Otherwise, stepwise addition of the outer core substituents is commonly observed for NTHi and has been reported for two clinical isolates of *H. parainfluenzae* [57].

The outer core of *H. parainfluenzae* strain T3T1 was established to be the novel structure α-Neu5,9Ac$_2$-(2→6)-β-Gal*p*NAc-(1→4)-β-Gal*p*-(1→3)-β-Fuc*p*NAc4N-(1→ elongating from GlcI. The terminal epitope α-Neu5Ac-(2→6)-α-D-Gal*p*NAc-(1→ was

described for the first time in human colon mucin, and is known as the sialyl-Tn-epitope [131, 132]. *O*-acetylated sialyl-Tn was reported in ovarian cancer where the acetyl group had a critical function in recognition by antibodies [133]. For *H. parainfluenzae* strain T3T1 it is tempting to speculate that the oligosaccharide, through molecular mimicry, contributes to evading the human immune response.

Interestingly, the outer core of strain 22 is also initiated by the sugar Fuc*p*NAc4N. To our knowledge the outer core structure, i.e. α–Neu5Ac-(2→6)-α-D-Gal*p*-(1→6)-β-D-Glc*p*-(1→3)-β-D-Fuc*p*NAc4N-(1→, has not been reported in LPS before. The terminal structure α–Neu5Ac-(2→6)-α-D-Gal*p*-(1→ has been previously been reported for *N. meningitidis* immunotype L1 where it was part of the sialylated P^K epitope [134].

S. pneumoniae has been shown to excrete a neuraminidase into its environment that cleaved terminal sialic acid bound to galactose and *N*-acetyl-galactoseamine on LPS from *H. influenzae* and *N. meningitidis* [135]. In our studies, we treated native LPS of *H. parainfluenzae* strain T3T1 with neuraminidase from *Clostridium perfringens*. This neuraminidase cleaves terminal sialic acid as it is found in NTHi and also *H. parainfluenzae* strain 22. *C. perfringens* neuraminidase failed to cleave the O-acetylated terminal sialic acid of strain T3T1. Therefore, LPS from strain T3T1 might also be protected from desialylation by the neuraminidase from *S. pneumoniae*. In this way *H. parainfluenzae* strain T3T1 might have an advantage in competing for the same ecological niche.

H. parainfluenzae strain 13 was established to express virtually the same lipid A and inner core as all other *H. parainfluenzae* strains investigated so far. One important difference is the occurrence of α-Glc*p* substitution at O-3 of HepII. The O-antigen was established to consist of the novel structure, [→6)-[*O*Ac→3]-β-D-Gal*f*-(1→3)-[*P*Etn→6]-β-D-Glc*p*NAc-(1→]. Genetic analysis suggested an ABC dependent transport mechanism for the O-repeating unit. The non-stoichiometric substitution of the O-antigen with *O*-acetyl groups and *P*Etn modulates the hydrophilicity of the molecule and could play a role in colonization. This could be the basis of further investigations.

Substitution of HepII with α-Glc*p* is also observed in NTHi LPS. In NTHi the phase variable gene *lic2C* encodes for the glucosyltransferase responsible for this addition. Though, *H. parainfluenzae* strain 13 expresses the same structure, *lic2C* has not been identified in its genome. *H. parainfluenzae* strain 20 showed also further variation in the core oligosaccharide. The genetic basis of core variation has to be established in future studies.

H. parainfluenzae can synthesize smooth and rough type LPS. As in the case of untested bacteria and/or mutants (e.g. *H. parainfluenzae* strain 13 and 20*lic1+*) a more efficient LPS extraction could be obtained by combining the PCP and hot phenol-water method. Extraction with PCP will result in the rough type LPS while extraction with hot phenol-water will make the O-antigen containing LPS available [136]. Future studies should take this into account.

While bacterial infections remain one of the main causes for global morbidity and mortality, bacterial resistance towards antibiotics is still increasing. In this thesis we aimed to understand host-pathogen interactions. In our approach we elucidated the LPS

structures and the genes involved in the LPS biosynthetic pathway of two closely related bacteria that colonize the human nasopharynx, namely *H. parainfluenzae* and *H. influenzae.*

The outer membrane molecule LPS exposes the carbohydrate part into the environment. Hence, LPS plays a major role in determining the virulence and commensal behavior of the bacteria. *H. influenzae* lacks the O-antigen chain, something it has in common with other bacteria living in the upper respiratory tract. *H. parainfluenzae* is closely related to *H. influenzae*. This relationship is represented in the structures of the LPS inner core and lipid A that are virtually identical. The commensal strains of *H. parainfluenzae* investigated by us did not display any of the virulence factors observed in *H. influenzae*. Prominent features of NTHi LPS are truncated versions of the highest glycoforms due to phase variation, molecular mimicry and non-carbohydrate substituents such as *P*Cho. The lack of *P*Cho in the LPS of *H. parainfluenzae* could indicate a different colonization mechanism and is currently under investigation .

6 Acknowledgements

7 References

1. Hari G. Garg, MKC, Charles A. Hales; *Carbohydrate Chemistry, Biology and Medical Applications*, ed. Hari G. Garg, MKC, Charles A. Hales. **2008**, Oxford, United Kingdom: ELSEVIER. 432.
2. Moyes, RB; Reynolds, J; Breakwell, DP; *Differential staining of bacteria: gram stain.* Curr Protoc Microbiol, **2009**. Appendix 3: p. Appendix 3C.
3. Bartholomew, JW Finkelstein, H; *Relationship of cell wall staining to gram differentiation.* J Bacteriol, **1958**. 75(1): p. 77-84.
4. Steeghs, L; den Hartog, R; den Boer, A; Zomer, B; Roholl, P; van der Ley, P; *Meningitis bacterium is viable without endotoxin.* Nature, **1998**. 392(6675): p. 449-50.
5. Frirdich, E Whitfield, C; *Lipopolysaccharide inner core oligosaccharide structure and outer membrane stability in human pathogens belonging to the Enterobacteriaceae.* J Endotoxin Res, **2005**. 11(3): p. 133-44.
6. Harvey, HA; Swords, WE; Apicella, MA; *The mimicry of human glycolipids and glycosphingolipids by the lipooligosaccharides of pathogenic neisseria and haemophilus.* J Autoimmun, **2001**. 16(3): p. 257-62.
7. Houliston, RS; Koga, M; Li, J; Jarrell, HC; Richards, JC; Vitiazeva, V; Schweda, EK; Yuki, N; Gilbert, M; *A Haemophilus influenzae strain associated with Fisher syndrome expresses a novel disialylated ganglioside mimic.* Biochemistry, **2007**. 46(27): p. 8164-71.
8. Pittman, M; *Variation and Type Specificity in the Bacterial Species Hemophilus Influenzae.* J Exp Med, **1931**. 53(4): p. 471-92.
9. Koga, M; Gilbert, M; Li, J; Koike, S; Takahashi, M; Furukawa, K; Hirata, K; Yuki, N; *Antecedent infections in Fisher syndrome: a common pathogenesis of molecular mimicry.* Neurology, **2005**. 64(9): p. 1605-11.
10. Campagnari, AA; Gupta, MR; Dudas, KC; Murphy, TF; Apicella, MA; *Antigenic diversity of lipooligosaccharides of nontypable Haemophilus influenzae.* Infect Immun, **1987**. 55(4): p. 882-7.
11. Zwahlen, A; Rubin, LG; Moxon, ER; *Contribution of lipopolysaccharide to pathogenicity of Haemophilus influenzae: comparative virulence of genetically-related strains in rats.* Microb Pathog, **1986**. 1(5): p. 465-73.
12. Schweda, EK; Richards, JC; Hood, DW; Moxon, ER; *Expression and structural diversity of the lipopolysaccharide of Haemophilus influenzae: implication in virulence.* Int J Med Microbiol, **2007**. 297(5): p. 297-306.
13. Wenger, JD; *Epidemiology of Haemophilus influenzae type b disease and impact of Haemophilus influenzae type b conjugate vaccines in the United States and Canada.* Pediatr Infect Dis J, **1998**. 17(9 Suppl): p. S132-6.
14. Cowgill, KD; Ndiritu, M; Nyiro, J; Slack, MP; Chiphatsi, S; Ismail, A; Kamau, T; Mwangi, I; English, M; Newton, CR; Feikin, DR; Scott, JA; *Effectiveness of Haemophilus influenzae type b Conjugate vaccine introduction into routine childhood immunization in Kenya.* JAMA, **2006**. 296(6): p. 671-8.
15. Lewis, RF; Kisakye, A; Gessner, BD; Duku, C; Odipio, JB; Iriso, R; Nansera, D; Braka, F; Makumbi, I; Kekitiinwa, A; *Action for child survival: elimination of Haemophilus influenzae type b meningitis in Uganda.* Bull World Health Organ, **2008**. 86(4): p. 292-301.
16. Adegbola, RA; Secka, O; Lahai, G; Lloyd-Evans, N; Njie, A; Usen, S; Oluwalana, C; Obaro, S; Weber, M; Corrah, T; Mulholland, K; McAdam, K; Greenwood, B; Milligan, PJ; *Elimination of Haemophilus influenzae type b (Hib) disease from*

The Gambia after the introduction of routine immunisation with a Hib conjugate vaccine: a prospective study. Lancet, **2005**. 366(9480): p. 144-50.

17. Adegbola, RA; Usen, SO; Weber, M; Lloyd-Evans, N; Jobe, K; Mulholland, K; McAdam, KP; Greenwood, BM; Milligan, PJ; *Haemophilus influenzae type b meningitis in The Gambia after introduction of a conjugate vaccine.* Lancet, **1999**. 354(9184): p. 1091-2.

18. Darras-Joly, C; Lortholary, O; Mainardi, JL; Etienne, J; Guillevin, L; Acar, J; *Haemophilus endocarditis: report of 42 cases in adults and review. Haemophilus Endocarditis Study Group.* Clin Infect Dis, **1997**. 24(6): p. 1087-94.

19. Hamed, KA; Dormitzer, PR; Su, CK; Relman, DA; *Haemophilus parainfluenzae endocarditis: application of a molecular approach for identification of pathogenic bacterial species.* Clin Infect Dis, **1994**. 19(4): p. 677-83.

20. Nwaohiri, N; Urban, C; Gluck, J; Ahluwalia, M; Wehbeh, W; *Tricuspid valve endocarditis caused by Haemophilus parainfluenzae: a case report and review of the literature.* Diagn Microbiol Infect Dis, **2009**. 64(2): p. 216-9.

21. Jellicoe, PA; Cohen, A; Campbell, P; *Haemophilus parainfluenzae complicating total hip arthroplasty: a rapid failure.* J Arthroplasty, **2002**. 17(1): p. 114-6.

22. Betriu, C; Coronel, F; Martin, P; Picazo, JJ; *Peritonitis caused by Haemophilus parainfluenzae in a patient undergoing continuous ambulatory peritoneal dialysis.* J Clin Microbiol, **1999**. 37(9): p. 3074-5.

23. Pillai, A; Mitchell, JL; Hill, SL; Stockley, RA; *A case of Haemophilus parainfluenzae pneumonia.* Thorax, **2000**. 55(7): p. 623-4.

24. Frankard, J; Rodriguez-Villalobos, H; Struelens, MJ; Jacobs, F; *Haemophilus parainfluenzae: an underdiagnosed pathogen of biliary tract infections?* Eur J Clin Microbiol Infect Dis, **2004**. 23(1): p. 46-8.

25. Lau, SK; Woo, PC; Chan, BY; Fung, AM; Que, TL; Yuen, KY; *Haemophilus segnis polymicrobial and monomicrobial bacteraemia identified by 16S ribosomal RNA gene sequencing.* J Med Microbiol, **2002**. 51(8): p. 635-40.

26. Helander, IM; Lindner, B; Brade, H; Altmann, K; Lindberg, AA; Rietschel, ET; Zahringer, U; *Chemical structure of the lipopolysaccharide of Haemophilus influenzae strain I-69 Rd-/b+. Description of a novel deep-rough chemotype.* Eur J Biochem, **1988**. 177(3): p. 483-92.

27. Mikhail, I; Yildirim, HH; Lindahl, EC; Schweda, EK; *Structural characterization of lipid A from nontypeable and type f Haemophilus influenzae: variability of fatty acid substitution.* Anal Biochem, **2005**. 340(2): p. 303-16.

28. Inzana, TJ; Seifert, WE, Jr.; Williams, RP; *Composition and antigenic activity of the oligosaccharide moiety of Haemophilus influenzae type b lipooligosaccharide.* Infect Immun, **1985**. 48(2): p. 324-30.

29. Zamze, SE; Ferguson, MA; Moxon, ER; Dwek, RA; Rademacher, TW; *Identification of phosphorylated 3-deoxy-manno-octulosonic acid as a component of Haemophilus influenzae lipopolysaccharide.* Biochem J, **1987**. 245(2): p. 583-7.

30. Coleman, WG, Jr.; *The rfaD gene codes for ADP-L-glycero-D-manno-heptose-6-epimerase. An enzyme required for lipopolysaccharide core biosynthesis.* J Biol Chem, **1983**. 258(3): p. 1985-90.

31. Schweda, EK; Richards, JC; Hood, DW; Moxon, ER; *Expression and structural diversity of the lipopolysaccharide of Haemophilus influenzae: implication in virulence*, in *Int J Med Microbiol*. **2007**. p. 297-306.

32. Månsson, M; *The structural diversity of lipopolysaccharides expressed by genetically defined clinical isolates of Haemophilus influenzae*, in *Laboratory Medicine*. **2003**, Karolinska Institutet: Stockholm, Sweden. p. 52.

33. Virji, M; Weiser, JN; Lindberg, AA; Moxon, ER; *Antigenic similarities in lipopolysaccharides of Haemophilus and Neisseria and expression of a digalactoside structure also present on human cells.* Microb Pathog, **1990**. 9(6): p. 441-50.

34. Marti-Lliteras, P; Lopez-Gomez, A; Mauro, S; Hood, DW; Viadas, C; Calatayud, L; Morey, P; Servin, A; Linares, J; Oliver, A; Bengoechea, JA; Garmendia, J; *Nontypable Haemophilus influenzae Displays a Prevalent Surface Structure Molecular Pattern in Clinical Isolates.* PLoS One, **2011**. 6(6): p. e21133.

35. Griffin, R; Cox, AD; Makepeace, K; Richards, JC; Moxon, ER; Hood, DW; *The role of lex2 in lipopolysaccharide biosynthesis in Haemophilus influenzae strains RM7004 and RM153.* Microbiology, **2003**. 149(Pt 11): p. 3165-75.

36. Moxon, R; Bayliss, C; Hood, D; *Bacterial contingency loci: the role of simple sequence DNA repeats in bacterial adaptation.* Annu Rev Genet, **2006**. 40: p. 307-33.

37. Appelmelk, BJ; Shiberu, B; Trinks, C; Tapsi, N; Zheng, PY; Verboom, T; Maaskant, J; Hokke, CH; Schiphorst, WE; Blanchard, D; Simoons-Smit, IM; van den Eijnden, DH; Vandenbroucke-Grauls, CM; *Phase variation in Helicobacter pylori lipopolysaccharide.* Infect Immun, **1998**. 66(1): p. 70-6.

38. Tettelin, H; Saunders, NJ; Heidelberg, J; Jeffries, AC; Nelson, KE; Eisen, JA; Ketchum, KA; Hood, DW; Peden, JF; Dodson, RJ; Nelson, WC; Gwinn, ML; DeBoy, R; Peterson, JD; Hickey, EK; Haft, DH; Salzberg, SL; White, O; Fleischmann, RD; Dougherty, BA; Mason, T; Ciecko, A; Parksey, DS; Blair, E; Cittone, H; Clark, EB; Cotton, MD; Utterback, TR; Khouri, H; Qin, H; Vamathevan, J; Gill, J; Scarlato, V; Masignani, V; Pizza, M; Grandi, G; Sun, L; Smith, HO; Fraser, CM; Moxon, ER; Rappuoli, R; Venter, JC; *Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.* Science, **2000**. 287(5459): p. 1809-15.

39. Inzana, TJ; Hensley, J; McQuiston, J; Lesse, AJ; Campagnari, AA; Boyle, SM; Apicella, MA; *Phase variation and conservation of lipooligosaccharide epitopes in Haemophilus somnus.* Infect Immun, **1997**. 65(11): p. 4675-81.

40. Fleischmann, RD; Adams, MD; White, O; Clayton, RA; Kirkness, EF; Kerlavage, AR; Bult, CJ; Tomb, JF; Dougherty, BA; Merrick, JM; et al.; *Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.* Science, **1995**. 269(5223): p. 496-512.

41. Erwin, AL; Nelson, KL; Mhlanga-Mutangadura, T; Bonthuis, PJ; Geelhood, JL; Morlin, G; Unrath, WC; Campos, J; Crook, DW; Farley, MM; Henderson, FW; Jacobs, RF; Muhlemann, K; Satola, SW; van Alphen, L; Golomb, M; Smith, AL; *Characterization of genetic and phenotypic diversity of invasive nontypeable Haemophilus influenzae.* Infect Immun, **2005**. 73(9): p. 5853-63.

42. Harrison, A; Dyer, DW; Gillaspy, A; Ray, WC; Mungur, R; Carson, MB; Zhong, H; Gipson, J; Gipson, M; Johnson, LS; Lewis, L; Bakaletz, LO; Munson, RS, Jr.; *Genomic sequence of an otitis media isolate of nontypeable Haemophilus influenzae: comparative study with H. influenzae serotype d, strain KW20.* J Bacteriol, **2005**. 187(13): p. 4627-36.

43. Cody, AJ; Field, D; Feil, EJ; Stringer, S; Deadman, ME; Tsolaki, AG; Gratz, B; Bouchet, V; Goldstein, R; Hood, DW; Moxon, ER; *High rates of recombination in otitis media isolates of non-typeable Haemophilus influenzae.* Infect Genet Evol, **2003**. 3(1): p. 57-66.

44. Moxon, ER; Hood, DW; Saunders, NJ; Schweda, EK; Richards, JC; *Functional genomics of pathogenic bacteria.* Philos Trans R Soc Lond B Biol Sci, **2002**. 357(1417): p. 109-16.

45. Hood, DW; Cox, AD; Wakarchuk, WW; Schur, M; Schweda, EK; Walsh, SL; Deadman, ME; Martin, A; Moxon, ER; Richards, JC; *Genetic basis for expression of the major globotetraose-containing lipopolysaccharide from H. influenzae strain Rd (RM118).* Glycobiology, **2001**. 11(11): p. 957-67.

46. Engskog, MK; Deadman, M; Li, J; Hood, DW; Schweda, EK; *Detailed structural features of lipopolysaccharide glycoforms in nontypeable Haemophilus influenzae strain 2019.* Carbohydr Res, **2011**. 346(10): p. 1241-9.

47. Deadman, ME; Hermant, P; Engskog, M; Makepeace, K; Moxon, ER; Schweda, EK; Hood, DW; *Lex2B, a phase-variable glycosyltransferase, adds either a glucose or a galactose to Haemophilus influenzae lipopolysaccharide.* Infect Immun, **2009**. 77(6): p. 2376-84.

48. Engskog, MK; Yildirim, HH; Li, J; Richards, JC; Deadman, M; Hood, DW; Schweda, EK; *A dual role for the lex2 locus: identification of galactosyltransferase activity in non-typeable Haemophilus influenzae strains 1124 and 2019.* Carbohydr Res, **2009**. 344(5): p. 632-41.

49. Yildirim, HH; Li, J; Richards, JC; Hood, DW; Moxon, ER; Schweda, EK; *An alternate pattern for globoside oligosaccharide expression in Haemophilus influenzae lipopolysaccharide: structural diversity in nontypeable strain 1124.* Biochemistry, **2005**. 44(13): p. 5207-24.

50. Erwin, AL; Bonthuis, PJ; Geelhood, JL; Nelson, KL; McCrea, KW; Gilsdorf, JR; Smith, AL; *Heterogeneity in tandem octanucleotides within Haemophilus influenzae lipopolysaccharide biosynthetic gene losA affects serum resistance.* Infect Immun, **2006**. 74(6): p. 3408-14.

51. Hood, DW; Deadman, ME; Engskog, MK; Vitiazeva, V; Makepeace, K; Schweda, EK; Moxon, R; *Genes required for the synthesis of heptose-containing oligosaccharide outer core extensions in Haemophilus influenzae lipopolysaccharide.* Microbiology, **2010**. 156(Pt 11): p. 3421-31.

52. Månsson, M; Hood, DW; Moxon, ER; Schweda, EK; *Structural diversity in lipopolysaccharide expression in nontypeable Haemophilus influenzae. Identification of L-glycerol-D-manno-heptose in the outer-core region in three clinical isolates.* Eur J Biochem, **2003**. 270(4): p. 610-24.

53. Rahman, MM; Gu, XX; Tsai, CM; Kolli, VS; Carlson, RW; *The structural heterogeneity of the lipooligosaccharide (LOS) expressed by pathogenic non-typeable Haemophilus influenzae strain NTHi 9274.* Glycobiology, **1999**. 9(12): p. 1371-80.

54. Månsson, M; Hood, DW; Moxon, ER; Schweda, EK; *Structural characterization of a novel branching pattern in the lipopolysaccharide from nontypeable Haemophilus influenzae.* Eur J Biochem, **2003**. 270(14): p. 2979-91.

55. Lundström, SL; Li, J; Deadman, ME; Hood, DW; Moxon, ER; Schweda, EK; *Structural analysis of the lipopolysaccharide from nontypeable Haemophilus influenzae strain R2846.* Biochemistry, **2008**. 47(22): p. 6025-38.

56. Schweda, EK; Jonasson, JA; Jansson, PE; *Structural studies of lipooligosaccharides from Haemophilus ducreyi ITM 5535, ITM 3147, and a fresh clinical isolate, ACY1: evidence for intrastrain heterogeneity with the production of mutually exclusive sialylated or elongated glycoforms.* J Bacteriol, **1995**. 177(18): p. 5316-21.

57. Pollard, A; St Michael, F; Connor, L; Nichols, W; Cox, A; *Structural characterization of Haemophilus parainfluenzae lipooligosaccharide and elucidation of its role in adherence using an outer core mutant.* Can J Microbiol, **2008**. 54(11): p. 906-17.

58. Hood, DW; Randle, G; Cox, AD; Makepeace, K; Li, J; Schweda, EK; Richards, JC; Moxon, ER; *Biosynthesis of cryptic lipopolysaccharide glycoforms in*

Haemophilus influenzae involves a mechanism similar to that required for O-antigen synthesis. J Bacteriol, **2004**. 186(21): p. 7429-39.

59. Månsson, M; Bauer, SH; Hood, DW; Richards, JC; Moxon, ER; Schweda, EK; *A new structural type for Haemophilus influenzae lipopolysaccharide. Structural analysis of the lipopolysaccharide from nontypeable Haemophilus influenzae strain 486.* Eur J Biochem, **2001**. 268(7): p. 2148-59.

60. Lundström, SL; Twelkmeyer, B; Sagemark, MK; Li, J; Richards, JC; Hood, DW; Moxon, ER; Schweda, EK; *Novel globoside-like oligosaccharide expression patterns in nontypeable Haemophilus influenzae lipopolysaccharide.* FEBS J, **2007**. 274(18): p. 4886-903.

61. Masoud, H; Moxon, ER; Martin, A; Krajcarski, D; Richards, JC; *Structure of the variable and conserved lipopolysaccharide oligosaccharide epitopes expressed by Haemophilus influenzae serotype b strain Eagan.* Biochemistry, **1997**. 36(8): p. 2091-103.

62. Masoud, H; Martin, A; Thibault, P; Moxon, ER; Richards, JC; *Structure of extended lipopolysaccharide glycoforms containing two globotriose units in Haemophilus influenzae serotype b strain RM7004.* Biochemistry, **2003**. 42(15): p. 4463-75.

63. Masoud, H; Moxon, ER; Richards, JC; *Structural elucidation of the major Hex4 lipopolysaccharide glycoform from the lgtC mutant of Haemophilus influenzae strain Eagan.* Carbohydr Res, **2008**. 343(9): p. 1424-34.

64. Deadman, ME; Lundström, SL; Schweda, EK; Moxon, ER; Hood, DW; *Specific amino acids of the glycosyltransferase LpsA direct the addition of glucose or galactose to the terminal inner core heptose of Haemophilus influenzae lipopolysaccharide via alternative linkages.* J Biol Chem, **2006**. 281(40): p. 29455-67.

65. Risberg, A; Masoud, H; Martin, A; Richards, JC; Moxon, ER; Schweda, EK; *Structural analysis of the lipopolysaccharide oligosaccharide epitopes expressed by a capsule-deficient strain of Haemophilus influenzae Rd.* Eur J Biochem, **1999**. 261(1): p. 171-80.

66. Schweda, EK; Li, J; Moxon, ER; Richards, JC; *Structural analysis of lipopolysaccharide oligosaccharide epitopes expressed by non-typeable Haemophilus influenzae strain 176.* Carbohydr Res, **2002**. 337(5): p. 409-20.

67. Yildirim, HH; Li, J; Richards, JC; Hood, DW; Moxon, ER; Schweda, EK; *Complex O-acetylation in non-typeable Haemophilus influenzae lipopolysaccharide: evidence for a novel site of O-acetylation.* Carbohydr Res, **2005**. 340(17): p. 2598-611.

68. Landerholm, MK; Li, J; Richards, JC; Hood, DW; Moxon, ER; Schweda, EK; *Characterization of novel structural features in the lipopolysaccharide of nondisease associated nontypeable Haemophilus influenzae.* Eur J Biochem, **2004**. 271(5): p. 941-53.

69. Schweda, EK; Brisson, JR; Alvelius, G; Martin, A; Weiser, JN; Hood, DW; Moxon, ER; Richards, JC; *Characterization of the phosphocholine-substituted oligosaccharide in lipopolysaccharides of type b Haemophilus influenzae.* Eur J Biochem, **2000**. 267(12): p. 3902-13.

70. Weiser, JN; Lindberg, AA; Manning, EJ; Hansen, EJ; Moxon, ER; *Identification of a chromosomal locus for expression of lipopolysaccharide epitopes in Haemophilus influenzae.* Infect Immun, **1989**. 57(10): p. 3045-52.

71. Lysenko, E; Richards, JC; Cox, AD; Stewart, A; Martin, A; Kapoor, M; Weiser, JN; *The position of phosphorylcholine on the lipopolysaccharide of Haemophilus influenzae affects binding and sensitivity to C-reactive protein-mediated killing.* Mol Microbiol, **2000**. 35(1): p. 234-45.

72. Weiser, JN; Shchepetov, M; Chong, ST; *Decoration of lipopolysaccharide with phosphorylcholine: a phase-variable characteristic of Haemophilus influenzae.* Infect Immun, **1997**. 65(3): p. 943-50.

73. Fox, KL; Li, J; Schweda, EK; Vitiazeva, V; Makepeace, K; Jennings, MP; Moxon, ER; Hood, DW; *Duplicate copies of lic1 direct the addition of multiple phosphocholine residues in the lipopolysaccharide of Haemophilus influenzae.* Infect Immun, **2008**. 76(2): p. 588-600.

74. Weiser, JN Pan, N; *Adaptation of Haemophilus influenzae to acquired and innate humoral immunity based on phase variation of lipopolysaccharide.* Mol Microbiol, **1998**. 30(4): p. 767-75.

75. Weiser, JN; Pan, N; McGowan, KL; Musher, D; Martin, A; Richards, J; *Phosphorylcholine on the lipopolysaccharide of Haemophilus influenzae contributes to persistence in the respiratory tract and sensitivity to serum killing mediated by C-reactive protein.* J Exp Med, **1998**. 187(4): p. 631-40.

76. Serino, L Virji, M; *Phosphorylcholine decoration of lipopolysaccharide differentiates commensal Neisseriae from pathogenic strains: identification of licA-type genes in commensal Neisseriae.* Mol Microbiol, **2000**. 35(6): p. 1550-9.

77. Kim, JO Weiser, JN; *Association of intrastrain phase variation in quantity of capsular polysaccharide and teichoic acid with the virulence of Streptococcus pneumoniae.* J Infect Dis, **1998**. 177(2): p. 368-77.

78. Weiser, JN; Goldberg, JB; Pan, N; Wilson, L; Virji, M; *The phosphorylcholine epitope undergoes phase variation on a 43-kilodalton protein in Pseudomonas aeruginosa and on pili of Neisseria meningitidis and Neisseria gonorrhoeae.* Infect Immun, **1998**. 66(9): p. 4263-7.

79. Bauer, SH; Månsson, M; Hood, DW; Richards, JC; Moxon, ER; Schweda, EK; *A rapid and sensitive procedure for determination of 5-N-acetyl neuraminic acid in lipopolysaccharides of Haemophilus influenzae: a survey of 24 non-typeable H. influenzae strains.* Carbohydr Res, **2001**. 335(4): p. 251-60.

80. Vimr, E; Lichtensteiger, C; Steenbergen, S; *Sialic acid metabolism's dual function in Haemophilus influenzae.* Mol Microbiol, **2000**. 36(5): p. 1113-23.

81. Hood, DW; Cox, AD; Gilbert, M; Makepeace, K; Walsh, S; Deadman, ME; Cody, A; Martin, A; Månsson, M; Schweda, EK; Brisson, JR; Richards, JC; Moxon, ER; Wakarchuk, WW; *Identification of a lipopolysaccharide alpha-2,3-sialyltransferase from Haemophilus influenzae.* Mol Microbiol, **2001**. 39(2): p. 341-50.

82. Fox, KL; Cox, AD; Gilbert, M; Wakarchuk, WW; Li, J; Makepeace, K; Richards, JC; Moxon, ER; Hood, DW; *Identification of a bifunctional lipopolysaccharide sialyltransferase in Haemophilus influenzae: incorporation of disialic acid.* J Biol Chem, **2006**. 281(52): p. 40024-32.

83. Hood, DW; Makepeace, K; Deadman, ME; Rest, RF; Thibault, P; Martin, A; Richards, JC; Moxon, ER; *Sialic acid in the lipopolysaccharide of Haemophilus influenzae: strain distribution, influence on serum resistance and structural characterization.* Mol Microbiol, **1999**. 33(4): p. 679-92.

84. Bouchet, V; Hood, DW; Li, J; Brisson, JR; Randle, GA; Martin, A; Li, Z; Goldstein, R; Schweda, EK; Pelton, SI; Richards, JC; Moxon, ER; *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, **2003**. 100(15): p. 8898-903.

85. Lundström, SL; Li, J; Månsson, M; Figueira, M; Leroy, M; Goldstein, R; Hood, DW; Moxon, ER; Richards, JC; Schweda, EK; *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, **2008**. 76(7): p. 3255-67.

86. Tinnert, AS; Månsson, M; Yildirim, HH; Hood, DW; Schweda, EK; *Structural investigation of lipopolysaccharides from nontypeable Haemophilus influenzae: investigation of inner-core phosphoethanolamine addition in NTHi strain 981.* Carbohydr Res, **2005**. 340(11): p. 1900-7.

87. Fox, KL; Yildirim, HH; Deadman, ME; Schweda, EK; Moxon, ER; Hood, DW; *Novel lipopolysaccharide biosynthetic genes containing tetranucleotide repeats in Haemophilus influenzae, identification of a gene for adding O-acetyl groups.* Mol Microbiol, **2005**. 58(1): p. 207-16.

88. Li, J; Bauer, SH; Månsson, M; Moxon, ER; Richards, JC; Schweda, EK; *Glycine is a common substituent of the inner core in Haemophilus influenzae lipopolysaccharide.* Glycobiology, **2001**. 11(12): p. 1009-15.

89. Roberts, MC; Mintz, CS; Morse, SA; *Characterization of Haemophilus parainfluenzae strains with low-Mr or ladder-like lipopolysaccharides.* J Gen Microbiol, **1986**. 132(3): p. 611-6.

90. Zhou, D; Stephens, DS; Gibson, BW; Engstrom, JJ; McAllister, CF; Lee, FK; Apicella, MA; *Lipooligosaccharide biosynthesis in pathogenic Neisseria. Cloning, identification, and characterization of the phosphoglucomutase gene.* J Biol Chem, **1994**. 269(15): p. 11162-9.

91. Swords, WE; Buscher, BA; Ver Steeg Ii, K; Preston, A; Nichols, WA; Weiser, JN; Gibson, BW; Apicella, MA; *Non-typeable Haemophilus influenzae adhere to and invade human bronchial epithelial cells via an interaction of lipooligosaccharide with the PAF receptor.* Mol Microbiol, **2000**. 37(1): p. 13-27.

92. Vitiazeva, V; Twelkmeyer, B; Young, R; Hood, DW; Schweda, EK; *Structural studies of the lipopolysaccharide from Haemophilus parainfluenzae strain 20.* Carbohydr Res, **2011**.

93. Ruiz, N; Kahne, D; Silhavy, TJ; *Transport of lipopolysaccharide across the cell envelope: the long road of discovery.* Nat Rev Microbiol, **2009**. 7(9): p. 677-83.

94. Whitfield, C; *Biosynthesis and assembly of capsular polysaccharides in Escherichia coli.* Annu Rev Biochem, **2006**. 75: p. 39-68.

95. Cuthbertson, L; Kos, V; Whitfield, C; *ABC transporters involved in export of cell surface glycoconjugates.* Microbiol Mol Biol Rev, **2010**. 74(3): p. 341-62.

96. Wang, L; Liu, D; Reeves, PR; *C-terminal half of Salmonella enterica WbaP (RfbP) is the galactosyl-1-phosphate transferase domain catalyzing the first step of O-antigen synthesis.* J Bacteriol, **1996**. 178(9): p. 2598-604.

97. Lehrer, J; Vigeant, KA; Tatar, LD; Valvano, MA; *Functional characterization and membrane topology of Escherichia coli WecA, a sugar-phosphate transferase initiating the biosynthesis of enterobacterial common antigen and O-antigen lipopolysaccharide.* J Bacteriol, **2007**. 189(7): p. 2618-28.

98. Hood, DW; Deadman, ME; Allen, T; Masoud, H; Martin, A; Brisson, JR; Fleischmann, R; Venter, JC; Richards, JC; Moxon, ER; *Use of the complete genome sequence information of Haemophilus influenzae strain Rd to investigate lipopolysaccharide biosynthesis.* Mol Microbiol, **1996**. 22(5): p. 951-65.

99. Herriott, RM; Meyer, EM; Vogt, M; *Defined nongrowth media for stage II development of competence in Haemophilus influenzae.* J Bacteriol, **1970**. 101(2): p. 517-24.

100. Galanos, C; Lüderitz, O; Westphal, O; *A new method for the extraction of R lipopolysaccharides.* Eur J Biochem, **1969**. 9(2): p. 245-9.

101. Westphal, O, Jann, K. ; *Bacterial lipopolysaccharides extraction with phenol-water and further applications of the procedure.* Methods in Carbohydrate Chemistry, **1965**. 5: p. 83-91.

102. Dzieciatkowska, M; Schweda, EK; Moxon, ER; Richards, JC; Li, J; *Characterization of intact lipopolysaccharides from the Haemophilus influenzae strain RM 118 using electrophoresis-assisted open-tubular liquid chromatography-mass spectrometry.* Electrophoresis, **2008**. 29(10): p. 2171-81.

103. Holst, O; Brade, L; Kosma, P; Brade, H; *Structure, serological specificity, and synthesis of artificial glycoconjugates representing the genus-specific lipopolysaccharide epitope of Chlamydia spp.* J Bacteriol, **1991**. 173(6): p. 1862-6.

104. Danan. A., M, M., Sarfati, S.R., Szabó, P.; *Synthesis and Behaviour under Acidic Conditions of 2-Deoxy-D-arabino-hexopyranose and 3-Deoxy-2-ketoaldonic Acids Bearing O-Phosphono or O-Glycosyl Substitutents at Positions β to the Carbonyl Function.* Journal of the Chemical Society. Perkin transactions. I, **1982**: p. 1272-1282.

105. Auzanneau, F-I, Charon, D., Szabó, L.; *Phosphorylated Sugars. Part 27. Synthesis and Reactions, in acidic medium, of 5-O-substituted Methyl 3-Deoxy-α-D-manno-oct-2-ulopyranosidonic Acid 4-Phosphates.* Journal of the Chemical Society, Perkin transactions, I, **1991**: p. 509-517.

106. Schweda, EK; Hegedus, OE; Borrelli, S; Lindberg, AA; Weiser, JN; Maskell, DJ; Moxon, ER; *Structural studies of the saccharide part of the cell envelope lipopolysaccharide from Haemophilus influenzae strain AH1-3 (lic3+).* Carbohydr Res, **1993**. 246: p. 319-30.

107. Mawhinney, TP Chance, DL; *Hydrolysis of sialic acids and O-acetylated sialic acids with propionic acid.* Anal Biochem, **1994**. 223(1): p. 164-7.

108. Kamerling, JP; Schauer, R; Shukla, AK; Stoll, S; Van Halbeek, H; Vliegenthart, JF; *Migration of O-acetyl groups in N,O-acetylneuraminic acids.* Eur J Biochem, **1987**. 162(3): p. 601-7.

109. Blakeney, AB Stone, BA; *Methylation of carbohydrates with lithium methylsulphinyl carbanion.* Carbohydr Res, **1985**. 140(2): p. 319-24.

110. Hakomori, S; *A Rapid Permethylation of Glycolipid, and Polysaccharide Catalyzed by Methylsulfinyl Carbanion in Dimethyl Sulfoxide.* J Biochem, **1964**. 55: p. 205-8.

111. Prehm, P; *Methylation of carbohydrates by methyl trifluoromethanesulfonate in trimethyl phosphate.* Carbohydrate Research, **1980**. 78(2): p. 372-374.

112. Jawahar S. Sawardeker, JHS, Allene Jeanes; *Quantitative Determination of Monosaccharides as Their Alditol Acetates by Gas Liquid Chromatography.* Analytical Chemistry, **1965**. 37(12): p. 1602-1604.

113. Weissbach, A Hurwitz, J; *The formation of 2-keto-3-deoxyheptonic acid in extracts of Escherichia coli B. I. Identification.* J Biol Chem, **1959**. 234(4): p. 705-9.

114. Gerwig, GJ; Kamerling, JP; Vliegenthart, JF; *Determination of the absolute configuration of mono-saccharides in complex carbohydrates by capillary G.L.C.* Carbohydr Res, **1979**. 77: p. 10-7.

115. Schweda, EK Richards, JC; *Profiling LPS glycoforms of non-typeable Haemophilus influenzae by multiple-stage tandem mass spectrometry.* Methods Mol Biol, **2010**. 600: p. 79-92.

116. Ruhaak, LR; Deelder, AM; Wuhrer, M; *Oligosaccharide analysis by graphitized carbon liquid chromatography-mass spectrometry.* Anal Bioanal Chem, **2009**. 394(1): p. 163-74.

117. Hashii, N; Kawasaki, N; Itoh, S; Harazono, A; Matsuishi, Y; Hayakawa, T; Kawanishi, T; *Specific detection of Lewis x-carbohydrates in biological samples*

using liquid chromatography/multiple-stage tandem mass spectrometry. Rapid Commun Mass Spectrom, **2005**. 19(22): p. 3315-21.

118. Wuhrer, M; Deelder, AM; van der Burgt, YE; *Mass spectrometric glycan rearrangements.* Mass Spectrom Rev, **2011**. 30(4): p. 664-80.

119. Brüll, LP; Kovacik, V; Thomas-Oates, JE; Heerma, W; Haverkamp, J; *Sodium-cationized oligosaccharides do not appear to undergo 'internal residue loss' rearrangement processes on tandem mass spectrometry.* Rapid Commun Mass Spectrom, **1998**. 12(20): p. 1520-32.

120. Schweda, EK Richards, JC; *Structural profiling of short-chain lipopolysaccharides from Haemophilus influenzae.* Methods Mol Med, **2003**. 71: p. 161-83.

121. Li, J; Cox, AD; Hood, DW; Schweda, EK; Moxon, ER; Richards, JC; *Electrophoretic and mass spectrometric strategies for profiling bacterial lipopolysaccharides.* Mol Biosyst, **2005**. 1(1): p. 46-52.

122. Li, J; Dzieciatkowska, M; Hood, DW; Cox, AD; Schweda, EK; Moxon, ER; Richards, JC; *Structural characterization of sialylated glycoforms of H. influenzae by electrospray mass spectrometry: fragmentation of protonated and sodiated O-deacylated lipopolysaccharides.* Rapid Commun Mass Spectrom, **2007**. 21(6): p. 952-60.

123. Dzieciatkowska, M; Brochu, D; van Belkum, A; Heikema, AP; Yuki, N; Houliston, RS; Richards, JC; Gilbert, M; Li, J; *Mass spectrometric analysis of intact lipooligosaccharide: direct evidence for O-acetylated sialic acids and discovery of O-linked glycine expressed by Campylobacter jejuni.* Biochemistry, **2007**. 46(50): p. 14704-14.

124. Schweda, EK; Twelkmeyer, B; Li, J; *Profiling structural elements of short-chain lipopolysaccharide of non-typeable Haemophilus influenzae.* Innate Immun, **2008**. 14(4): p. 199-211.

125. Keeler, J; *Understanding NMR Spectroscopy.* 1 ed. **2005**: John Wiley and Sons Ltd.

126. Duus, J; Gotfredsen, CH; Bock, K; *Carbohydrate structural determination by NMR spectroscopy: modern methods and limitations.* Chem Rev, **2000**. 100(12): p. 4589-614.

127. Pettigrew, MM; Foxman, B; Marrs, CF; Gilsdorf, JR; *Identification of the lipooligosaccharide biosynthesis gene lic2B as a putative virulence factor in strains of nontypeable Haemophilus influenzae that cause otitis media.* Infect Immun, **2002**. 70(7): p. 3551-6.

128. Månsson, M; Hood, DW; Li, J; Richards, JC; Moxon, ER; Schweda, EK; *Structural analysis of the lipopolysaccharide from nontypeable Haemophilus influenzae strain 1003.* Eur J Biochem, **2002**. 269(3): p. 808-18.

129. Wong, SM; St Michael, F; Cox, A; Ram, S; Akerley, BJ; *ArcA-regulated glycosyltransferase lic2B promotes complement evasion and pathogenesis of nontypeable Haemophilus influenzae.* Infect Immun, **2011**. 79(5): p. 1971-83.

130. Griffin, R; Bayliss, CD; Herbert, MA; Cox, AD; Makepeace, K; Richards, JC; Hood, DW; Moxon, ER; *Digalactoside expression in the lipopolysaccharide of Haemophilus influenzae and its role in intravascular survival.* Infect Immun, **2005**. 73(10): p. 7022-6.

131. Podolsky, DK; *Oligosaccharide structures of human colonic mucin.* J Biol Chem, **1985**. 260(14): p. 8262-71.

132. Podolsky, DK; *Oligosaccharide structures of isolated human colonic mucin species.* J Biol Chem, **1985**. 260(29): p. 15510-5.

133. Toba, S; Tenno, M; Kurosaka, A; *An O-acetylated sialyl-Tn is involved in ovarian cancer-associated antigenicity.* Biochem Biophys Res Commun, **2000**. 271(2): p. 281-6.

134. Wakarchuk, WW; Gilbert, M; Martin, A; Wu, Y; Brisson, JR; Thibault, P; Richards, JC; *Structure of an alpha-2,6-sialylated lipooligosaccharide from Neisseria meningitidis immunotype L1.* Eur J Biochem, **1998**. 254(3): p. 626-33.

135. Shakhnovich, EA; King, SJ; Weiser, JN; *Neuraminidase expressed by Streptococcus pneumoniae desialylates the lipopolysaccharide of Neisseria meningitidis and Haemophilus influenzae: a paradigm for interbacterial competition among pathogens of the human respiratory tract.* Infect Immun, **2002**. 70(12): p. 7161-4.

136. De Castro, C; Parrilli, M; Holst, O; Molinaro, A; *Microbe-associated molecular patterns in innate immunity: Extraction and chemical analysis of gram-negative bacterial lipopolysaccharides.* Methods Enzymol, **2010**. 480: p. 89-115.